Nucleic Acids Research, 2025, **53**, gkaf110
https://doi.org/10.1093/nar/gkaf110
Molecular and Structural Biology

Structural basis of ssDNA-guided NADase activation of prokaryotic SPARTA system

Rong Hu, Chenmin Guo, Xiaotian Liu, Yuanyi Lin, Zixin Yang, Zhixin Li, Ye Yang, Erman Ma, Yiyi Li, Jiyun Chen*, Liang Liu *

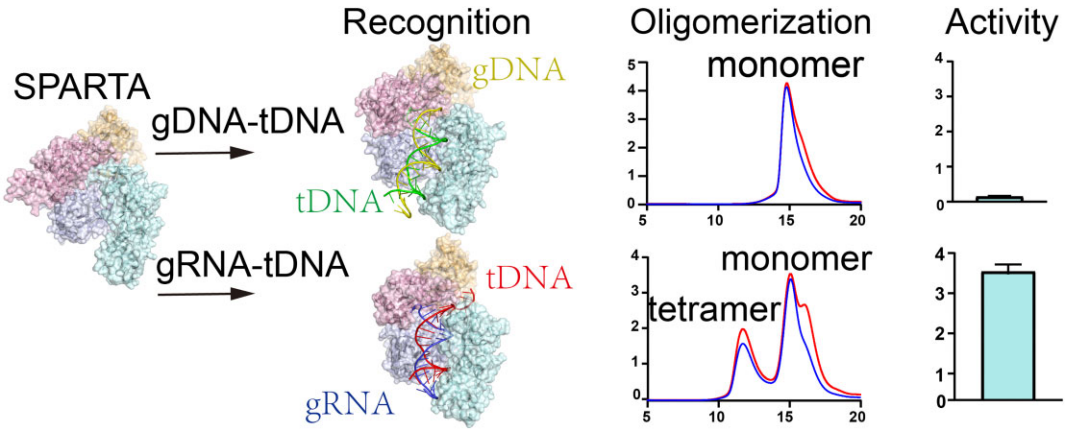
State Key Laboratory of Cellular Stress Biology, School of Life Sciences, Faculty of Medicine and Life Sciences, Xiamen University, No. 4221, Xiang'an South Road, Xiamen 361102, China

*To whom correspondence should be addressed. Email: liangliu2019@xmu.edu.cn
Correspondence may also be addressed to Jiyun Chen. Email: chenjiyun@xmu.edu.cn

Abstract

Short prokaryotic Argonaute and the associated TIR-APAZ (SPARTA) proteins constitute a prokaryotic immune system, mediating RNA- or DNA-guided target single-stranded DNA (ssDNA) to activate NADase activity and induce cell death by degrading NAD^+ in response to invading plasmids. Although the guide RNA-mediated targeting mechanism of SPARTA has been established, the functional role and mechanisms of guide DNA-mediated SPARTA remain poorly understood. Here, we report two crystal structures of *Crenotalea thermophila* SPARTA complexes with 5′-phosphorylated 21-nt guide DNA and complementary target ssDNA lengths of 15 or 20 nt. The structures demonstrate specific recognition of the 5′-OH or 3′-OH groups in target DNA by SPARTA, while not recognizing the 5′-P group in guide DNA. This suggests distinct recognition models for guide DNA and guide RNA, indicating different activation mechanisms. Furthermore, these two structures reveal disparate models for recognizing guide DNA and target DNA, providing insights into the length requirement for SPARTA activation.

Graphical abstract



Introduction

Argonaute proteins recruit guide RNA or DNA (gRNA, gDNA) to bind or cleave target RNA or DNA (tRNA, tDNA) in both eukaryotes and prokaryotes [1–6]. Eukaryotic Ago proteins (eAgos) recruit short RNA molecules to recognize and cleave target RNA, which is the core process of RNA interference [1, 3, 7, 8]. In prokaryotes, bacteria possess a homologous eAgo protein called pAgo that serves diverse functions as an antiviral defense system against bacteriophage infections and nucleic acid invasions [9–16]. pAgo can be classified into three clades: long-A pAgo, long-B pAgo, and short pAgo clades. Long pAgo shares conserved N (N-terminal), PAZ (PIWI-Argonaute-Zwille), MID (middle), and PIWI (P-element-induced wimpy testis) domains with eAgo [17–19]. Short pAgo only contains the MID and PIWI domains while lacking an N-terminal region and PAZ domain responsible for forming and closing DNA binding channels [20].

An APAZ domain, which shares homology with the N domain of long pAgos, is encoded within the same operon as the short pAgo [18]. Typically, the APAZ domain is commonly fused to domains such as TIR (Toll–interleukin-1

Received: September 2, 2024. **Revised:** January 17, 2025. **Editorial Decision:** January 31, 2025. **Accepted:** February 5, 2025

receptor), SIR2 (silent information regulator 2), or DUF4365 [16, 17, 21]. Recently, studies have demonstrated that short pAgo proteins associated with TIR-APAZ or SIR2-APAZ proteins form complexes with single-strand (ss) DNA or RNA guides and ssDNA targets, constituting a prokaryotic immune system known as SPARTA (pAgo-TIR-APAZ) or SPARSA (pAgo-SIR2-APAZ) [13, 20, 22]. The activation of SPARTA or SPARSA can attenuate the viability of DNA-invaded cells and induce their demise by depleting $NAD(P)^+$, thereby safeguarding the uninvaded cells. Several recent studies have shed light on the activation mechanism and structural basis underlying the RNA-guided and invading DNA-targeted SPARTA and SPARSA phage defense systems [21, 23–29]. It has been observed that SPARTA can assemble into an active tetramer while also existing in inactive monomeric and dimeric states. The critical steps of SPARTA activation involve the interaction between Ago–Ago dimers induced by gRNA–tDNA duplex recognition, as well as the rearrangement and packaging of the TIR domain to form NADase active sites. In contrast to SPARTA, SPARSA exhibits a distinct activation mechanism dependent on conformational changes in its SIR2 domain, particularly involving a loop shift away from NAD^+. This allows water molecules to undergo nucleophilic attack on C1′ of NAD^+ for NADase activation [21].

Although RNA-guided DNA target binding by Argonautes is more commonly found in bacteria [30, 31], certain prokaryotic Agonaute proteins also exhibit specificity for ssDNA-guided DNA targets binding [9, 19, 32–35]. The structural basis and molecular mechanisms underlying the recognition of DNA guides and interference with DNA targets by long pAgos have been extensively elucidated in *Pyrococcus furiosus* [32], *Thermus thermophiles* [35], *Clostridium butyricum* [36], and *Methanocaldococcus jannaschii* [33]. The long pAgos are a diverse group of endonucleases that possess the ability to utilize gDNA for precise cleavage of tDNA, a process commonly referred to as DNA interference [11, 34–36]. The cleavage of the target is dependent on the PIWI domain, which contains an RNase H catalytic fold with the DEDX (where X is commonly Asp or His) tetrad [2, 9, 35]. In comparison to long pAgos, SPARTA also exhibits specific recognition of 5′-phosphate (5′-P) gDNA and its tDNA or tRNA. SPARTA can be activated by both guide ssDNA–tRNA and guide ssDNA–target ssDNA duplexes. Despite possessing the PIWI domain, SPARTA lacks the catalytic DEDX tetrad and therefore does not exhibit cleavage activity toward targets [20]. It is evident that the immune system of SPARTA has a significantly different mechanism compared to that of long pAgos in defense against invading nucleic acids [20]. Although there have been investigations into and identification of the molecular basis for gRNA-mediated activation of SPARTA, our understanding regarding the mechanisms and structural basis underlying gDNA-mediated tDNA recognition and activation of SPARTA remains limited.

In this study, we present crystal structures of the *Crenotalea thermophila* (Crt) SPARTA complex bound to gDNA and tDNA of different lengths, revealing both inactive and active states of the SPARTA system. These structures elucidate the molecular basis for gDNA recognition and the mechanism underlying NADase activation induced by gDNA–tDNA duplex formation, thereby providing a comprehensive understanding of the antiphage defense function in short Ago immune systems.

Materials and methods

Bacterial strains

Escherichia coli strain DH5α was used for plasmid reconstruction, while *E. coli* BL21 Star (DE3) (Novagen) was utilized for protein expression.

Protein expression and purification

For the expression and purification of the *Crenotalea thermophila* (Crt) and *Maribacter polysiphoniae* (Map) SPARTA (TIR-APAZ and short pAgo proteins) binary complex proteins, the tir-apaz and ago DNA sequences for *Crenotalea thermophila* and *Maribacter polysiphoniae* were synthesized by Generay Biotech. The *Crenotalea thermophila* tir-apaz gene was cloned into a modified pET28a-HisSUMO expression vector, while the *Crenotalea thermophila* ago gene was cloned into a pET28a-SUMO expression vector. Both constructs have a HisSUMO or SUMO tag at the N-terminal end, followed by an ubiquitin-like protein 1 (Ulp1) protease cleavage site using ligation-independent cloning. The CrtTIR-APAZ and short CrtpAgo proteins were co-expressed in *E. coli* BL21 Star (DE3). The *Maribacter polysiphoniae* tir-apaz gene was cloned into the pCold-HisSUMO vector using homologous recombination with Ulp1 protease cleavage site, and the *Maribacter polysiphoniae* ago gene was cloned into a pET28a-SUMO expression vector. The MapTIR-APAZ and short MappAgo proteins were co-expressed in *E. coli* BL21 Star (DE3). The transformed constructs were introduced into *E. coli* BL21 Star (DE3), and the cell cultures were grown at 37°C until reaching an OD_{600} of 0.8–1.0. Subsequently, isopropyl-1-thio-β-D-galactopyranoside was added to a final concentration of 0.2 mM, and the cultures were incubated at 16°C for 14 h. The bacterial solutions were collected and resuspended in a buffer containing 20 mM Tris–HCl (pH 7.5) and 500 mM NaCl. After sonication and centrifugation, the supernatants were incubated with Ni Sepharose (GE Healthcare). The target proteins were then eluted using a buffer containing 20 mM Tris–HCl (pH 7.5), 300 mM NaCl, and 300 mM imidazole. The eluted proteins were digested with Ulp1 protease at 4°C for 2 h. Subsequently, the proteins were purified using a Heparin HP column (GE Healthcare) and further purified for uniformity using size exclusion chromatography (SEC) on a Superdex 200 Increase 10/300 column (Cytiva Life Sciences) in a buffer containing 20 mM Tris–HCl (pH 7.5), 150 mM NaCl, and 2 mM DTT. Indicated mutants and truncates were prepared in the same way.

Reconstitution and SEC analysis of the SPARTA–gDNA/gRNA–tDNA complex

For the reconstitution of the SPARTA–gDNA–tDNA or SPARTA–gRNA–tDNA complexes, the purified Crt/MapSPARTA protein was incubated with 21-nt 5′-phosphorylated gDNA or gRNA at a molar ratio of 1:1.05. Following a 30-min incubation at 4°C, the sample was supplemented with tDNA in a molar ratio of 1.1:1 to SPARTA. The mixture was then incubated for an additional 30 min to assemble the SPARTA–gDNA–tDNA complex or SPARTA–gRNA–tDNA complex. The complex was assayed by Superdex 200 Increase 10/300 column (Cytiva Life Sciences). The SEC buffer contained 20 mM Tris–HCl (pH 7.5), 150 mM NaCl, and 2 mM DTT. The peak fractions were collected separately and analyzed using sodium

dodecyl sulfate–polyacrylamide gel electrophoresis (SDS–PAGE). The gRNA and gDNA were chemically synthesized from GenScript. The different lengths of DNA targets were synthesized from Sangon.

Urea-PAGE analysis

The Crt/MapSPARTA complex samples elution from the SEC peak were incubated with 2× loading buffer (8 M urea, 0.05% bromophenol blue) and analyzed by 20% denaturing urea-PAGE with 1× tris-borate-ethylenediaminetetraacetic acid (TBE) buffer for 100 min. After electrophoresis, the gels were visualized by toluidine blue staining.

SDS–PAGE analysis

The Crt/MapSPARTA protein samples after purification by Ni Sepharose, Heparin HP, and SEC were analyzed using 12% SDS–PAGE with 1× TG buffer (0.025 M Tris, 0.192 M glycine, 0.1% SDS, pH 8.3), then stained with the buffer (2.5 g/l R250, 40% ethyl alcohol, 10% acetic acid in ddH_2O), and washed with the destaining solution (40% ethyl alcohol, 10% acetic acid in ddH_2O).

Electrophoretic mobility shift assay

Electrophoretic mobility shift assays (EMSAs) used gDNA and tDNA purchased from Sangon Biotech. Four micromolar 5′-phosphorylated gDNA was incubated with CrtSPARTA or CrtSPARTATIR$^{1–421}$ for 15 min. Subsequently, the mixture was supplemented with increasing concentrations of 21-nt tDNA in a buffer containing 20 mM Tris–HCl (pH 7.5) and 150 mM NaCl in a total volume of 10 μl. The incubations were carried out at 4°C for an additional 30 min. One microliter of loading buffer for prestained DNA was added to the incubation samples. The incubation samples were then analyzed using native polyacrylamide gels of 6% in Tris–Glycine buffer (pH 8.5). The electrophoresis was conducted at 100 V for 40 min at 4°C. The gels were visualized using the Chemi-Doc XRS+ electrophoresis apparatus (Bio-Rad).

In vitro NADase activity assays

The *in vitro* NADase activity assays were performed as previously described [20]. The ϵ-NAD$^+$ (nicotinamide 1,N^6-ethenoadenine dinucleotide) *in vitro* NADase activity assays were conducted in a reaction buffer containing 10 mM MES (pH 6.8), 150 mM NaCl, and 5 mM $MgCl_2$. Three micromolar CrtSPARTA or CrtSPARTA$^{1–421}$ was incubated with gDNA at 55°C for 15 min. Subsequently, the DNA target was introduced to the reaction mixture, and the mixture was further incubated at 55°C for 1 h. The molar ratio of CrtSPARTA or CrtSPARTATIR$^{1–421}$ to gDNA and tDNA was maintained at a 1:1:1 stoichiometry. Following incubation, ϵ-NAD$^+$ (Sigma–Aldrich) was added to the reaction system. The final concentration of ϵ-NAD$^+$ was 50 μM in a total volume of 30 μl. The NADase activity assays were conducted at 42°C in wells of a white low-volume 384-well plate (Corning) and monitored using the Tecan E plex system (TECAN). The excitation wavelength was set at 310 nm, while the emission wavelength was set at 410 nm. The experiments were conducted in triplicate.

Crystallization and data collection

The preliminary crystallization screening of Apo CrtSPARTA and CrtSPARTA$^{1–421}$–gDNA–tDNA was conducted using crystallization screening kits (Hampton Research, Molecular Dimensions, and QIAGEN) at 16°C. The screened crystal conditions were further optimized using the hanging drop vapor diffusion method at 16°C. The protein-to-reservoir solution ratio was 1:1, with a total volume of 2 μl. The crystals of Apo CrtSPARTA were grown in a solution containing 2.0 M ammonium sulfate and 0.1 M Tris–HCl (pH 8.0). The gDNA and 16-nt tDNA-bound CrtSPARTA$^{1–421}$ complex crystals were crystallized in 0.1 M BIS-TRIS propane (pH 9.0) and 8% (w/v) polyethylene glycol (PEG) 20 000. The gDNA and 21-nt tDNA-bound CrtSPARTA$^{1–421}$ complex crystals were obtained from a solution containing 2% (v/v) Tacsimate (pH 7.0), 5% (v/v) 2-propanol, 0.1 M imidazole (pH 7.0), and 8% (w/v) PEG 3350. The gDNA and 19-nt tDNA-bound CrtSPARTA$^{1–421}$ complex crystals were obtained from a solution containing 0.1 M BICINE (pH 8.5) and 4.375% (v/v) PEG 4000. The gDNA and 20-nt tDNA-bound CrtSPARTA$^{1–421}$ complex crystals were obtained from a solution containing 0.1 M BICINE (pH 8.5) and 4.625% (v/v) PEG 4000. The Apo-CrtSPARTA crystals were soaked in cryoprotectants supplemented with 20% (v/v) ethylene glycol, while the CrtSPARTA$^{1–421}$–gDNA–tDNA crystals were cryoprotected using corresponding reservoir buffers supplemented with 20% (v/v) glycerol and flash frozen in liquid nitrogen. The X-ray diffraction datasets were collected at 100 K at beamlines BL-18U1, BL-19U1, BL-02U1, and BL-10U2 at Shanghai Synchrotron Radiation Facility (SSRF) in China.

Structure determination

The X-ray diffraction datasets were processed using HKL 3000 package [37]. The Apo-CrtSPARTA complex structure was determined by molecular replacement with PHASER using the predicted structures by AlphaFold (accession codes: AF-A0A1I7NFD7-F1 and AF-A0A1I7NFG5-F1) as the search models [38]. The 16- and 21-nt tDNA-bound CrtSPARTA$^{1–421}$–gDNA–tDNA complex structures were determined by molecular replacement using the structure of Apo CrtSPARTA as the search model. The 19- and 20-nt tDNA-bound CrtSPARTA$^{1–421}$–DNA–tDNA complex structures were determined by molecular replacement using the structure of 21-nt tDNA-bound CrtSPARTA$^{1–421}$–gDNA–tDNA complex as the search model. Atomic models were built manually in COOT and refined using PHENIX [39, 40]. All data processing and structure refinement statistics are summarized in Table 1 and Supplementary Table S1. Structure figures were prepared using PyMOL (http://www.pymol.org/).

Results

C-terminal tail of TIR-APAZ regulates gDNA binding

Previous studies have demonstrated that SPARTA exhibits superior activation efficiency toward guide ssRNA compared to guide ssDNA [20], prompting us to hypothesize that this discrepancy may arise from the differential binding affinity of SPARTA toward these two types of guides. To investigate this hypothesis, we conducted EMSA and SEC assays to compare the binding capacity of CrtSPARTA and *Maribacter polysiphoniae* (Map) SPARTA with guide ssRNA and guide ssDNA.

Table 1. Crystallographic data collection and refinement statistics

	Apo SPARTA	SPARTA–gDNA–16-nt tDNA	SPARTA–gDNA–21-nt tDNA
Data collection[a]			
Space group	$P6_522$	$C222_1$	$P3_221$
Cell dimensions			
a, *b*, *c* (Å)	198.4, 198.4, 184.4	201.2, 287.3, 111.8	133.6, 133.6, 167.3
α, β, γ (°)	90.0, 90.0, 120.0	90.0, 90.0, 90.0	90.0, 90.0, 120.0
Resolution (Å)	50.00–3.18 (3.26–3.18)	50.00–3.85 (3.92–3.85)	50.00–3.36 (3.42–3.36)
R_{pim}	0.052 (1.097)	0.170 (0.518)	0.052 (0.276)
I/σ_I	12.5 (0.8)	3.5 (1.3)	14.8 (3.0)
Completeness (%)	99.9 (100)	99.6 (99.4)	100 (100)
Redundancy	39.6 (40.0)	4.6 (4.0)	17.3 (14.8)
Refinement			
Resolution (Å)	3.18	3.85	3.36
No. of reflections	36 416	38 276	22 264
R_{work}/R_{free}	0.2742/0.2993	0.2468/0.2648	0.2470/0.2872
No. of atoms			
Protein	7761	14 536	7279
Nucleic acid	–	1514	858
Ion/ligand	2	0	1
Water	7	0	25
B-factors (Å^2)			
Protein	113.6	58.7	57.7
Nucleic acid	–	99.1	75.4
Ion/ligand	63.9	–	43.3
Water	89.5	–	23.1
R.m.s. deviations			
Bond length (Å)	0.013	0.004	0.015
Bond angles (°)	1.646	0.807	1.753
Ramachandran plot			
Favored region	94.94	96.04	95.49
Allowed region	4.85	3.74	4.06
Outlier region	0.21	0.23	0.45

[a]Highest resolution shell is shown in parentheses.

Our findings demonstrate that both gRNA and gDNA can interact with SPARTA to form relatively stable complexes (Supplementary Fig. S1A–C), indicating that SPARTA efficiently recognizes both gRNA and gDNA without a significant difference in binding affinity.

To investigate whether there is a recognition model disparity between gRNA and gDNA by SPARTA, we first determined the crystal structure of Apo CrtSPARTA at a resolution of 3.18 Å (Fig. 1A and B, and Table 1). CrtSPARTA comprises two components: Ago and TIR-APAZ. Ago consists of MID and PIWI domains, while TIR-APAZ consists of TIR and APAZ domains, and a C-terminal tail (Fig. 1A). Through electrostatic potential analysis, we discovered that the MID, PIWI, and APAZ domains form a predominantly positively charged channel that accommodates the C-terminal tail of TIR-APAZ (Fig. 1C). The C-terminal tail primarily contains Asp and Glu residues (Fig. 1D), exhibiting a highly negatively charged surface (Fig. 1E). Previous structural determination of the gRNA-bound CrtSPARTA complex has revealed that gRNA is located within the positively charged channel formed by MID, PIWI, and APAZ domains [25, 29, 41]. A structural comparison between Apo-CrtSPARTA complex and gRNA-bound complex demonstrates clashes between nucleotides in the gRNA with residues from the C-terminal tail of TIR-APAZ (Supplementary Fig. S1D). This observation suggests that inhibition of gRNA or gDNA binding may be attributed to interactions involving the C-terminal tail of TIR-APAZ.

To investigate the impact of the C-terminal tail of TIR-APAZ on gRNA and gDNA binding, we generated a truncated version of TIR-APAZ lacking the C-terminal tail and compared its binding ability with wild-type (WT) TIR-APAZ. Truncation of the C-terminal tail did not affect the formation of a stable complex between TIR-APAZ and Ago protein. Our EMSA and SEC analyses reveal that CrtSPARTA$^{1–421}$ exhibits higher affinity for gRNA and gDNA compared to CrtSPARTA^{WT} (Supplementary Fig. S1A–C). Additionally, SEC data indicate that MapSPARTA$^{1–420}$ displays stronger capability than MapSPARTA^{WT} in accommodating gDNA (Supplementary Fig. S1A). Furthermore, we assessed tDNA binding ability between SPARTA variants with or without C-terminal tail truncation in the presence of gDNA. Similar to our findings regarding gDNA binding, truncation of the C-terminal tail enhances tDNA binding capacity in CrtSPARTA (Fig. 1F and G). Both CrtSPARTA TIR$^{1–421}$ and MapSPARTA TIR$^{1–420}$ form more homogeneous complexes with both gDNA and tDNA compared to their respective WT counterparts (Fig. 1F). The findings suggest that the C-terminal tails of TIR-APAZ may mimic nucleic acids, thereby occupying the binding channel for gRNA or gDNA, consequently hindering their recognition by SPARTA.

DNA target length requirement for ssDNA-guided SPARTA activation

We investigated the length requirement of tDNA for CrtSPARTA$^{1–421}$ and gDNA recognition through SEC experiments. The results demonstrate that CrtSPARTA$^{1–421}$ and gDNA can bind to 5-, 10-, 15-, and 21-nt tDNAs, forming stable complexes (Fig. 2A). This suggests that the recognition

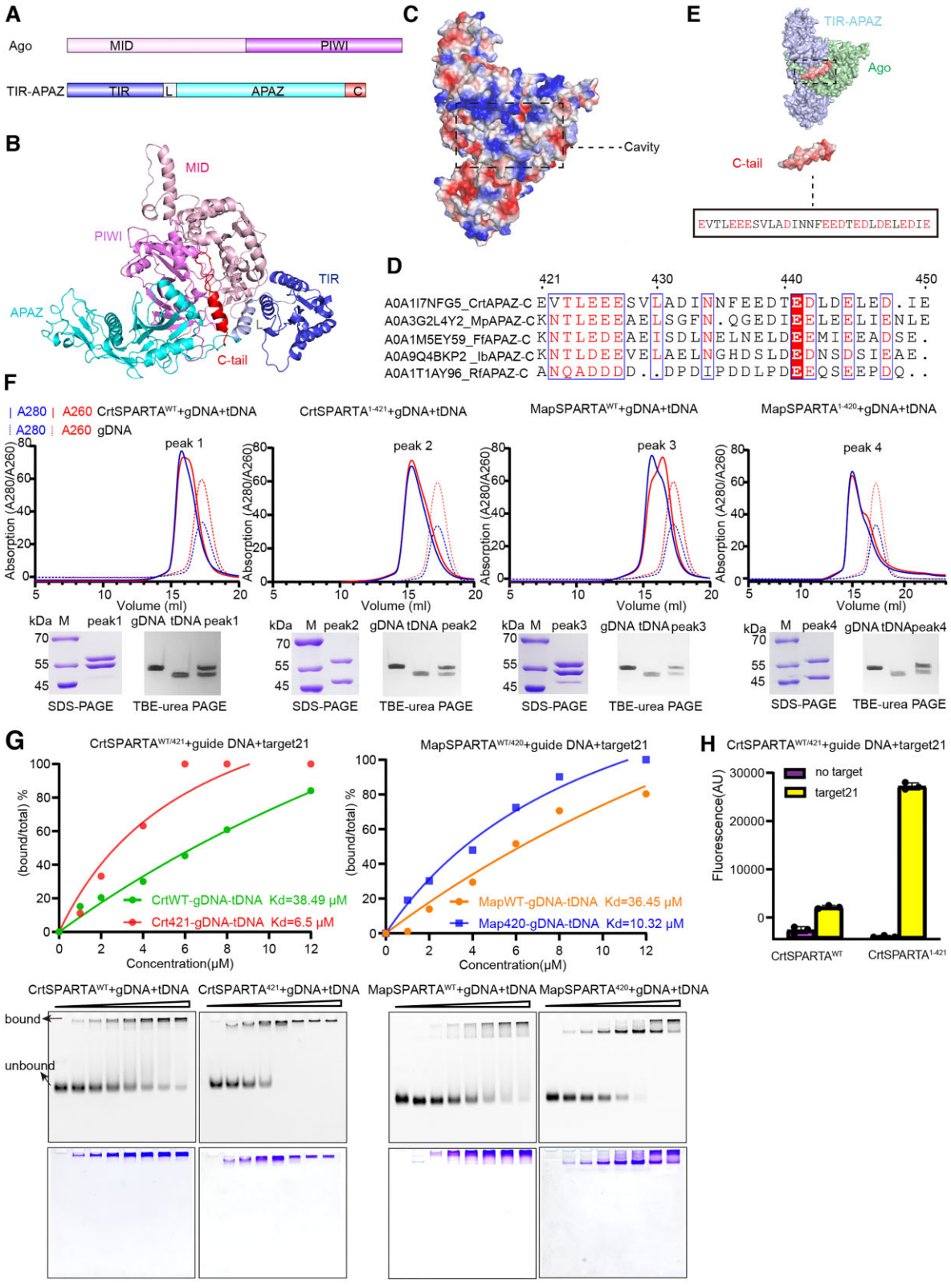


Figure 1. C-terminal tail of TIR-APAZ regulates gDNA binding and NADase activation. (**A**) Domain organization of CrtAgo and CrtTIR-APAZ. (**B**) The overall structure of CrtSPARTA complex. Color-coded as defined in panel (A). (**C**) The surface electrostatic potential of the CrtSPARTA. The nucleic acid binding cavity is highlighted by dashed line. (**D**) Sequence alignments of the C-terminal tail of APAZ domains. Each sequence is indicated with its UniProt number and the abbreviated name of the organism. The conserved residues are highlighted with a box. (**E**) The negatively charged C-terminal tail of APAZ domain occupies the nucleic acid binding channel. The C-terminal tail consists mainly of negatively charged amino acids. (**F**) SEC experiments assaying the binding ability of CrtSPARTA, CrtSPARTA$^{1–421}$, MapSPARTA, and MapSPARTA$^{1–420}$ to gDNA and tDNA. The solid curve is the sample containing SPARTA, gDNA, and tDNA, and the dashed curve is the sample containing only gDNA and tDNA. The peak fraction is analyzed by SDS–PAGE and TBE–urea PAGE. (**G**) EMSA experiments assaying the binding ability of CrtSPARTA, CrtSPARTA$^{1–421}$, MapSPARTA, and MapSPARTA$^{1–420}$ to gDNA and tDNA. (**H**) Comparison of DNA-guided NADase activity between CrtSPARTA^{WT} and CrtSPARTA$^{1–421}$. Data are shown as mean ± standard deviation ($n = 3$ independent experiments).

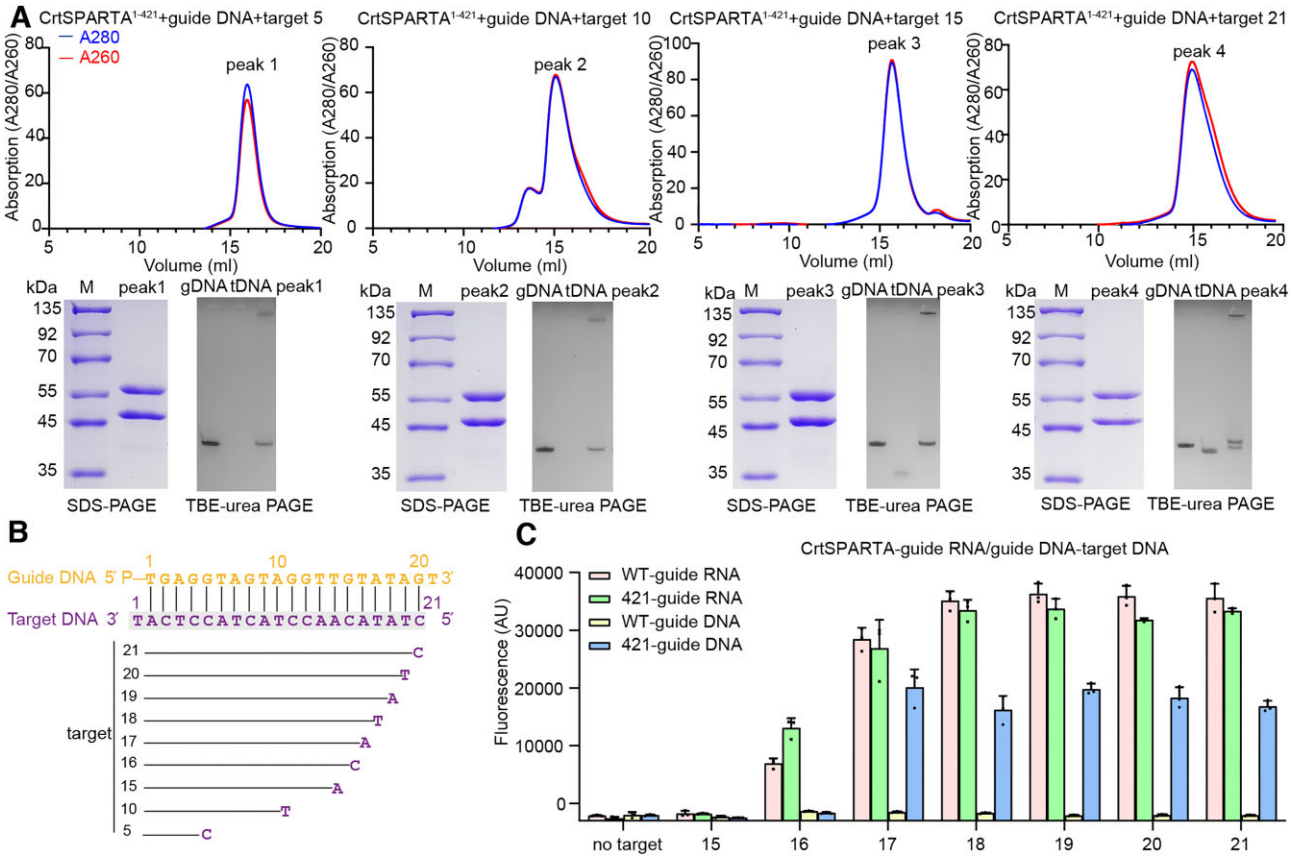


Figure 2. The length of DNA target affects NADase activation. (**A**) SEC experiments assaying the binding ability of CrtSPARTA$^{1–421}$ to gDNA and tDNA of different lengths (5, 10, 15, and 21 nt). The peak fraction is analyzed by SDS–PAGE and TBE–urea PAGE. (**B**) The schematic representation of the gDNA and tDNA of different lengths. (**C**) NADase activity assays with 21-nt gDNA or gRNA and different lengths of tDNA. Data are shown as mean ± standard deviation ($n = 3$ independent experiments).

of DNA targets by SPARTA$^{1–421}$ and gDNA is not influenced by their length.

The activation of ssRNA-guided SPARTA has been reported to have a tDNA length requirement. To investigate whether this length requirement also applies to ssDNA-guided SPARTA activation and whether there are any differences in the length requirements between ssRNA- and ssDNA-guided SPARTA activation, we conducted NADase activity assays using gDNA and tDNA with varying lengths. Our results demonstrate that activation of CrtSPARTA$^{1–421}$ requires a minimum tDNA length of 17 nt (there are 16 complementary nucleotides between tDNA and gDNA) (Fig. 2B and C, and Supplementary Fig. S2). None of the tested tDNAs ranging from 15 to 21 nt can activate ssDNA-guided CrtSPARTA^{WT}. Although a 16-nt tDNA (15 nt complementary with gRNA) shows weak NADase activity for both ssRNA-guided CrtSPARTA$^{1–421}$ and CrtSPARTA^{WT}, it is not sufficient for robust activation. In contrast, high NADase activity in ssRNA-guided CrtSPARTA$^{1–421}$ and CrtSPARTA^{WT} requires a complemented tDNA longer than 16 nt. These findings indicate that the length of the tDNA plays a crucial role in SPARTA activation, necessitating at least 17 nt (16-nt complementarity with guide) for achieving relatively complete SPARTA activation.

Similar to the gRNA and tDNA-bound SPARTA complex [23–29, 42], the gDNA and tDNA-bound SPARTA complex also shows a characteristic two-lobed architecture, with the CrtTIR domain and the C-terminal APAZ domain forming arms that envelop the MID and PIWI domains of the short pAgo (Fig. 3B). The MID, PIWI, and APAZ domains form a positively charged channel for binding the guide–target duplex (Fig. 3F). The guide–target duplex is further stabilized by a network of direct hydrogen bonds between the SPARTA and the phosphate groups of DNA (Fig. 3C and Supplementary Fig. S3C and D). The gDNA is recognized by residues of Arg72, Asn484, Ser391, and Arg201 through the formation of hydrogen bonds between phosphate groups of nucleotides and side chains of residues. The tDNA is bound by CrtSPARTA primarily through hydrogen bond interactions between the phosphate groups of nucleotides T0′–C8′ and specific residues Thr228, Arg481, Asp474, Glu476, Lys395, Asn439, His430, Tyr285, and His358. Furthermore, the bases A1′ and T0′ are stabilized by stacking interactions with the side chains of Tyr148, His207, and Leu252. The 5′-hydroxyl group of nucleotide T0′ is bound by the side chain of Thr255 while being coordinated by Mg^{2+} ion as well as side chains of Asn468 and main chain of Ile507.

Different gRNA and gDNA recognition models

Although the CrtSPARTA–gRNA–tDNA and CrtSPART–gDNA–tDNA complexes exhibit similar overall architecture (Supplementary Fig. S3E), they demonstrate distinct modes of recognition for gRNA and gDNA. The structure of the SPARTA complex bound to gRNA and tDNA reveals that the 5′-end phosphate of the first guide nucleotide flips out and

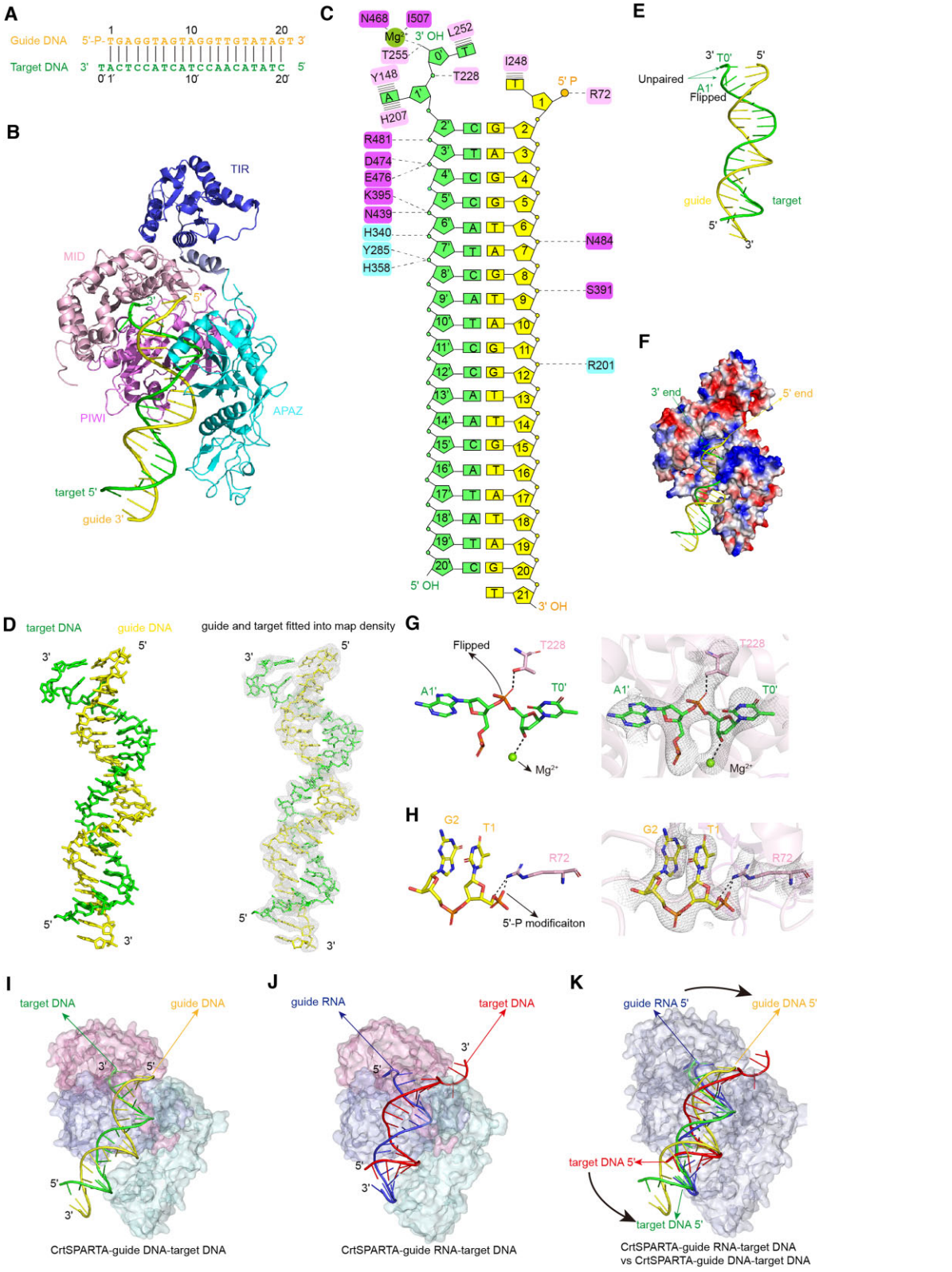


Figure 3. Overall structure of CrtSPARTA$^{1–421}$ in complex with gDNA and 21-nt tDNA. (**A**) Schematic representation of 5′-P gDNA and tDNA oligonucleotides used for crystallization. (**B**) The overall structure of gDNA and 21-nt tDNA-bound CrtSPARTA$^{1–421}$ complex. (**C**) Schematic representation of the intermolecular contacts between SPARTA and guide–target duplex in the complex structure. The base, backbone sugar, and phosphate group of the gDNA and tDNA are shown as rectangles, pentagons, and circles, respectively. Protein residues are color-coded by their respective domains. (**D**) Electron density map of the gDNA–tDNA duplex. The 2mFo-Fc electron density map of the gDNA–tDNA duplex is shown as a mesh (contoured at 1.0σ). (**E**) Structure of the gDNA–tDNA duplex. (**F**) The gDNA–tDNA duplex binds in the positively charged channel of the CrtSPARTA. (**G**) Detailed interactions between the 3′-end nucleotides (A1′–T0′) of tDNA and the MID binding pocket of SPARTA. (**H**) Detailed interactions between the 5′-end nucleotides (T1–G2) of gDNA and the MID domain of SPARTA. The surface representation of CrtSPARTA reveals distinct spatial localization of the gDNA–tDNA duplex (**I**) and gRNA–tDNA duplex (**J**). (**K**) The superposition of the gDNA–tDNA duplex with the gRNA–tDNA duplex in CrtSPARTA complexes reveals distinct models for nucleic acid recognition.

anchors in the MID pocket. Additionally, an Mg^{2+} ion is coordinated with the 5′-phosphate group within the MID binding pocket [21, 24, 25, 27, 31]. However, in the CrtSPARTA–gDNA–tDNA complex structure, it is evident that the MID pocket and Mg^{2+} ion binding network accommodate a hydroxyl group at the 3′-end of the target strand (Fig. 3G and Supplementary Fig. S4A). Meanwhile, the second nucleotide at 3′-terminal of target strand is flipped away from its complementary base in gDNA (Fig. 3G). In contrast, the 5′-end of the guide is positioned on the groove formed between MID and APAZ domains (Fig. 3B and H), which is distant from the MID binding pocket. The first unpaired nucleotide at gDNA's 5′-end extends outwardly without contacting its corresponding base on target strand (Supplementary Fig. S4A). Notably, Arg72 within the MID domain specifically recognizes the phosphorylated group attached to gDNA's 5′-end, which provides a structural basis for recognition of gDNA's terminus (Fig. 3H). This unique orientation between target and guide strands distinguishes itself from all previously reported guide and target-bound pAgo and SPARTA complex structures (Supplementary Figs S4 and S5).

Further comparison of the CrtSPARTA–gDNA–tDNA (Fig. 3I) and CrtSPARTA–gRNA–tDNA (Fig. 3J) structures reveals distinct spatial disparities in the positioning of guide and target strands between the two architectures. In the CrtSPARTA–gDNA–tDNA structure, the 3′-end of the target strand is situated between MID and PIWI domains. Conversely, this position serves as an anchor for the 5′-end of the guide strand in the CrtSPARTA–gRNA–tDNA structure. The superposition of CrtSPARTA–gDNA–tDNA and CrtSPARTA–gRNA–tDNA structures demonstrates a position exchange of guide and target strands between the two structures (Fig. 3K). Overall, despite their structural and functional similarities, these two SPARTA complexes exhibit opposite orientations for their guide and target strands.

Structural superposition of SPARTA in complex with RNA–DNA heteroduplex and SPARTA in complex with DNA–DNA duplex reveals that there are no significant conformational changes in SPARTA between the two complexes (Fig. 3I–K and Supplementary Fig. S3E). Both RNA–DNA hybrid and DNA–DNA duplex bind to the positively charged channel of SPARTA (Fig. 3F and K), adopting a B-form double-stranded DNA (dsDNA) duplex structure similar to that of canonical DNA–DNA duplex (Fig. 3K and Supplementary Fig. S3B). Interaction analysis between SPARTA and RNA–DNA duplexes [23–25, 27–29, 42], as well as between SPARTA and DNA–DNA duplexes, reveals that SPARTA primarily recognizes the phosphate backbone of these duplex structures (Fig. 3C). Consequently, SPARTA can accommodate the distinct RNA–DNA hybrid and DNA–DNA duplex substrates.

15-bp gDNA–tDNA match complex structure

Our biochemical characterization reveals that the minimal complementary length of tDNA and gDNA required for CrtSPARTA NADase activation exceeds 15 bp. To illuminate the structural basis underlying the DNA target length necessary for SPARTA activation, we determined the crystal structure of CrtSPARTA in complex with a 21-nt 5′-P gDNA and a 16-nt tDNA at 3.85 Å resolution (Fig. 4A and B, and Table 1). There are 15-bp complementary nucleotides between gDNA and tDNA sequences. The nucleotides T1–G15 of gDNA are complementary to the nucleotide A1′–C15′ of tDNA. Biochemical data indicate that the 15-bp gDNA–tDNA match complex exhibits negligible NADase activity. The 15-bp gDNA–tDNA match CrtSPARTA complex adopts a two-lobed architecture (Supplementary Fig. S6A), where the central channel formed between the lobes accommodates the binding of the gDNA–tDNA duplex. Despite the gDNA and tDNA sequences containing 15-bp complementary nucleotides, only a 14-bp A-form DNA duplex is formed between them (Supplementary Fig. S6B). Specifically, T1–T14 of gDNA form Watson–Crick base pairs with A1′–A14′ of tDNA (Fig. 4C and D), while G15 of gDNA and C15′ of tDNA undergo conformational changes resulting in their spatial separation (Fig. 4C, E, and F). Notably, PIWI domain stabilizes G15 nucleotide in gDNA, whereas MID domain accommodates C15′ nucleotide in tDNA.

The bases of nucleotides G15 and C15′ are stabilized by residues Met435 and Gln205 through stacking and hydrogen bond interactions (Fig. 4C and F), respectively. The phosphate group of nucleotide G15 is recognized by the side chain of Arg72 (Fig. 4C). Importantly, the 5′-hydroxyl group of nucleotide C15′ of tDNA locates in the MID binding pocket (Fig. 4C and F), which is fixed by the hydroxyl of Ile223 and coordinated by the Mg^{2+} ion as well as the side chains of His207, Gln222, Lys263, and Asn468 and main chain of Ile507. Additionally, Arg243, Arg225, and Thr228 form hydrogen bonds with the phosphate group of nucleotide A14′ in tDNA to prevent base pair formation between T14 in gDNA and A14′ in tDNA. The 14-bp gDNA–tDNA duplex is stabilized by SPARTA through extensive hydrogen bond interactions (Fig. 4C and Supplementary Fig. S6C). Base pairs 1–10 are bound by APAZ and PIWI domain via interactions with the phosphate backbone and side chains of specific residues. Base pairs 11–14 are predominantly recognized by MID and PIWI domains through hydrogen bonds interaction involving Asn468, Arg481, Gly475, Ser431, Arg72, and Lys247 residues. Furthermore, G15–T21 nucleotides from gDNA extend outside SPARTA along the duplex binding channel (Fig. 4B and C).

Conformation changes upon gDNA–tDNA loading

To gain deeper insights into the recognition mechanism of CrtSPARTA toward the guide and target DNA duplex, we conducted a comparative analysis between the Apo-CrtSPARTA structure and the 20-nt gDNA–tDNA match CrtSPARTA complex. Superimposition of these two structures reveals a modest conformational change in CrtSPARTA upon DNA binding, with a core root-mean-square deviation (RMSD) of 1.88 Å for 837 Cα atoms (Fig. 5A). Further examination of individual domains demonstrates minimal conformational alterations in the MID, TIR, and PIWI domains, while significant rearrangements are observed in the APAZ domain (Fig. 5B). The APAZ domain exhibits substantial displacement from both MID and PIWI domains to create an expanded channel for accommodating the guide and target DNA duplex (Fig. 5C). These conformational changes within the APAZ domain contribute to enhanced stability of the gDNA–tDNA duplex formation. Specifically, two β-strands connected by a loop (residues ranging from Asn281 to Glu296) shift by ∼4 Å to anchor within the minor groove of gDNA–tDNA duplex (Fig. 5D), whereas an α-helix region (residues ranging from Leu350 to Asn370) undergoes moderate rearrangement to stabilize interactions with major groove nucleotides of gDNA–tDNA duplex (Fig. 5E). Additionally, analysis of Apo-CrtSPARTA complex structure reveals that

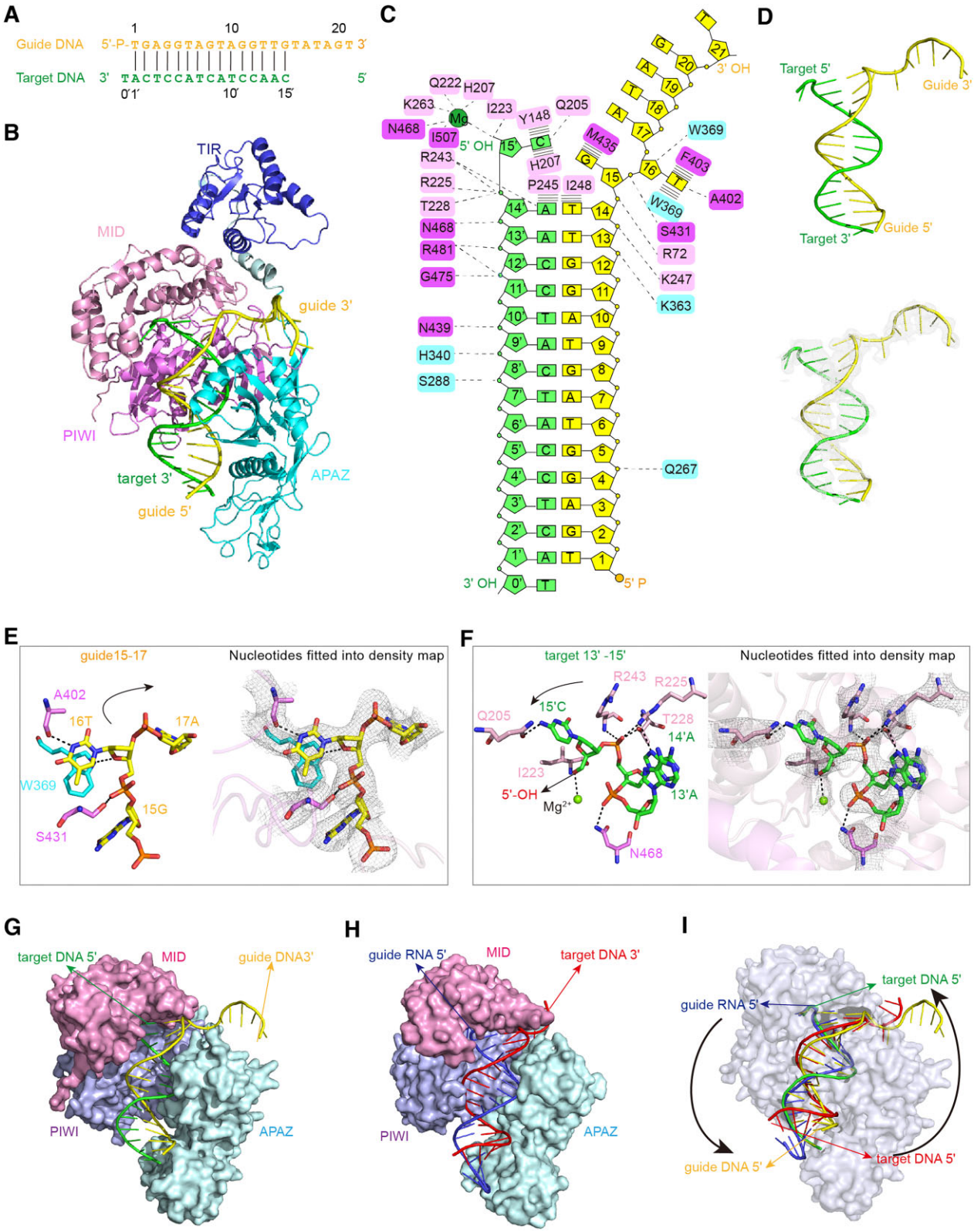


Figure 4. Overall architecture of CrtSPARTA$^{1–421}$ in complex with gDNA and 16-nt tDNA. (**A**) The sequences of 5′-P gDNA and 16-nt tDNA were used for crystallization. (**B**) The overall structure of gDNA and 16-nt tDNA-bound CrtSPARTA$^{1–421}$ complex. (**C**) Schematic representation of the intermolecular contacts between SPARTA and guide–target duplex in the complex structure. The base, backbone sugar, and phosphate group of the gDNA and tDNA are shown as rectangles, pentagons, and circles, respectively. Protein residues are color-coded by their respective domains. (**D**) Electron density map of the gDNA–tDNA duplex. The 2mFo-Fc electron density map of the gDNA–tDNA duplex is shown as a mesh (contoured at 1.0σ). (**E**) Detailed interactions between the nucleotides G15–A17 of gDNA and the MID and PIWI domains of SPARTA. (**F**) Detailed interactions between the 5′-end nucleotides (C15′–A13′) of tDNA and the MID binding pocket of SPARTA. The surface representation of CrtSPARTA reveals distinct spatial localization of the 15-bp gDNA–tDNA duplex (**G**) and 20-bp gRNA–tDNA duplex (**H**). (**I**) The superposition of the 15-bp gDNA–tDNA duplex with the 20-bp gRNA–tDNA duplex in CrtSPARTA complexes reveals distinct models for nucleic acid recognition.

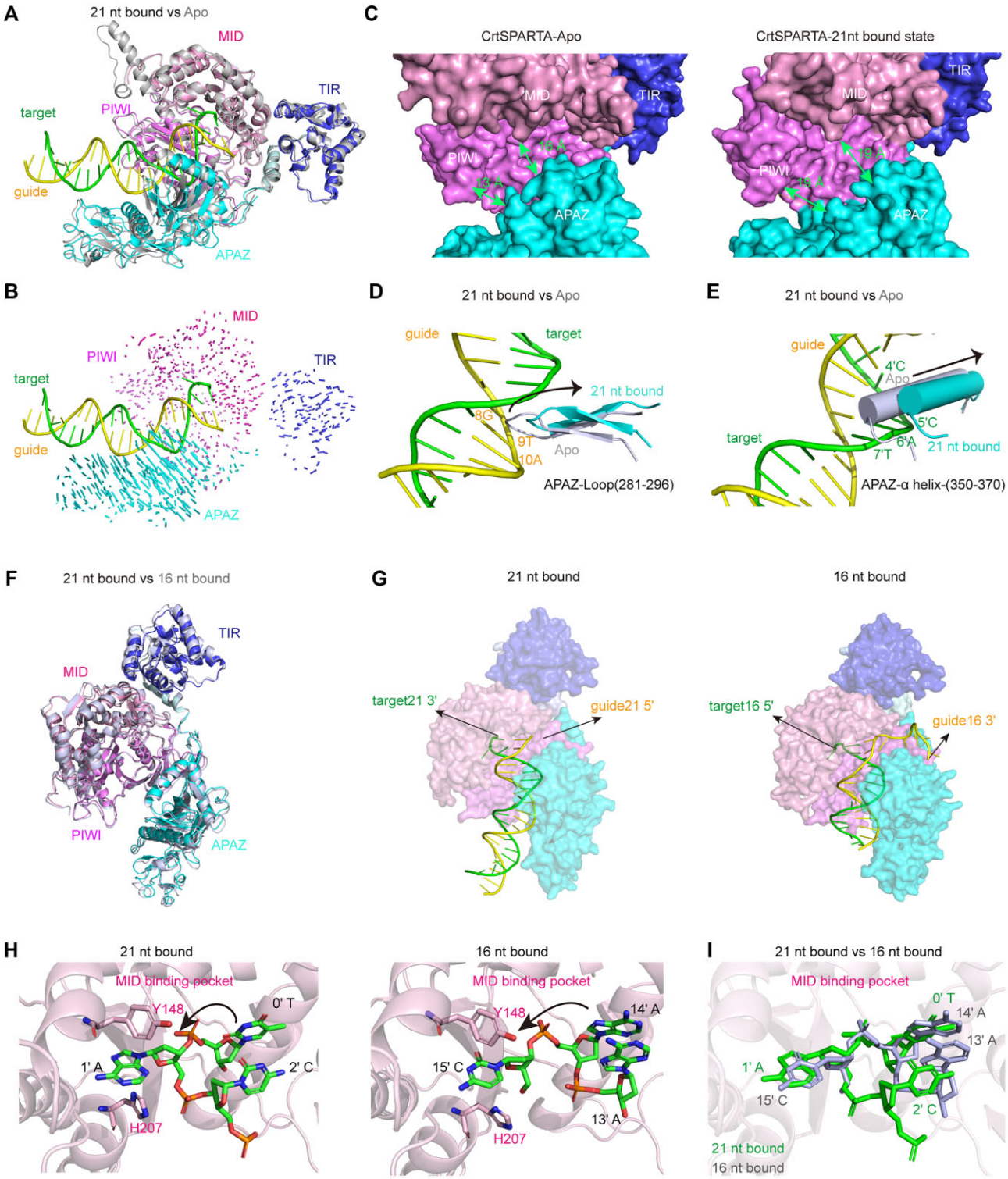


Figure 5. Conformation changes upon the guide and target binding. (**A**) Structure superposition of the Apo SPARTA with the 21-nt tDNA-bound SPARTA (color-coded as defined in Fig. 3B). (**B**) Structural comparison between the Apo SPARTA with the 21-nt tDNA-bound SPARTA. The Cα–Cα vector map is shown as lines. The vector lengths proportionally represent the domain movement between the two complexes being compared. (**C**) Structural comparison of the nucleic acid binding channel before (left panel) and after (right panel) gDNA–tDNA duplex binding. The diameter of the DNA duplex binding channel within the SPARTA is indicated. (**D**) Structure superposition of two β-strands (residues ranging from Asn281 to Glu296) in the Apo SPARTA and the 21-nt tDNA-bound SPARTA (color-coded as defined in Fig. 3B). (**E**) Structure superposition of an α-helix (residues ranging from Leu350 to Asn370) in the Apo-SPARTA and the 21-nt tDNA-bound SPARTA (color-coded as defined in Fig. 3B). (**F**) Structure superposition of SPARTA in the 16-nt tDNA-bound SPARTA complex and the 21-nt tDNA-bound SPARTA complex (color-coded as defined in Fig. 3B). (**G**) The surface representation of CrtSPARTA reveals distinct orientation of the gDNA and tDNA in the 21-nt tDNA-bound SPARTA complex (left panel) and the 16-nt tDNA-bound SPARTA complex (right panel). (**H**) Structural comparison of the MID binding pocket in the 21-nt tDNA-bound SPARTA complex (left panel) and the 16-nt tDNA-bound SPARTA complex (right panel). (**I**) Structure superposition of the nucleotides of 21-nt DNA bound in the MID pocket with the nucleotides of 16-nt DNA bound in the MID pocket.

the C-terminal tail region of TIR-APAZ protein is positioned between MID and APAZ domains, occupying part of the binding channel for gDNA–tDNA duplex (Supplementary Fig. S7A). However, superimposition with the structure containing a 20-bp gDNA–tDNA duplex shows steric clashes between this C-terminal tail region and bound nucleic acids (Supplementary Fig. S7B), suggesting its requirement for extensive rearrangement upon DNA duplex binding process so as to create sufficient space for accommodating complete formation of gDNA–tDNA.

Distinct recognition for 16- and 21-nt tDNA

To investigate the variations in NADase activity with different lengths of DNA targets, we compared the structures of CrtSPARTA complexes bound to 16-nt and 21-nt tDNA. Superimposing the two CrtSPARTA molecules in these structures reveals a core RMSD of 1.28 Å for 890 Cα atoms (Fig. 5F), indicating moderate conformational differences between them. In particular, the APAZ and TIR domains are not well aligned in these two structures. Furthermore, modest structural differences are observed when comparing the binding channels of gDNA–tDNA duplexes in these two structures (Supplementary Fig. S7C). The alignment of gDNA–tDNA duplexes also shows poor overlap between these two structures (Supplementary Fig. S7D), suggesting significant conformational disparities.

Remarkably, significant differences in the orientation of gDNA and tDNA between the 16-nt and 21-nt tDNA-bound CrtSPARTA complexes were observed through further structural comparison (Fig. 5G and Supplementary Fig. S7E). In the 16-nt tDNA-bound CrtSPARTA complex, the 5′-end nucleotides of the tDNA are inserted into the MID binding pocket (Fig. 5H), while the gDNA's 3′-end nucleotides reside in the groove between APAZ and MID domains. Both the 3′-end nucleotides of tDNA and the 5′-end nucleotides of gDNA extend along the dsDNA binding channel toward the exterior of the CrtSPARTA complex. Conversely, in the 21-nt tDNA-bound CrtSPARTA complex, it is observed that the MID binding pocket accommodates the 3′-end nucleotides of tDNA (Fig. 5H and I), whereas anchoring occurs for gDNA's 5′-end nucleotides within a groove formed by APAZ and MID domains. The nucleotides of 3′-end of tDNA and the 5′-end of gDNA stretch toward protein's outer surface. Furthermore, our structural analysis reveals distinct recognition patterns of the nucleotides at the 5′-end of 16-nt tDNA and the 3′-end of 21-nt tDNA within the MID binding pocket of CrtSPARTA (Fig. 5H and I). In the structure with bound 16-nt tDNA, the first nucleotide C15′ at the 5′-end disrupts the gDNA–tDNA duplex and engages in stacking interactions with Tyr148 and His207. Conversely, in the structure with bound 21-nt tDNA, this stacking interaction is formed by the second nucleotide A1′ at the 3′-end with Tyr148 and His207. The differential recognition exhibited by CrtSPARTA toward these two DNA lengths may be associated with their disparate efficiencies for NADase activation.

Discussion

Similar to certain long pAgo proteins, CrtSPARTA exhibits the ability to employ ssDNA guides for targeting foreign DNA sequences and conferring immunity in prokaryotes. Through structural and biochemical investigations, this study provides valuable insights into the mechanism underlying its ssDNA-guided DNA targeting, elucidating the processes involved in both DNA guide recognition and DNA target recognition associated with NADase activation. Furthermore, our findings offer a structural explanation for the required length of DNA targets for CrtSPARTA activation. These mechanisms not only enhance our understanding of the immune function of SPARTA system in prokaryotes but also contribute to the development and application of DNA detection tools utilizing ssDNA guidance.

We have determined the structure of the SPARTA complex bound to 5′-P gDNA and tDNA, revealing a distinct recognition model for guide and target compared to the 5′-P gRNA-bound SPARTA complex structure. The structural basis underlying gRNA and tDNA recognition by the SPARTA complex has been extensively investigated [23–25, 27–29, 42]. Specifically, the conserved MID binding pocket exhibits specific recognition of the 5′-P of gRNA, explaining its preference for the 5′-P over the 5′-OH. In both CrtSPARTA–gRNA–tDNA and MapSPARTA–gRNA–tDNA complex structures, Lys211 and Ile223 residues within the MID domain cooperate with Mg^{2+} to specifically recognize and stabilize the first nucleotide at the 5′-end of gRNA [24, 27]. The initial nucleotide at the 5′-terminal of gRNA undergoes conformational flipping and establishes interactions with conserved residues in MID domain to establish a robust hydrogen network, thereby further enhancing the stability of gRNA's 5′-terminal binding (Supplementary Figs S4C and D and S5). Interestingly, in contrast to these findings, our CrtSPARTA–gDNA–tDNA (20-nt match) complex structure reveals that while not recognizing the 5′-P group of gDNA, it accommodates instead for binding with tDNA's 3′-OH group. The second nucleotide, rather than the first, at the 5′-terminal of gDNA undergoes a conformational change and forms stacking interactions with the side chains of residues Tyr148 and His207 within the MID domain, thereby enhancing the stability of gDNA binding at its 5′-terminal. Furthermore, the 5′-P of the first nucleotide in gDNA is distantly positioned from the conserved MID binding pocket and establishes a hydrogen bond with residue Arg72 within this domain. The 3′-OH group of tDNA binds to the MID binding pocket through coordination and stabilization facilitated by Mg^{2+}, while recognition of the 5′-P group in gDNA by the MID domain occurs independently of Mg^{2+}. These observations highlight that Mg^{2+} plays an essential role in recognizing 5′-P groups specifically in gRNA but not in gDNA.

The distinct recognition model of gRNA and gDNA may be associated with the differential NADase activity of the SPARTA complex. A previous study investigated the NADase activity of MapSPARTA and CrtSPARTA using both gRNA and gDNA with a 5′-P group or 5′-OH group [20]. While the gRNA with a 5′-P group exhibits superior activation efficiency compared to the gDNA with a 5′-P group, the gRNA with a 5′-OH group displays lower activation efficiency than its corresponding gDNA counterpart. Furthermore, the NADase activity of SPARTA is significantly enhanced by utilizing a 5′-P guided RNA rather than a 5′-OH guided RNA. However, there is minimal disparity in NADase activity when employing either a 5′-P guided DNA or a 5′-OH guided DNA [20]. These observations indicate that SPARTA demonstrates distinct recognition mechanisms for phosphate group in gRNA and hydroxyl group in gDNA, which aligns with our structural findings. Notably, our binding assays indicate

that there is no significant difference in binding affinity between SPARTA and 5′-P DNA versus SPARTA and 5′-OH DNA (Supplementary Fig. S9). It is plausible that the binding affinity between SPARTA and DNA is influenced not only by the 5′- or 3′-end phosphate or hydroxyl groups but also by the phosphate groups and bases of nucleotides in other regions of DNA. This conclusion is supported by structural observations showing multiple interactions between SPARTA and various nucleotides along the DNA strand.

The SEC results demonstrate that the gRNA significantly promotes the oligomerization of SPARTA in the presence of tDNA (Supplementary Fig. S10). In contrast, the gDNA does not induce noticeable oligomerization of SPARTA under the same conditions. These differences in the oligomerization of SPARTA may be attributed to distinct recognition patterns of SPARTA toward gDNA and gRNA. Our NADase activity assays indicate that the RNA-guided NADase activity of SPARTA is substantially higher than its DNA-guided NADase activity. Given that oligomerization is crucial for SPARTA's catalytic function, the greater extent of oligomerization observed with RNA-guided SPARTA compared to DNA-guided SPARTA may explain the higher RNA-guided NADase activity relative to the DNA-guided activity.

The prokaryotic Ago proteins can be classified into three major clades: short pAgos, long-A pAgos, and long-B pAgos [17, 18, 43]. Among these, four long pAgos from *Thermus thermophiles* (Tt), *Methanocaldococcus jannaschii* (Mj), *Clostridium butyricum* (Cb), and *Pyrococcus furiosus* (Pf) have been identified as DNA-guided nucleases that target and cleave DNA or RNA [9, 33–36, 44]. Structural studies have revealed that the recognition and coordination of gDNA in the MID binding pocket of these four pAgos require the presence of a 5′-P group and Mg^{2+} ions (Supplementary Fig. S8A and B) [2, 32, 34, 36, 44]. *Pseudooceanicola lipolyticus* pAgo (PliAgo) is also a DNA-guided long pAgo, however, its MID binding pocket exhibiting independence from divalent cations and the canonical set of contacts for 5′-end interactions [8]. On the other hand, CrtAgo belongs to the class of short pAgo proteins, which also function as DNA-guided effector. However, it utilizes the MID binding pocket to anchor the 3′-OH of tDNA instead of the 5′-P of gDNA, thereby demonstrating distinct recognition models for gDNA and tDNA compared to the long pAgos. Numerous long and short pAgos have been identified as RNA-guided immune systems that rely on a 5′-P group for gRNA recognition (Supplementary Fig. S8C and D) [23–29, 31, 42, 45, 46]. The recognition of the 5′-OH of gRNA by *Marinitoga piezophila* (Mp) Ago has been elucidated in a recent study [30, 47]. However, it is noteworthy that the MID pocket of MpAgo exhibits distinct characteristics in comparison to CrtSPARTA's recognition of the 3′-OH in tDNA, as it does not require Mg^{2+} coordination for 5′-OH binding.

The requirement for target length in guide recognition and effector enzyme activation is commonly observed in short pAgos, long pAgos, and CRISPR–Cas enzymes [2, 48–50]. Although short pAgos exhibit a similar target length requirement (usually 13–18 nt) to long Agos and CRISPR–Cas enzymes for effector enzyme activation, they may possess distinct mechanisms of target recognition and enzyme activation. In contrast to long pAgos and CRISPR–Cas enzymes, the target length does not significantly affect the recognition of gRNA or gDNA. However, it does influence the conformational changes of the effector enzyme required for forming an active state. Our structural data suggest that in short pAgo complexes, the tDNA length may impact its recognition by gDNA and effector enzyme, thereby affecting their activity.

The discrepancy in the MID domain's recognition of the 5′-end of 16-nt tDNA and the 3′-end of 21-nt tDNA may be attributed to the different lengths of tDNA. In the 21-nt tDNA-bound SPARTA complex, nucleotides T16–G20 of gDNA form base pairs with nucleotides A16′–C20′ of tDNA. Conversely, in the 16-nt tDNA-bound SPARTA complex, nucleotides T16–G20 of gDNA remain in a single-stranded form, which can be specifically recognized by SPARTA. Specifically, nucleotide T16 forms multiple hydrophobic bonds and stacking interactions with SPARTA, which subsequently trigger specific interactions between nucleotide G15 of gDNA and SPARTA. These interactions enable SPARTA to specifically recognize the 5′-end rather than the 3′-end of the 16-nt tDNA.

Notably, the guide and target sequences utilized in our study are identical to those previously reported for structural and functional studies [20, 24, 28]. Despite using the same target sequence, we observed significant structural differences between the SPARTA–gDNA–tDNA complex and the SPARTA–gRNA–tDNA complex. Specifically, in the SPARTA–gRNA–tDNA complex, the gRNA is bound to the MID domain, whereas in the SPARTA–gDNA–tDNA complex, the tDNA is bound to the MID domain. To better illustrate these structural differences, we refer to the DNA bound to the MID domain as tDNA. However, it is widely accepted that the guide DNA/RNA is recognized by the MID domain of Ago. Therefore, the DNA bound to the MID domain should be designated as gDNA, while its complementary strand should be referred to as tDNA.

Furthermore, since the determined 16-nt and 21-nt tDNA-bound SPARTA complex structures are crystal structures, it is possible that the observed differences in the relative orientations of gDNA and tDNA in our SPARTA complex structure, compared to those in previously determined SPARTA complexes containing gRNA and tDNA, may be a crystallographic artifact without functional significance. To address this concern, we also attempted to determine the structures of SPARTA complexes bound to different lengths of tDNA and successfully solved the crystal structures of SPARTA in complex with gDNA and 19-nt or 20-nt tDNA (Supplementary Fig. S11 and Supplementary Table S1). Although these structures exhibit lower resolution, their unit cell parameters and space groups are consistent with those of the 21-nt tDNA-bound complex, suggesting that the three complexes (19-nt, 20-nt, and 21-nt tDNAs) likely share similar structural features.

In conclusion, our structural analysis of gDNA and tDNAs with varying lengths, along with accompanying biochemical data, provides valuable insights into the recognition mechanism between gDNA and tDNA. This facilitates a better understanding of the activation mechanism of DNA-guided SPARTA and its role in prokaryotic immunity against invading DNA.

Acknowledgements

We are grateful to the staff of the BL-18U1, BL-19U1, BL-10U2, and BL02U1 beamlines at the National Center for Protein Sciences Shanghai (NCPSS) at Shanghai Synchrotron Radiation Facility (SSRF). We thank Y. Wu, J. Hu, and W. Zheng for technical support.

Author contributions: Rong Hu (Conceptualization, Data curation, Formal analysis, Investigation. Methodology, Writing—original draft), Chenmin Guo (Investigation, Visualization, Methodology), Xiaotian Liu (Investigation, Visualization, Methodology), Yuanyi Lin (Investigation, Methodology), Zixin Yang (Investigation, Methodology), Zhixin Li (Investigation, Methodology), Ye Yang (Investigation, Methodology), Erman Ma (Investigation, Methodology), Yiyi Li (Investigation, Methodology), Jiyun Chen (Conceptualization, Data curation, Project administration, Funding acquisition, Writing—review and editing), and Liang Liu (Conceptualization, Data curation, Formal analysis, Supervision, Funding acquisition, Investigation, Visualization, Methodology, Writing—original draft, Project administration, Writing—review and editing)

Supplementary data

Supplementary data is available at NAR online.

Conflict of interest

None declared.

Funding

This work was supported by the National Natural Science Foundation of China (32171286, 32371346, 32022047 to L.L., and 32301007 to J.C.); Natural Science Foundation of Xiamen, China (3502Z20227020 to J.C.); Natural Science Foundation of Fujian Province (2024J011007, 2023J01023 to L.L., and 2023J05008 to J.C.). Funding to pay the Open Access publication charges for this article was provided by the National Natural Science Foundation of China (32301007 to J.C.) and the Natural Science Foundation of Fujian Province (2024J011007 to L.L.).

Data availability

The atomic models of the Apo-CrtSPARTA complex, the Crt–SPARTA–gDNA–target21 DNA complex, the Crt–SPARTA–gDNA–target20 DNA complex, the Crt–SPARTA–gDNA–target19 DNA complex, and the Crt–SPARTA–gDNA–target16 DNA complex have been deposited into the Protein Data Bank with accession codes 8Z8Y, 8Z92, 9L9X, 9L9W, and 8Z96. Publicly available protein additional atomic accession codes in the study: 4NCA (TtAgo), 6QZK (CbAgo), 6DBP (RsAgo), 8PVV (AfAgo), 8IFK (CrtSPARTA–gRNA–tDNA), 8SQU (MapSPARTA–gRNA–tDNA), 8JPX (PfAgo), 3HM9 (TtAgo), 5I4A (MpAgo), 5G5T (MjAgo), and 5UX0 (MpAgo).

References

1. Peters L, Meister G. Argonaute proteins: mediators of RNA silencing. *Cell* 2007;**26**:1856–71. https://doi.org/10.1016/j.molcel.2007.05.001
2. Wang Y, Juranek S, Li H *et al.* Nucleation, propagation and cleavage of target RNAs in Ago silencing complexes. *Nature* 2009;**461**:754–61. https://doi.org/10.1038/nature08434
3. Iwakawa HO, Tomari Y. Life of RISC: formation, action, and degradation of RNA-induced silencing complex. *Mol Cell* 2022;**82**:30–43. https://doi.org/10.1016/j.molcel.2021.11.026
4. Wang L, Xie X, Lv B *et al.* A bacterial Argonaute with efficient DNA and RNA cleavage activity guided by small DNA and RNA. *Cell Rep* 2022;**41**:111533. https://doi.org/10.1016/j.celrep.2022.111533
5. Kim SY, Jung Y, Lim D. Argonaute system of *Kordia jejudonensis* is a heterodimeric nucleic acid-guided nuclease. *Biochem Biophys Res Commun* 2020;**525**:755–8. https://doi.org/10.1016/j.bbrc.2020.02.145
6. Huang S, Wang K, Mayo SL. Genome manipulation by guide-directed Argonaute cleavage. *Nucleic Acids Res* 2023;**51**:4078–85. https://doi.org/10.1093/nar/gkad188
7. Ji-Joon Song SKS, Hannon GJ, Joshua-Tor L. Crystal structure of Argonaute and its implications for RISCS licer activity. *Science* 2004;**305**:1434–7. https://doi.org/10.1126/science.1102514
8. Lisitskaya L, Shin Y, Agapov A *et al.* Programmable RNA targeting by bacterial Argonaute nucleases with unconventional guide binding and cleavage specificity. *Nat Commun* 2022;**13**:4624. https://doi.org/10.1038/s41467-022-32079-5
9. Swarts DC, Jore MM, Westra ER *et al.* DNA-guided DNA interference by a prokaryotic Argonaute. *Nature* 2014;**507**:258–61. https://doi.org/10.1038/nature12971
10. Sashital DG. Prokaryotic Argonaute uses an all-in-one mechanism to provide host defense. *Mol Cell* 2017;**65**:957–8. https://doi.org/10.1016/j.molcel.2017.03.002
11. Kuzmenko A, Oguienko A, Esyunina D *et al.* DNA targeting and interference by a bacterial Argonaute nuclease. *Nature* 2020;**587**:632–7. https://doi.org/10.1038/s41586-020-2605-1
12. Zeng Z, Chen Y, Pinilla-Redondo R *et al.* A short prokaryotic Argonaute activates membrane effector to confer antiviral defense. *Cell Host Microbe* 2022;**30**:930–43. https://doi.org/10.1016/j.chom.2022.04.015
13. Zaremba M, Dakineviciene D, Golovinas E *et al.* Short prokaryotic Argonautes provide defence against incoming mobile genetic elements through NAD^+ depletion. *Nat Microbiol* 2022;**7**:1857–69. https://doi.org/10.1038/s41564-022-01239-0
14. Lei J, Sheng G, Cheung PP *et al.* Two symmetric arginine residues play distinct roles in *Thermus thermophilus* Argonaute DNA guide strand-mediated DNA target cleavage. *Proc Natl Acad Sci USA* 2019;**116**:845–53. https://doi.org/10.1073/pnas.1817041116
15. Olovnikov I, Chan K, Sachidanandam R *et al.* Bacterial Argonaute samples the transcriptome to identify foreign DNA. *Mol Cell* 2013;**51**:594–605. https://doi.org/10.1016/j.molcel.2013.08.014
16. Makarova KS, Wolf YI, van der Oost J *et al.* Prokaryotic homologs of Argonaute proteins are predicted to function as key components of a novel system of defense against mobile genetic elements. *Biol Direct* 2009;**4**:29. https://doi.org/10.1186/1745-6150-4-29
17. Sergei Ryazansky AK, Aravina AA. The expanded universe of prokaryotic Argonaute proteins. *mBio* 2018;**9**:1–20.
18. Koopal B, Mutte SK, Swarts DC. A long look at short prokaryotic Argonautes. *Trends Cell Biol* 2023;**33**:605–18. https://doi.org/10.1016/j.tcb.2022.10.005
19. Swarts DC, Makarova K, Wang Y *et al.* The evolutionary journey of Argonaute proteins. *Nat Struct Mol Biol* 2014;**21**:743–53.https://doi.org/10.1038/nsmb.2879
20. Koopal B, Potocnik A, Mutte SK *et al.* Short prokaryotic Argonaute systems trigger cell death upon detection of invading DNA. *Cell* 2022;**185**:1471–86. https://doi.org/10.1016/j.cell.2022.03.012
21. Zhen X, Xu X, Ye L *et al.* Structural basis of antiphage immunity generated by a prokaryotic Argonaute-associated SPARSA system. *Nat Commun* 2024;**15**:450. https://doi.org/10.1038/s41467-023-44660-7
22. Garb J, Lopatina A, Bernheim A *et al.* Multiple phage resistance systems inhibit infection via SIR2-dependent NAD^+ depletion. *Nat Microbiol* 2022;**7**:1849–56. https://doi.org/10.1038/s41564-022-01207-8
23. Finocchio G, Koopal B, Potocnik A *et al.* Target DNA-dependent activation mechanism of the prokaryotic immune system SPARTA. *Nucleic Acids Res* 2024;**52**:2012–29. https://doi.org/10.1093/nar/gkad1248

24. Gao X, Shang K, Zhu K *et al.* Nucleic-acid-triggered NADase activation of a short prokaryotic Argonaute. *Nature* 2024;**625**:822–31. https://doi.org/10.1038/s41586-023-06665-6
25. Guo L, Huang P, Li Z *et al.* Auto-inhibition and activation of a short Argonaute-associated TIR-APAZ defense system. *Nat Chem Biol* 2024;**20**:512–20. https://doi.org/10.1038/s41589-023-01478-0
26. Guo M, Zhu Y, Lin Z *et al.* Cryo-EM structure of the ssDNA-activated SPARTA complex. *Cell Res* 2023;**33**:731–4. https://doi.org/10.1038/s41422-023-00850-y
27. Shen Z, Yang XY, Xia S *et al.* Oligomerization-mediated activation of a short prokaryotic Argonaute. *Nature* 2023;**621**:154–61. https://doi.org/10.1038/s41586-023-06456-z
28. Wang X, Li X, Yu G *et al.* Structural insights into mechanisms of Argonaute protein-associated NADase activation in bacterial immunity. *Cell Res* 2023;**33**:699–711. https://doi.org/10.1038/s41422-023-00839-7
29. Zhang JT, Wei XY, Cui N *et al.* Target ssDNA activates the NADase activity of prokaryotic SPARTA immune system. *Nat Chem Biol* 2024;**20**:503–11. https://doi.org/10.1038/s41589-023-01479-z
30. Doxzen KW, Doudna JA. DNA recognition by an RNA-guided bacterial Argonaute. *PLoS One* 2017;**12**:e0177097. https://doi.org/10.1371/journal.pone.0177097
31. Liu Y, Esyunina D, Olovnikov I *et al.* Accommodation of helical imperfections in *Rhodobacter sphaeroides* Argonaute ternary complexes with guide RNA and target DNA. *Cell Rep* 2018;**24**:453–62. https://doi.org/10.1016/j.celrep.2018.06.021
32. Swarts DC, Hegge JW, Hinojo I *et al.* Argonaute of the archaeon *Pyrococcus furiosus* is a DNA-guided nuclease that targets cognate DNA. *Nucleic Acids Res* 2015;**43**:5120–9. https://doi.org/10.1093/nar/gkv415
33. Zander A, Willkomm S, Ofer S *et al.* Guide-independent DNA cleavage by archaeal Argonaute from *Methanocaldococcus jannaschii*. *Nat Microbiol* 2017;**2**:17034. https://doi.org/10.1038/nmicrobiol.2017.34
34. Willkomm S, Oellig CA, Zander A *et al.* Structural and mechanistic insights into an archaeal DNA-guided Argonaute protein. *Nat Microbiol* 2017;**2**:17035. https://doi.org/10.1038/nmicrobiol.2017.35
35. Sheng G, Zhao H, Wang J *et al.* Structure-based cleavage mechanism of *Thermus thermophilus* Argonaute DNA guide strand-mediated DNA target cleavage. *Proc Natl Acad Sci USA* 2014;**111**:652–7. https://doi.org/10.1073/pnas.1321032111
36. Hegge JW, Swarts DC, Chandradoss SD *et al.* DNA-guided DNA cleavage at moderate temperatures by *Clostridium butyricum* Argonaute. *Nucleic Acids Res* 2019;**47**:5809–21. https://doi.org/10.1093/nar/gkz306
37. Otwinowski Z, Minor W. Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol* 1997;**276**:307–26. https://doi.org/10.1016/S0076-6879(97)76066-X
38. McCoy AJ, Grosse-Kunstleve RW, Adams PD *et al.* Phaser crystallographic software. *J Appl Crystallogr* 2007;**40**:658–74. https://doi.org/10.1107/S0021889807021206
39. Emsley P, Lohkamp B, Scott WG *et al.* Features and development of Coot. *Acta Crystallogr D Biol Crystallogr* 2010;**66**:486–501. https://doi.org/10.1107/S0907444910007493
40. Adams P, Grosse-Kunstleve R, Hung L *et al.* PHENIX: building new software for automated crystallographic structure determination. *Acta Crystallogr* 2002;**58**:1948–54.
41. Kottur J, Malik R, Aggarwal AK. Nucleic acid mediated activation of a short prokaryotic Argonaute immune system. *Nat Commun* 2024;**15**:4852. https://doi.org/10.1038/s41467-024-49271-4
42. Dongchun Ni XL, Stahlberg H, Ekundayo B. Activation mechanism of a short Argonaute-TIR prokaryotic immune system. *Sci Adv* 2023;**9**:eadh9002.
43. Hegge JW, Swarts DC, van der Oost J. Prokaryotic Argonaute proteins: novel genome-editing tools? *Nat Rev Microbiol* 2018;**16**:5–11. https://doi.org/10.1038/nrmicro.2017.73
44. Wang L, Chen W, Zhang C *et al.* Molecular mechanism for target recognition, dimerization, and activation of *Pyrococcus furiosus* Argonaute. *Mol Cell* 2024;**84**:675–86. https://doi.org/10.1016/j.molcel.2024.01.004
45. Manakova E, Golovinas E, Poceviciute R *et al.* The missing part: the *Archaeoglobus fulgidus* Argonaute forms a functional heterodimer with an N-L1-L2 domain protein. *Nucleic Acids Res* 2024;**52**:2530–45. https://doi.org/10.1093/nar/gkad1241
46. Bastiaanssen C, Bobadilla Ugarte P, Kim K *et al.* RNA-guided RNA silencing by an Asgard archaeal Argonaute. *Nat Commun* 2024;**15**:5499. https://doi.org/10.1038/s41467-024-49452-1
47. Kaya E, Doxzen KW, Knoll KR *et al.* A bacterial Argonaute with noncanonical guide RNA specificity. *Proc Natl Acad Sci USA* 2016;**113**:4057–62. https://doi.org/10.1073/pnas.1524385113
48. Liu L, Li X, Ma J *et al.* The molecular architecture for RNA-guided RNA cleavage by Cas13a. *Cell* 2017;**170**:714–26. https://doi.org/10.1016/j.cell.2017.06.050
49. Chen J, Chen Y, Huang L *et al. Trans*-nuclease activity of Cas9 activated by DNA or RNA target binding. *Nat Biotechnol* 2024. https://doi.org/10.1038/s41587-024-02255-7
50. Stella S, Mesa P, Thomsen J *et al.* Conformational activation promotes CRISPR–Cas12a catalysis and resetting of the endonuclease activity. *Cell* 2018;**175**:1856–71. https://doi.org/10.1016/j.cell.2018.10.045

Received: September 2, 2024. **Revised:** January 17, 2025. **Editorial Decision:** January 31, 2025. **Accepted:** February 5, 2025