Target DNA-dependent activation mechanism of the prokaryotic immune system SPARTA

Finocchio, Giada ; Koopal, Balwina ; Potocnik, Ana ; Heijstek, Clint ; Westphal, Adrie H ; Jinek, Martin ; Swarts, Daan C

DOI: https://doi.org/10.1093/nar/gkad1248

Nucleic Acids Research, 2024, **52**, 2012–2029
https://doi.org/10.1093/nar/gkad1248
Advance access publication date: 15 January 2024
Structural Biology

Target DNA-dependent activation mechanism of the prokaryotic immune system SPARTA

Giada Finocchio [1,†], Balwina Koopal [2,†], Ana Potocnik [2], Clint Heijstek [2], Adrie H. Westphal [2], Martin Jinek [1,*] and Daan C. Swarts [2,*]

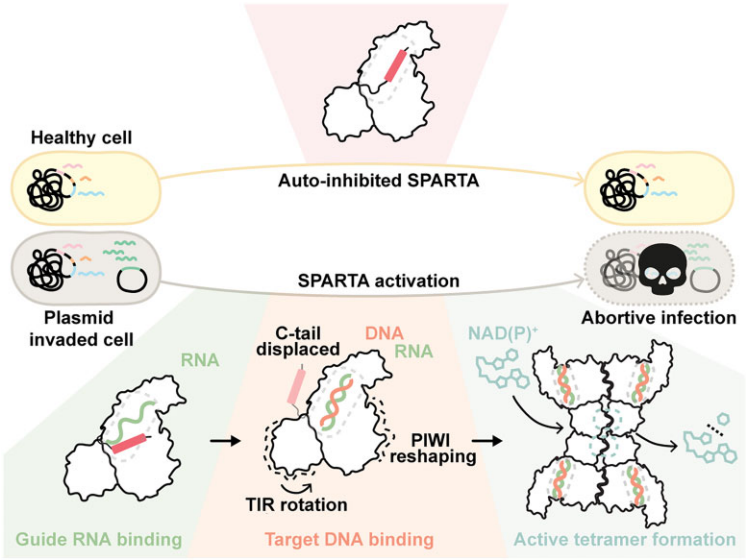
[1]Department of Biochemistry, University of Zurich, 8057 Zurich, Switzerland
[2]Laboratory of Biochemistry, Wageningen University, 6708 WE Wageningen, the Netherlands

[*]To whom correspondence should be addressed. Tel: +31 6 1432 7706; Email: daan.swarts@wur.nl
Correspondence may also be address to Martin Jinek. Tel: +41 44 635 5572; Email: jinek@bioc.uzh.ch
[†]The first two authors should be regarded as Joint First Authors.

Abstract

In both prokaryotic and eukaryotic innate immune systems, TIR domains function as NADases that degrade the key metabolite NAD^+ or generate signaling molecules. Catalytic activation of TIR domains requires oligomerization, but how this is achieved varies in distinct immune systems. In the Short prokaryotic Argonaute (pAgo)/TIR-APAZ (SPARTA) immune system, TIR NADase activity is triggered upon guide RNA-mediated recognition of invading DNA by an unknown mechanism. Here, we describe cryo-EM structures of SPARTA in the inactive monomeric and target DNA-activated tetrameric states. The monomeric SPARTA structure reveals that in the absence of target DNA, a C-terminal tail of TIR-APAZ occupies the nucleic acid binding cleft formed by the pAgo and TIR-APAZ subunits, inhibiting SPARTA activation. In the active tetrameric SPARTA complex, guide RNA-mediated target DNA binding displaces the C-terminal tail and induces conformational changes in pAgo that facilitate SPARTA-SPARTA dimerization. Concurrent release and rotation of one TIR domain allow it to form a composite NADase catalytic site with the other TIR domain within the dimer, and generate a self-complementary interface that mediates cooperative tetramerization. Combined, this study provides critical insights into the structural architecture of SPARTA and the molecular mechanism underlying target DNA-dependent oligomerization and catalytic activation.

Graphical abstract



Introduction

Argonaute proteins comprise a diverse family of proteins that confer immunity in all domains of life (1–3). Eukaryotic Argonautes (eAgos) associate with short (15–30 nt) RNA guides to target complementary RNA sequences. Catalytically active 'slicing' eAgos can cleave the targeted RNA, for example during antiviral defense, while catalytically inactive eAgos recruit accessory proteins to the targeted RNA to elicit transcriptional or post-transcriptional gene silencing (4). As such, eAgos can regulate native gene expression but also interfere with

Received: September 25, 2023. **Revised:** December 14, 2023. **Editorial Decision:** December 15, 2023. **Accepted:** December 19, 2023

virus and transposon propagation (1). In contrast, archaea and bacteria encode a large variety of prokaryotic Argonaute proteins (pAgos) that, besides reported involvement in DNA replication (5,6) and DNA repair (7), mainly act as prokaryotic immune systems that interfere with invading plasmid and bacteriophage DNA (8–14).

Based on their phylogeny, pAgos are subdivided in long-A, long-B, and short pAgos (15,16). eAgos and long pAgos share a bilobed structural architecture (17). The MID-PIWI lobe consists of the middle (MID) and p-element induced wimpy testis (PIWI) domains (18), and is responsible for guide 5′ end binding and guide-mediated target binding (18–21). This lobe is connected through the linker 2 (L2) motif to the N-terminal lobe, which is comprised of the N-terminal (N) domain, the linker 1 (L1) motif and the PAZ (PIWI-ARGONAUTE-ZWILLE) domain. These domains have been implicated in guide loading, guide 3′ end binding, target binding specificity, and Ago protein turnover (19,21–23). In contrast to eAgos and long pAgos, short pAgos are comprised of the MID and PIWI domains only, and all short pAgos are catalytically inactive (15). Short pAgos are encoded in operons together with (and sometimes fused to) proteins comprising an Analog of PAZ (APAZ) domain, which displays homology to the N domain of long Agos (16). APAZ domains are fused to one of various 'effector' domains including (but not limited to) Sirtuin (Sir2), RecG/DHS-like, Mrr-like, DUF4365, and Toll/interleukin-1 receptor (TIR) domains (15,16). Upon guide RNA-mediated recognition of invading plasmid or bacteriophage DNA, short pAgo/TIR-APAZ (SPARTA), short pAgo/Sir2-APAZ (SPARSA) and short pAgo/DUF4365 (SPARDA) systems trigger abortive infection (12–14). SPARTA and SPARSA systems achieve abortive infection through enzymatic degradation of the central metabolite nicotinamide adenine dinucleotide (NAD^+) (13,14).

The SPARTA TIR domain is homologous to other TIR effector domains found in prokaryotic, plant, and animal immune systems (13,24–34). In these immune systems, TIR domains function in conjunction with variable 'sensor' proteins that assemble into higher-order structures upon recognition of a pathogen-derived molecular signature, thereby facilitating TIR domain oligomerization (24,26,29,32,34,35). Not all TIR domains have an enzymatic function; in scaffolding TIR domains, TIR oligomerization results in formation of signalosomes (36). In enzymatic TIR NADases, TIR oligomerization triggers catalytic activation whereby head-to-tail TIR domain interactions allow one TIR domain to insert its flexible 'BB-loop' into a pocket of a second TIR domain. This forms a composite catalytic site capable of NAD^+ binding and hydrolysis (26,29,32,34,35). Prokaryotic TIR effectors degrade NAD^+ either to generate secondary messenger molecules (27) or to trigger cell death and consequentially prevent invader propagation (13,28–31). The structural basis for TIR domain oligomerization has been described for various eukaryotic and prokaryotic immune systems (24,26,29,32,35). Notably, the interaction interfaces underlying TIR oligomerization are variable for distinct immune systems, which gives rise to various TIR oligomerization modes. In SPARTA systems, the short pAgo and TIR-APAZ form an inactive heterodimeric pAgo:TIR-APAZ complex with a 1:1 stoichiometry (13), hereafter referred to as the monomeric SPARTA complex. Guide RNA-mediated target DNA binding results in the tetramerization of four SPARTA complexes, yielding a pAgo:TIR-APAZ:guide RNA:target DNA complex with a 4:4:4:4 stoichiometry (13). This suggests that the TIR domains are activated upon oligomerization. However, the molecular mechanisms underlying SPARTA activation remain unknown.

Here, we report cryo-EM structures of the *Bacillales bacterium* (Bab)SPARTA system in its inactive monomeric and activated tetrameric forms. The structure of the monomeric BabSPARTA complex reveals an autoinhibited state in which the nucleic acid binding channel formed by the interaction of pAgo and TIR-APAZ subunits is occupied by the C-terminal tail (Ct) of TIR-APAZ. Interactions of the Ct with APAZ inhibit guide/target binding, while interactions between the MID and TIR domains keep TIR in an inhibited conformation. The structure of the activated BabSPARTA tetramer reveals that RNA-guided target DNA binding displaces the Ct and induces conformational changes that result in pAgo-mediated dimerization. This in turn releases one TIR domain and enables dimerization with the TIR domain of the other SPARTA protomer, generating a composite NADase catalytic site. The TIR-TIR dimer is further stabilized through cooperative interactions with the TIR-TIR dimer of another activated SPARTA dimer, resulting in the formation of a catalytically active tetramer. Together, these insights reveal the mechanisms underlying the RNA-guided target DNA recognition and catalytic activation of SPARTA systems.

Materials and methods

Plasmid availability

MapSPARTA expression vector (pBK086) and BabSPARTA expression vector (pAP007) have been deposited to AddGene (Plasmid #183145 and #206809). Other plasmids are available upon request.

Plasmid cloning for *in vivo* assays and protein production

All cloning procedures were performed as described previously (1) The genes encoding BabSPARTA (Coding contig: JADBKC010000166.1; BabAgo: MBE3571069.1.1; BabTIR-APAZ: MBE3571068.1) were ordered from Twist Bioscience as an artificial operon. This operon was cloned in the pET-His6-MBP-TEV-LIC cloning vector. Bacterial Artificial Chromosomes (BACs) encoding MapSPARTA or the catalytically inactive mutant $E77A^{TIR\text{-}APAZ}$ were previously constructed and described (1). All plasmids used in this study are listed on Mendeley Data (https://data.mendeley.com/datasets/w94r9mgx2d/draft?a=a848ee5e-14d9-4650-90bb-c79f218ecf30).

SPARTA protein expression and purification

Proteins were purified as described previously (1).

Cryo-EM sample preparation and data collection: BabSPARTA-gRNA

Purified BabSPARTA was mixed with a 5′-phosphorylated RNA guide (oBK084, 5′-UGACGGCUCUAAUCUAUUAGU-3′) in assembly buffer (20 mM HEPES pH 7.5, 125 mM KCl, 2 mM $MgCl_2$). The final sample contained 20 μM BabSPARTA and 30 μM gRNA (1:1.5 molar ratio) in a total volume of 80 μl. The volume was incubated at 50°C for 1 hour, centrifuged at 18000 rpm for 10′ at room temperature and directly used for cryo-EM grid preparation.

3.5 μl of sample was applied to a freshly glow discharged 200-mesh Au R1.2/1.3 grid (Quantifoil Micro Tools), incubated for 5 s, blotted for 6 s at 100% humidity, 4 °C, plunge frozen in liquid ethane (using a Vitrobot Mark IV plunger, FEI) and stored in liquid nitrogen. cryo-EM data collection was performed on a FEI Titan Krios G3i microscope (University of Zurich, Switzerland) operated at 300 kV and equipped with a Gatan K3 direct electron detector in super-resolution counting mode. A total of 14 332 movies were recorded at 130 000× magnification, resulting in a super-resolution pixel size of 0.325 Å. Each movie comprised 36 subframes with a total dose of 58.005 $e^-/Å^2$. Data acquisition was performed with EPU Automated Data Acquisition Software for Single Particle Analysis (ThermoFisher Scientific) with three shots per hole at -1.0 mm to -2.4 mm defocus (0.2 mm steps).

Data processing and model building: BabSPARTA-gRNA

The collected exposures were processed in cryoSPARC (v.4.2) (37). Patch Motion Correction and Patch CTF Correction were used to align and correct the imported 14332 movies. Blob picker (minimum particle diameter 90 Å, maximum particle diameter 180 Å) was used to select particles that were further extracted (extraction box size 360 pix, Fourier-cropped to box size 120 pix) and classified in 50 classes using 2D Classification. 22 classes, corresponding to 2 815 212 particles, were selected and given as input to a 2-classes Ab-Initio Reconstruction. One of the two classes, corresponding to 1 733 366 particles, was further classified using a 5-classes Heterogeneous Refinement. One of the 5 classes, corresponding to 904 721 particles, was used as input for Non-uniform Refinement, resulting in a 2.73 Å (GSFSC resolution) density. A summary of the processing workflow is shown in the Supplementary Figure S6.

An initial model of BabSPARTA was generated using AlphaFold2 ColabFold (38). The model was manually docked as rigid body in the BabSPARTA-gRNA cryo-EM density map using UCSF ChimeraX (39), followed by real space fitting with the Fit in Map function. The model was subjected to manual refinement against the corresponding cryo-EM map using the software Coot (40) and real space refine in Phenix (41). Secondary structure restraints, side chain rotamer restraints and Ramachandran restraints were used. The refinement statistics are listed on Mendeley Data (https://data.mendeley.com/datasets/w94r9mgx2d/draft?a=a848ee5e-14d9-4650-90bb-c79f218ecf30). The final model comprises one copy of BabTIR-APAZ$^{1\text{-}435}$(chain A) and one copy of BabAgo$^{1–151\ 204–507}$(chain B). Partial density for nucleotide 2 of the RNA guide was visible in the map, but not interpretable, probably due to partial occupancy. The molecule was therefore not modelled. Figures preparation of model and map was performed using UCSF ChimeraX (39).

Cryo-EM sample preparation and data collection: BabSPARTA-gRNA-tDNA

Purified BabSPARTA was mixed with a 5′-phosphorylated guide RNA (oBK084, UGACGGCUCUAAUCUAUUAGU) in assembly buffer (20 mM HEPES pH 7.5, 125 mM KCl, 2 mM $MgCl_2$). The mixture was incubated at 37°C for 30′ before adding the target DNA (oBK382, CAACTAATAGATTAGAGCCGTCAAT). The final sample contained 14.5 μM BabSPARTA, 22.5 μM gRNA and 22.5 μM tDNA (1:1.5:1.5 molar ratio) in a total volume of 100 μl. The volume was incubated at 50°C for 1 hour, centrifuged at 18 000 rpm for 10′ at room temperature and analyzed by size exclusion chromatography (Superdex 200 Increase 15/500 GL column, ÄKTA pure™ micro system). The fractions corresponding to the peak at the elution volume ∼1.3 mL were pooled together (total volume 300 μl) and concentrated to 30 μl using a 100 000 Da molecular weight cut-off centrifugal filter (Merck Millipore). *n*-Octyl-β-D-glucoside was freshly added to the sample to a final concentration of 0.07% m/V.

2.4 μl of sample was applied to a freshly glow discharged 200-mesh Cu + 2 nm C R1.2/1.3 grid (Quantifoil Micro Tools), blotted for 6 s at 100% humidity, 4°C, plunge frozen in liquid ethane (using a Vitrobot Mark IV plunger, FEI) and stored in liquid nitrogen. Cryo-EM data collection was performed on a FEI Titan Krios G3i microscope (University of Zurich, Switzerland) operated at 300 kV and equipped with a Gatan K3 direct electron detector in super-resolution counting mode. A total of 9041 movies were recorded at 130 000× magnification, resulting in a super-resolution pixel size of 0.325 Å. Each movie comprised 47 subframes with a total dose of 59.347 $e^-/Å^2$. Data acquisition was performed with EPU Automated Data Acquisition Software for Single Particle Analysis (ThermoFisher Scientific) with three shots per hole at –1.0 mm to –2.4 mm defocus (0.2 mm steps).

Data processing and model building: BabSPARTA-gRNA-tDNA

The collected exposures were processed in cryoSPARC (v.4.2) (37). Patch Motion Correction and Patch CTF Correction were used to align and correct the imported 9041 movies. Blob picker (minimum particle diameter 100 Å, maximum particle diameter 300 Å) was used to select particles that were further extracted (extraction box size 640 pixel, Fourier-cropped to box size 128 pixel) and classified in 50 classes using 2D Classification. Eighteen classes, corresponding to 865 048 particles, were selected and given as input to a 2-classes Ab-Initio Reconstruction. One of the two classes, corresponding to 486647 particles, was refined with Non-uniform Refinement. As some parts of the map showed signs of flexibility, particles and mask were used as input for a 3D variability analysis, solving 3 modes and using 5.5 Å as filter resolution. The results were analyzed with 3D Variability Display in cluster output mode with 5 clusters. One of the 5 cluster volumes, corresponding to 60319 particles, was used as input for Non-uniform Refinement, resulting in a 3.09 Å (GSFSC resolution) density. A summary of the processing workflow is shown in the Supplementary Figure S7.

Four copies of the BabSPARTA apo model were manually docked as rigid bodies in the cryo-EM density map of BabSPARTA-gRNA-tDNA using UCSF ChimeraX (39), followed by real space fitting with the Fit in Map function. The DNA and RNA molecules were manually built using the software Coot (40). The model was subjected to manual refinement in Coot (40) against the corresponding cryo-EM map and real space refine in Phenix (41). Secondary structure, side chain rotamer, Ramachandran and nucleic acid restraints (base pair, stacking plane and sugar pucker restraints calculated with the LibG (42) script)

were used. The refinement statistics are listed on Mendeley Data (https://data.mendeley.com/datasets/w94r9mgx2d/draft?a=a848ee5e-14d9-4650-90bb-c79f218ecf30). The final model comprises 4 copies of BabTIR-APAZ$^{1-418}$ (chains A, E, I, M), 4 copies of BabAgo$^{1-152\ 204–289\ 294–507}$ (chains B, F, J, N), four copies of the guide RNA$^{1-20}$ (chains C, G, K, O)-target DNA$^{3-20}$ (chains D, H, L, P) duplex and 4 Mg^{2+} ions (chains Q, R, S, T). Figures preparation of model and map was performed using UCSF ChimeraX (39).

Structure analysis

Structure interfaces were determined using the 'Protein interfaces, surfaces and assemblies' service PISA at the European Bioinformatics Institute. (http://www.ebi.ac.uk/pdbe/prot_int/pistart.html) (43).

Structural alignments were made with the PyMOL (44) function cealign, using truncated models for the domain alignments.

Sequence conservation analysis

Short pAgo and APAZ-domain containing proteins sequences were obtained from previous analyses and alignments described previously (13) and extended with sequences from BabAgo (MBE3571069.1) and BabTIR-APAZ (MBE3571068.1), KjAgo (WP_046756490.1) and KjMrr-APAZ (WP_046756489.1) (45), and NgaSPARDA (NgaAgo and NgaDUF4365-APAZ) (12). In total 285 short pAgos an 255 APAZ-domain containing proteins were used for multiple sequence alignments. Sequences were aligned in Jalview (V2.11.2.7) using MAFFT (46) with E-INS-I settings. Full sequence alignments can be found on Mendeley data (https://data.mendeley.com/datasets/w94r9mgx2d/draft?a=a848ee5e-14d9-4650-90bb-c79f218ecf30). For data visualization, the aligned sequences were limited to a subset of proteins, gaps were hidden, and sequence conservation was visualized together with structural features from the heterodimeric BabSPARTA complex using ESpript 3.0 (https://espript.ibcp.fr).

For C-terminal tail conservation analysis the sequence alignment of clade S1A, S2A, and S2B APAZ-domain was used. All (aligned) sequences up to BabTIR-APAZ residue 416 were removed (i.e. only sequences aligned with the sequence downstream of BabTIR-APAZ residue 417 were included). Of these sequences, length and Asp/Glu content was determined in Excel. C-terminal tail sequences and analysis can be found on Mendeley Data (https://data.mendeley.com/datasets/w94r9mgx2d/draft?a=a848ee5e-14d9-4650-90bb-c79f218ecf30).

Fluorescence polarization assays

A reaction mixture of purified BabSPARTA complex or the Δ417–450$^{TIR-APAZ}$ Ct truncation mutant and a guide RNA guide with a 3′ ATTO532 label was prepared on ice in 96-well half area microplates. The final reaction conditions were 10–600 nM protein, 50 nM guide RNA, 10 mM MES pH 6.5, 125 mM KCl, 2 mM $MgCl_2$ and 0.1% pluronic acid in a final volume of 50 µl. After assembly of the individual reactions, the plate was incubated at room temperature for 30 minutes and transferred to a SH1M2FG plate reader (Biotek). The parallel and perpendicular polarized fluorescence emission intensities, as well as the unpolarized fluorescence emission intensities were measured using an excitation wavelength of 530 nm (bandwidth 25 nm) and emission wavelength of 590 nm (bandwidth 35 nm). From the parallel and perpendicular intensities the anisotropy was calculated for every protein concentration. The measurements at 0 nM protein (six in total) were averaged, and this value was also taken as the anisotropy of the free gRNA (Af). For each replicate, the anisotropy of the bound guide RNA (Ab) was estimated by plotting the anisotropy over protein concentration. Quenching of the ATTO532 label by protein binding was estimated using the unpolarized fluorescence emission intensities. Using the obtained quenching factor, Af, and Ab, the fraction of bound gRNA was calculated for each measurement (47). The data was fitted to a simple model described by the equation derived for two component binding (47). The sequence of the 3′-ATTO532-labeled guide RNA is listed on Mendeley Data (https://data.mendeley.com/datasets/w94r9mgx2d/draft?a=a848ee5e-14d9-4650-90bb-c79f218ecf30). All assays were performed as technical triplicates, and error bars indicate standard deviations.

ε-NAD assays

To assess NADase activity *in vitro* we followed protocols described previously (13). A reaction mixture of purified BabSPARTA complex in SEC buffer, ε-NAD^+ (Nicotinamide 1,N6-ethenoadenine dinucleotide), RNA guide, and 5X reaction buffer (50 mM MES pH 6.5, 375 mM KCl, and 10 mM $MgCl_2$) was prepared on ice in 96-well or 384-well plates. The mixture was incubated at room temperature for 15 min, after which DNA target was added. Unless indicated otherwise, concentrations of each component were 1 µM SPARTA complex, 1 µM guide, 25 µM ε-NAD^+, 10 mM MES pH 6.5, 125 mM KCl and 2 mM $MgCl_2$ in a final volume of 60 or 20 µl.

In the assays to determine the preferred 5′ end nucleotide of the guide, the target concentration was 200 nM, whereas in the assay to determine the temperature optimum of BabSPARTA the target concentration was 1 µM. For chained target experiments, the concentration of the chained target concentration was 100 nM, while the concentration of 5′ or 3′ scrambled chained target as well as the single target was 200 nM. For assessing the effect of Ct truncation *in vitro*, either BabSPARTA or BabSPARTA Ct-truncation mutant Δ417–450$^{TIR-APAZ}$ was used, and target DNA was added to final concentrations ranging from 50 nM to 200 nM.

After addition of the target, the plate was transferred to a SH1M2FG plate reader (Biotek) preheated to 55°C. For kinetic measurements, fluorescence intensity was measured in kinetic mode using an excitation wavelength of 310 nm and emission wavelength of 410 nm. To determine the temperature optimum of BabSPARTA, reaction mixtures were prepared in PCR strips, and after addition of the target the reactions were transferred to a pre-heated PCR block. After 1 h incubation, the reactions were transferred to a 384-well plate and the fluorescence intensity was measured as described for kinetic measurements.

The sequences of all guide and target oligonucleotides used are listed on Mendeley Data (https://data.mendeley.com/datasets/w94r9mgx2d/draft?a=a848ee5e-14d9-4650-90bb-c79f218ecf30). All measurements were corrected with a control without ssDNA target. For assays with chained targets, fluorescence values were normalized to the minimum and maximum of each individual reaction. All assays were performed as technical triplicates, error bars in bar charts and shaded areas in line graphs indicate standard deviations.

In vivo total NAD measurement and growth assays

To assess the effect of MapSPARTA expression (from bacterial artificial chromosome (BAC) pBK138) and mutants thereof (pBK261-pBK273, pBK295-297, and pBK301) on total NAD concentration in *E. coli*, BW25113(DE3) we followed the protocol described by Koopal *et al.* (2022) (13). To activate MapSPARTA, a pUC-based high copy plasmid that encodes mRFP under control of a T7 promoter (pBK145) was included. Before induction of SPARTA expression, 3 mL LB supplemented with 100 μg/ml carbenicillin and 25 μg/ml chloramphenicol was inoculated with 30 μl overnight culture. After 2.5 h of growth at 37°C in a shaker incubator, all cultures were normalized by spinning down the amount of cells corresponding to 0.7 $OD_{600\,nm}$ units. The cells were resuspended in fresh LB containing 100 μg/ml carbenicillin, 25 μg/ml chloramphenicol and 0.2% arabinose to induce expression of MapSPARTA. The cultures were further incubated in a shaker incubator at 37°C. 1h after normalization and induction with arabinose, the cultures were induced with 0.25 mM IPTG to start transcription from pBK145. 4h after normalization and induction of MapSPARTA expression, the NAD concentration was determined as described before (13). At the same timepoint, the same cultures were diluted 1:200 in 200 μl LB in a 96-well plate. The 96-well plate was covered with a gas permeable moisture seal (4titude) and was transferred to a Synergy Neo2 or SH1M2FG plate reader (Biotek) preheated to 37°C. The $OD_{600\,nm}$ was measured in kinetic mode for 12 h. All experiments were performed as biological triplicates, error bars in bar charts and shaded areas in line graphs indicate standard deviations.

Protein pulldown of BabAgo using BabTIR-APAZ

6x-his-MBP-tagged BabSPARTA (pAP007) or mutants thereof (pBK274-pBK290) were transformed in *E. coli* BL21 Star (DE3). Single colonies were used to inoculate 1 ml LB cultures which were incubated o/n at 37°C at 180 RPM. The o/n cultures were used to inoculate 10 ml LB medium which were incubated at 37°C at 180 RPM. When an OD_{600} nm of ∼0.6 was reached, the cultures were incubated on ice for 10 min while the temperature of the incubator was lowered to 18°C. Protein expression was induced by addition of 0.25 mM IPTG, and cultures were further incubated at 18°C and 180 RPM for 16 h. Cells were harvested by centrifugation at 3500 × g at 4°C for 15 min. Cell pellets were resuspended in 0.5 ml amylose wash buffer (500 mM NaCl, protease inhibitors (200 μg/ml AEBSF, 1 μg/ml pepstatin), 20 mM HEPES pH 7.5), transferred to thin-walled 0.5 ml Eppendorf tubes, and lysed in a sonicator bath at 4°C (QSONICA Q700A-220 sonicator; amp 90%, 1 s ON/4 s OFF, 20 min total ON time). The lysates were centrifuged for 10 min at 20000 x g at 4°C. Clarified lysates were incubated for 1 hour at 4°C on a nutator with 40 μl amylose resin slurry (New England Biolabs) equilibrated with amylose wash buffer. After incubation, the resin of each sample was pelleted by centrifugation for 30 seconds at 2000 x g at 4°C, and resuspended in 1 ml fresh amylose wash buffer. The wash step was repeated 3 times. Finally, the resin of each sample was resuspended in 40 μl SDS-PAGE loading buffer and incubated for 5 min at 95°C. The samples were analyzed on SDS-PAGE using CBB staining, and the bands corresponding to MBP-BabTIR-APAZ and BabAgo were quantified using GelAnalyzer (version 19.1; www.gelanalyzer.com). The pull-down experiments were performed as biological triplicates, dots in result graphs represent measurements from individual experiments.

Statistical analyses

For *in vivo* total NAD measurements (4 h after induction), as well $OD_{600\,nm}$ measurements (9 h after induction), statistical analyses were performed in R version 4.1.0. The data in each experiment was subjected to a Shapiro-Wilk test on linear model residuals and Levene's test to confirm normal distribution of the data and homogeneity of variances, respectively. One-way ANOVA was used to test for significant effect of MapSPARTA variant on total NAD and $OD_{600\,nm}$, and Tukey's test was used for post-hoc pairwise comparisons.

In three experiments the assumption of normality was not met, which was solved for two experiments by using a $x^{1.25}$ transformation on the data. This concerns the statistical tests described in Figure 1, Figure 4, Supplementary Figure S7 and Supplementary Figure S11. For one experiment transforming the data was not successful (comparison of $OD_{600\,nm}$ between cells expressing MapSPARTA, $E77A^{TIR\text{-}APAZ}$ and $\Delta 417\text{–}450^{TIR\text{-}APAZ}$, no pUC-$mRFP^{\Delta RBS}$; see Supplementary Figure S7). Therefore, instead of a one-way ANOVA a Kruskal–Wallis rank sum test was used. As this test was not significant, no post-hoc test was performed for this experiment.

Results

Structural basis for short pAgo-APAZ heterodimerization

To gain insights into the molecular mechanism of the SPARTA system, we reconstituted the BabSPARTA complex (Figure 1A) together with a 21-nucleotide guide RNA in the absence of a target DNA for structural analysis by cryogenic electron microscopy. The resulting reconstruction, determined at a resolution of 2.6 Å, reveals a monomeric, binary BabAgo:BabTIR-APAZ SPARTA complex (Figure 1B, C, Supplementary Figure S1, and Table S1). This monomeric SPARTA complex reconstruction lacks interpretable density for the guide RNA (Supplementary Figure S1D), suggesting that inactive SPARTA binds guide RNAs with a low affinity or that SPARTA-bound guide RNAs are conformationally heterogeneous in the absence of target DNA. Alternatively, SPARTA systems might bind preformed guide RNA-target DNA duplexes or rely on other mechanisms for efficient guide RNA loading.

The monomeric BabSPARTA structure reveals that the BabTIR-APAZ subunit wraps around the BabAgo subunit (Figure 1B, C). BabAgo adopts a compact fold with a canonical arrangement of MID and PIWI domains, as previously observed in other pAgo and eAgo proteins (20,48–50) (Supplementary Figure S2). BabTIR-APAZ adopts an extended conformation, with the TIR and APAZ domains contacting the BabAgo MID and PIWI domains, respectively (Figure 1C). Structural superposition of the BabSPARTA complex with the long-A pAgo from *Thermus thermophilus* (TtAgo) reveals that the APAZ domain resembles the N domain and linker segments L1 and L2 of long-A pAgo proteins (Supplementary Figure S2). The TIR-MID domain interface extends over 633 $Å^2$ and features a network of salt-bridge and hydrogen-bonding interactions (Figure 1D). The TIR-MID in-

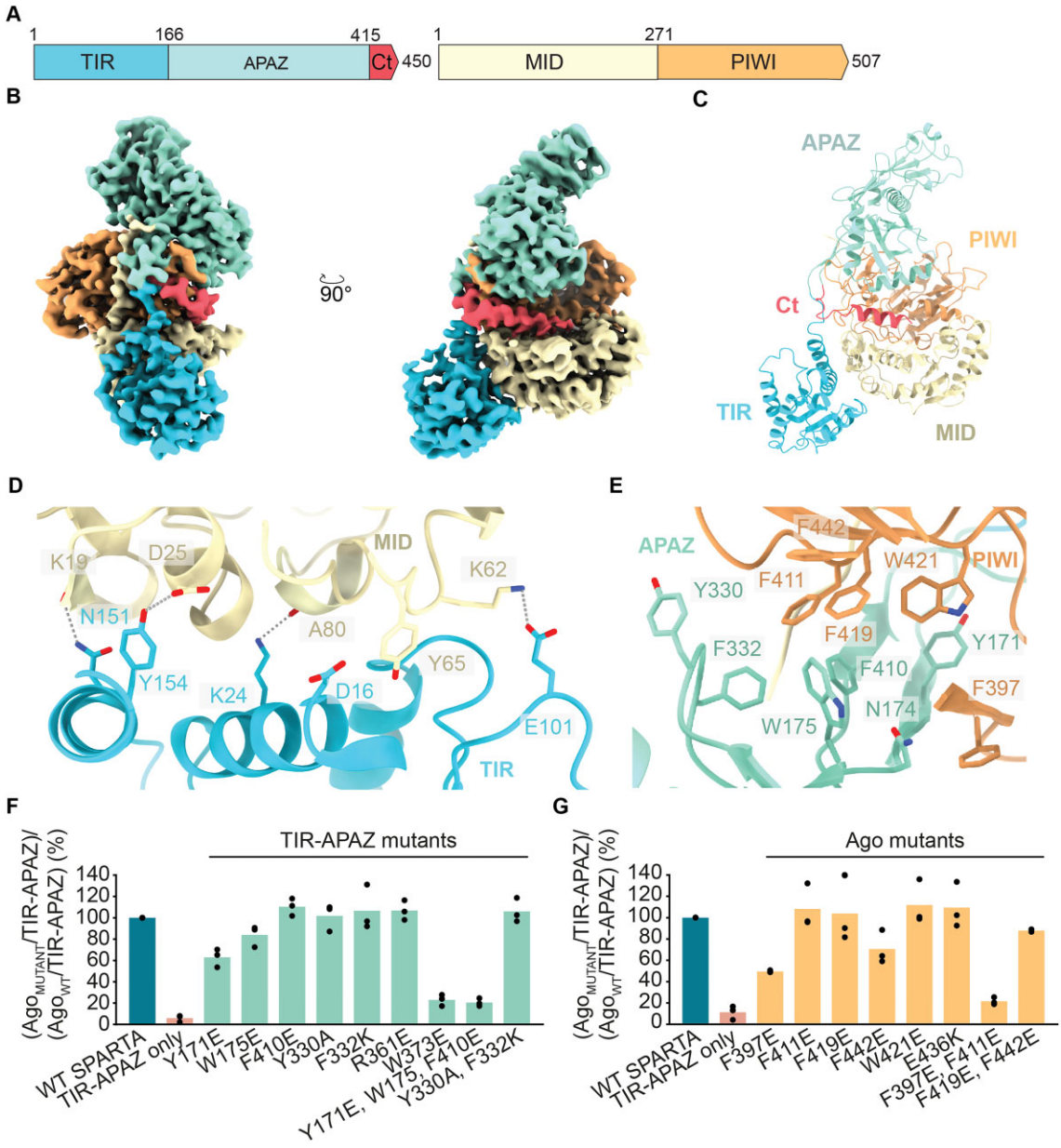


Figure 1. Molecular architecture of the monomeric apo-SPARTA complex. (**A**) Schematic diagram of the domain organization of the BabTIR-APAZ and BabAgo proteins. TIR, Toll-interleukin-1 receptor-like; APAZ, analog of PIWI-AGO-ZWILLE domain; Ct, C-terminal tail; MID, Middle; PIWI, P-element induced (PIWI, P-element induced wimpy testis domain.). (**B**) Cryo-electron microscopic (cryo-EM) densities of the BabSPARTA complex. Domains are colored according to the scheme in (A). (**C**) Cartoon representation of the overall structure of the Apo-BabSPARTA complex. Domains are colored according to the scheme in (A). (**D**) Close-up view of MID-TIR interactions. (**E**) Close-up view of PIWI-APAZ interactions. (F and G) Multiple residues at the PIWI-APAZ interface contribute to heterodimerization of TIR-APAZ and pAgo. 6xHis-MBP-BabTIR-APAZ or mutants thereof (**F**) were co-expressed in *E. coli* with BabAgo or mutants thereof (**G**) and proteins were purified using amylose affinity chromatography. Graphs show the percentage of the BabAgo/BabTIR-APAZ ratio normalized against the BabAgo/BabTIR-APAZ ratio of the WT proteins. Expression and consecutive pull-down experiments were performed in triplicates and data points reflect individual replicates. See also Supplementary Figure S3.

teraction sterically precludes the TIR domain from forming head-to-tail TIR-TIR interactions that are typically required for their catalytic activation (24,29,32,35), explaining why apo-SPARTA is catalytically inactive.

In turn, the APAZ-PIWI domain interface spans 1165 Å^2 and comprises extensive hydrophobic interactions involving numerous conserved aromatic residues (Figure 1E) as well as a single salt bridge involving Glu436^{pAgo} and Arg361$^{TIR\text{-}APAZ}$. Disruption of the salt bridge and various other surface mutants had limited to no effect on SPARTA heterodimer formation, as judged by pulldown assays, indicating that individual interface residues make limited contributions to the overall interaction (Figure 1F, G). However, BabSPARTA complex formation was at least partially disrupted in pAgo mutants F397E^{pAgo}, F442E^{pAgo}, and F397E/F411E^{pAgo}, as well as in TIR-APAZ mutants Y171E$^{TIR\text{-}APAZ}$, W175E$^{TIR\text{-}APAZ}$, W373E$^{TIR\text{-}APAZ}$ and Y171E/W175E/F410E$^{TIR\text{-}APAZ}$ (Figure 1F, G; Supplementary Figure S3). The corresponding residues in SPARTA, SPARSA, and SPARDA systems, all of which form heterodimeric complexes (12–14), are also strictly aromatic or hydrophobic (Supplementary Figures S4 and S5). This indicates that short

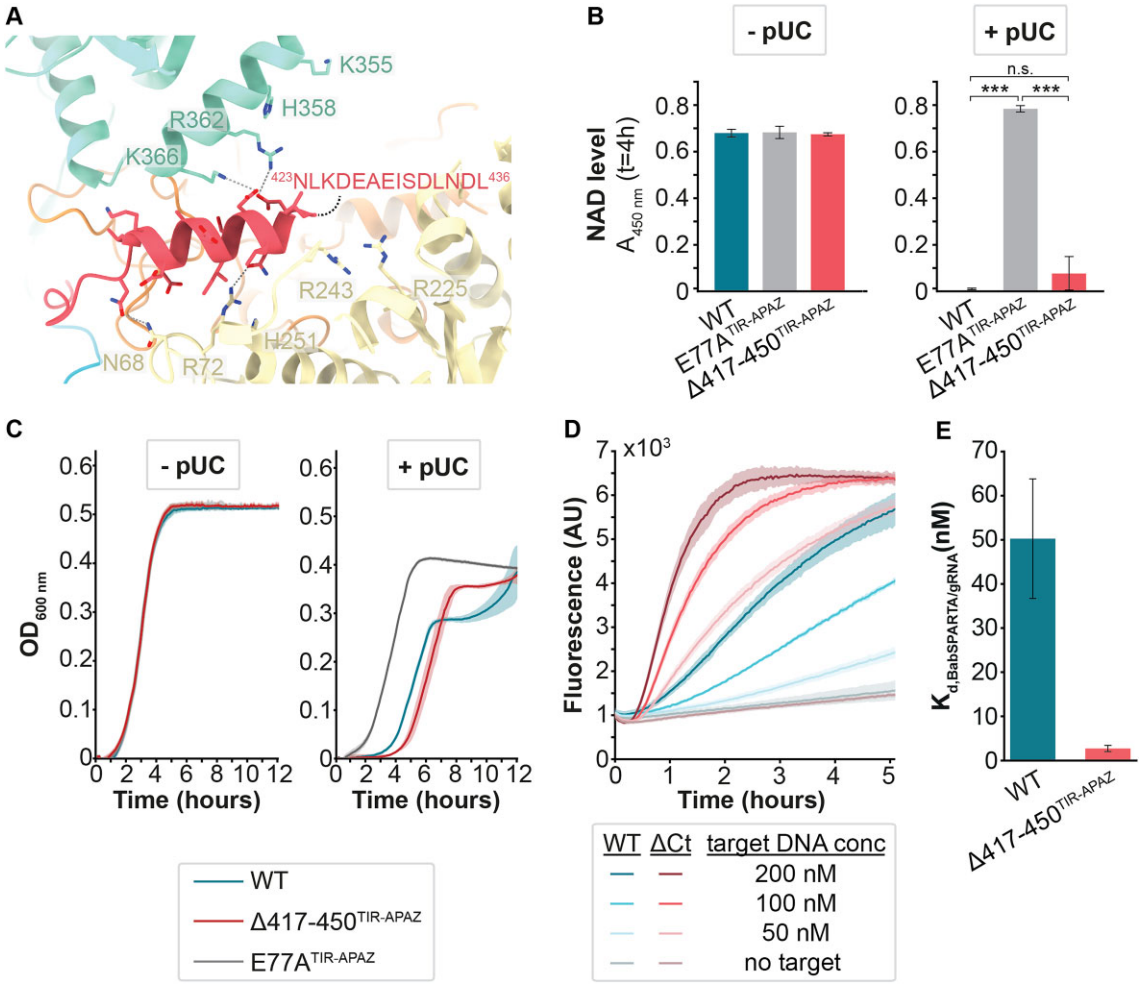


Figure 2. The C-terminal tail of TIR-APAZ blocks the nucleic acid binding cleft of the monomeric SPARTA complex. (**A**) Close-up view of the interactions of the C-terminal tail with the MID and APAZ domains. (**B**) Total NAD (NAD$^+$ + NADH) level in *E. coli* cultures expressing MapSPARTA, catalytic mutant E77A$^{TIR\text{-}APAZ}$, or C-terminal tail truncation mutant Δ417–450$^{TIR\text{-}APAZ}$ in the absence or presence of a highly transcribed high copy number plasmid (pUC-mRFP$^{\Delta RBS}$), determined 4 h after induction of MapSPARTA expression. The averages of three biological replicates are shown, error bars indicate standard deviations. ***$P < 0.001$. (**C**) Growth curves of *E. coli* cultures expressing MapSPARTA, catalytic mutant E77A$^{TIR\text{-}APAZ}$, or C-terminal tail truncation mutant Δ417–450$^{TIR\text{-}APAZ}$ in the absence or presence of pUC-mRFP$^{\Delta RBS}$. Cultures were started with pre-cultures in which MapSPARTA expression was induced 4 h earlier. The averages of three biological replicates are shown, shaded areas indicate standard deviations. (**D**) *In vitro* NADase activity of WT and C-terminally truncated BabSPARTA in the presence of varying concentrations of target ssDNA. Total fluorescence was measured over time after the addition of ε-NAD$^+$. The averages of three technical replicates are shown, shaded areas indicate standard deviations. (**E**) Affinity of WT and C-terminally truncated SPARTA for guide RNA. Equilibrium dissociation constants (K_d) were determined by measuring fluorescence polarization of a 3′-ATTO532-labelled guide RNA in the presence of varying concentrations of SPARTA or SPARTA mutant Δ417–450$^{TIR\text{-}APAZ}$. Relative Fluorescence Polarization (Supplementary Figure S7F) was used to calculate the K_d. The data points in the graph reflect averages of three technical replicates and error bars indicate standard deviations.

pAgo and TIR-APAZ proteins interact through cumulative hydrophobic/aromatic interactions and that such interactions are conserved in various short pAgos and their APAZ partners.

Monomeric SPARTA is autoinhibited by TIR-APAZ C-terminal tail

The relative spatial disposition of the BabTIR-APAZ and BabAgo MID and PIWI domains creates a deep, positively charged cleft that corresponds to binding site for the guide-target duplex in long pAgos (Figure 1B; Supplementary Figure S1C, D and Supplementary Figure S7A). Unlike long pAgos, in the monomeric BabSPARTA complex the cleft is occupied by the C-terminal tail (Ct) of TIR-APAZ (Figure 2A). Residues Asn423-Leu436$^{TIR\text{-}APAZ}$ form an alpha-helix, while the remainder of the Ct (residues 436–450$^{TIR\text{-}APAZ}$) is structurally disordered (Figure 2A). The Ct is highly acidic and makes numerous electrostatic contacts, via side chains of Asp426$^{TIR\text{-}APAZ}$, Glu427$^{TIR\text{-}APAZ}$, Glu429$^{TIR\text{-}APAZ}$, Asp432$^{TIR\text{-}APAZ}$ and Asp435$^{TIR\text{-}APAZ}$, and possibly additional acidic residues in the structurally disordered remainder of the Ct (residues Asn437-As451$^{TIR\text{-}APAZ}$), with positively charged residues in the APAZ (Lys366$^{TIR\text{-}APAZ}$, Arg362$^{TIR\text{-}APAZ}$) and MID (Arg72^{pAgo}, and Arg225^{pAgo}, Arg243^{pAgo} and His251^{pAgo}) domains (Figure 2A). The placement of the TIR-APAZ Ct in the inactive monomeric BabSPARTA complex suggests that it may sterically inhibit guide RNA and/or target DNA binding, thus modulating SPARTA activation in an autoinhibitory manner.

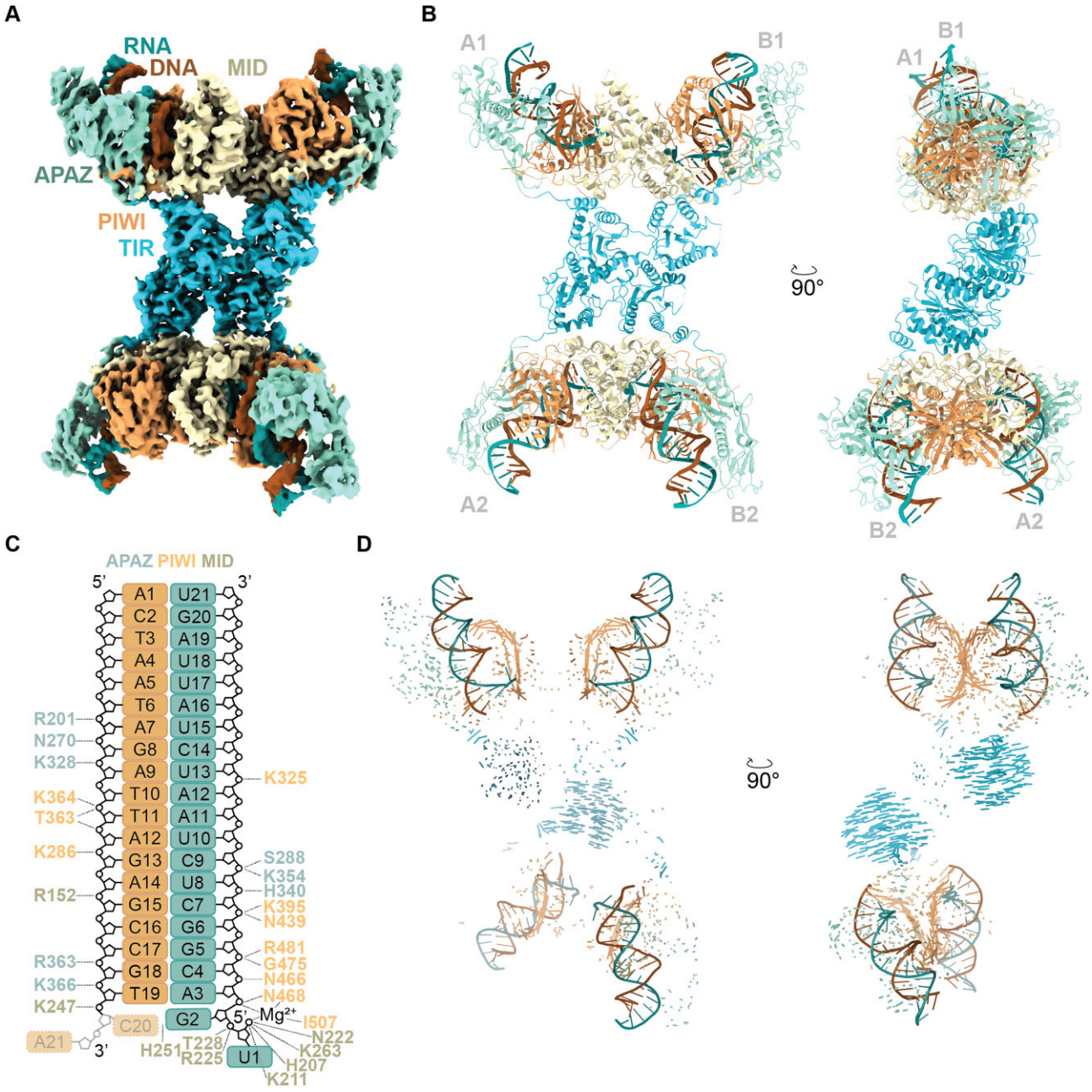


Figure 3. Molecular architecture of the activated tetrameric SPARTA complex. (**A**) Cryo-electron microscopy (cryo-EM) densities of the tetramer of guide RNA/target DNA duplex-bound BabSPARTA complexes. SPARTA protein domains are colored according to the scheme Figure 1A, the guide/target duplex is colored according to the scheme in panel (A). (**B**) Cartoon representation of the overall structure of the tetrameric BabSPARTA complex. (**C**) Schematic representation of the interactions of SPARTA with the guide RNA-target DNA duplex. Transparent and dotted rectangles indicate structurally disordered nucleotides. Hydrogen-bonding and electrostatic interactions are indicated with dotted lines. (**D**) Vector map displaying observed domain movement between the apo-BabSPARTA complex and each of the BabSPARTA protomers in the activated BabSPARTA complex. Vectors are colored corresponding to the domain coloring used in other panels.

To investigate the importance of the Ct in SPARTA systems, we tested the activation of SPARTA systems containing wild-type and C-terminally truncated (Ct-truncated) TIR-APAZ subunits *in vivo*. As BabSPARTA shows limited activity at 37°C (Supplementary Figure S6), we instead tested the activity of *Maribacter polysiphoniae* SPARTA (MapSPARTA), which displays 80% sequence identity with BabSPARTA, and Ct-truncated MapSPARTA mutants (Δ431–450$^{\text{TIR-APAZ}}$, Δ421–450$^{\text{TIR-APAZ}}$ and Δ417–450$^{\text{TIR-APAZ}}$). MapSPARTA was expressed from a bacterial artificial chromosome (BAC) in the absence or presence of a highly transcribed pUC-based plasmid (pUC-mRFP$^{\Delta\text{RBS}}$). SPARTA activity was assessed by monitoring cell growth and by measuring NAD^+ concentration in cell lysates. In the absence of invading DNA, MapSPARTA or Ct-truncated mutants thereof were not catalytically activated (Figure 2B and Supplementary Figure S7B–E). As observed previously (13), WT MapSPARTA depleted NAD^+ and substantially diminished the growth rate of *E. coli* in the presence of invading plasmid DNA, while MapSPARTA carrying a point mu-

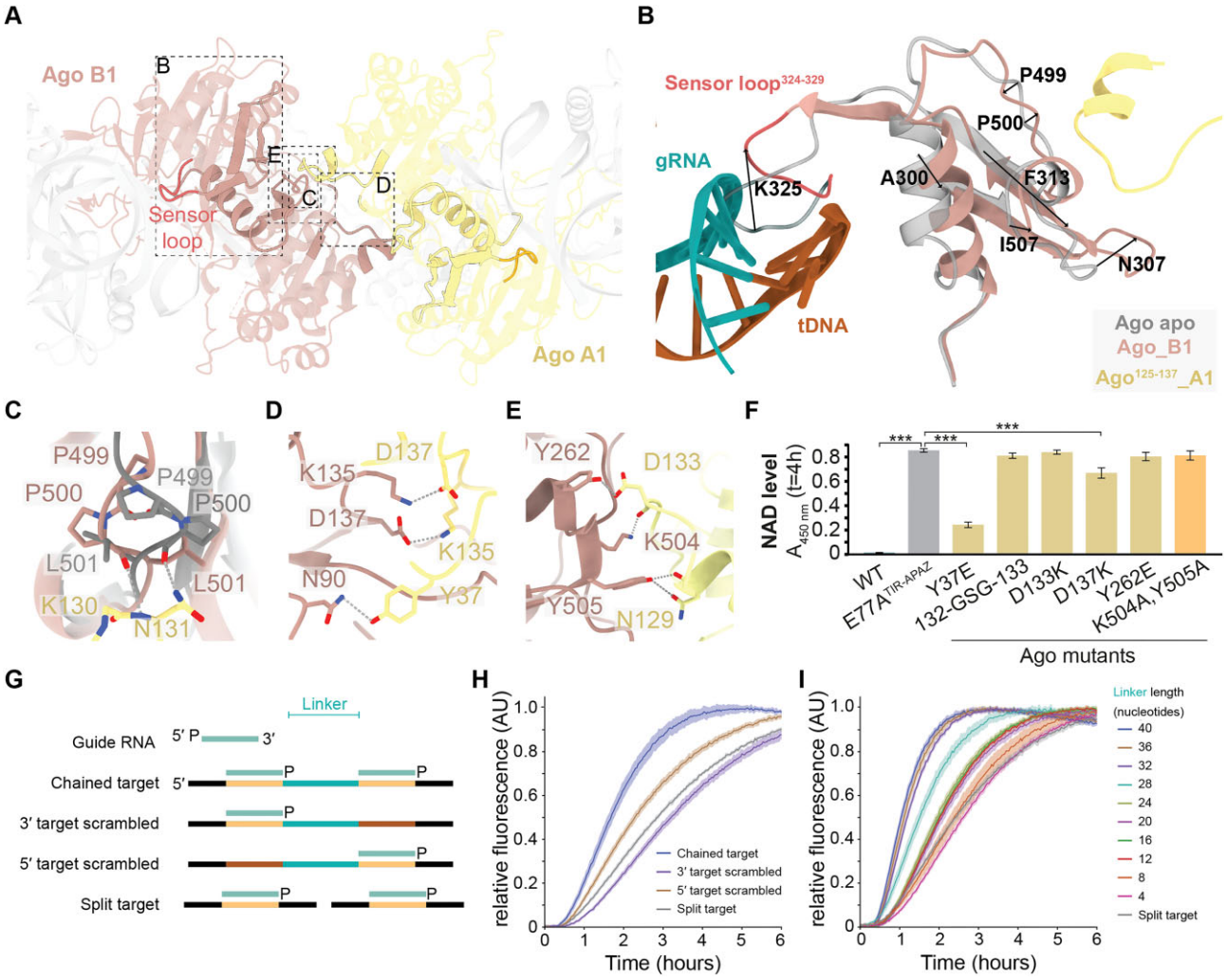


Figure 4. RNA-guided target DNA binding induces pAgo-pAgo interactions. (**A**) Guide RNA/target DNA duplex binding is associated with conformational changes of the sensor loop and other segments within the PIWI domain. (**B–E**) Structural rearrangements facilitate the formation of polar interactions and salt bridges between MID and PIWI domain residues. See also Supplementary Figure S10. (**F**) Mutational analyses reveal that pAgo-pAgo interaction residues are crucial for catalytic activation of SPARTA. The total NAD (NAD^+ + NADH) level was determined in *E. coli* cultures expressing MapSPARTA, catalytic mutant $E77A^{TIR\text{-}APAZ}$, or MapSPARTA with mutations at the pAgo-pAgo interface, in presence of a highly transcribed high copy number plasmid ($pUC\text{-}mRFP^{\Delta RBS}$). The averages of three biological replicates are shown, error bars indicate standard deviations. *$p < 0.05$, **$p < 0.01$; ***$p < 0.001$. See also Supplementary Figure S11. (G–I) SPARTA activation is increased by target sites co-localized in *cis*. BabSPARTA was mixed with guide RNA and ssDNA targets that contain a single target site (split target) or two target sites (chained target) complementary to the guide RNA, with equal abundance of binding sites in every condition (**G**). After addition of ε-NAD^+ fluorescence was measured over time. SPARTA shows increased NADase activity with chained targets compared to split targets and chained targets in which one of the target sites is scrambled (**H**). SPARTA activation by chained targets is influenced by the linker length that connect the two target sites (**I**). In panel H and I, measurements are corrected for a control without ssDNA target, and normalized to the minimum and maximum of each individual reaction. The averages of three technical replicates are shown, shadings indicate standard deviations.

tation in the TIR domain ($E77A^{TIR\text{-}APAZ}$) lacked NADase activity and did not affect cell growth (Figure 2B, C and Supplementary Figure S7B-E). Akin to the WT MapSPARTA, Ct-truncated mutants degraded NAD^+ in presence of the invading pUC plasmid (Figure 2B and Supplementary Figure S7C). However, expression of the Ct-truncated mutants in the presence of invading DNA delayed the onset of *E. coli* growth compared to expression of WT MapSPARTA, which suggests that the Ct-truncated mutant is hyperactive under these conditions (Figure 2C; Supplementary Figure S7B–E).

To further investigate whether the Ct controls SPARTA activity, we purified BabSPARTA in which the Ct was truncated ($\Delta 417\text{–}450^{TIR\text{-}APAZ}$) and compared its NADase activity to that of WT SPARTA at varying target DNA concentrations (Figure 2D). While the Ct-truncation mutant was inactive in absence of target DNA, it displayed faster activation and higher activity rates compared to WT SPARTA at the same target DNA concentrations. In agreement with the *in vivo* experiments, these results imply that the Ct is not required to keep SPARTA in an inhibited conformation in the absence of target DNA. Instead, the Ct might act as a competitive inhibitor of guide RNA and/or target DNA binding or as an allosteric inhibitor of target DNA-dependent conformational activation. To experimentally determine the affinity of WT SPARTA and the Ct-truncation mutant for guide RNA, we performed fluorescence polarization (FP) binding assays using a 16-nt long guide RNA with a 3′-ATTO532 label as ligand (Figure 2E, Supplementary Figure S7F). Quantitation

of the FP data yielded an equilibrium dissociation constant (K_d) of 50 (±13.5) nM for WT SPARTA-guide RNA complex, while truncation of the Ct reduces the K_d to 2.7 (±0.7) nM, implying that the Ct acts as a competitive inhibitor that reduces the affinity for guide RNA. Altogether, these results indicate that the Ct of TIR-APAZ is not essential to control the NADase activity of SPARTA, but that it plays an autoinhibitory role in modulating guide RNA binding, and possibly also in modulating target DNA binding.

Based on phylogeny, short pAgos can be subdivided in the S1A, S1B, S2A and S2B clades (16) (Supplementary Figure S7G). As clade S1A contains fused Sir2-APAZ-pAgo systems, we analyzed the conservation of the Ct in clades S1B (containing short pAgo/Sir2-APAZ (SPARSA) systems), S2A (containing SPARTA systems), and S2B (containing diverse short pAgo systems with various effector domains). This revealed that all APAZ domain-containing proteins in clades S2A and S2B have C-terminal extensions (on average 16–46 amino acids long) that contain a large proportion (21–45%) of negatively charged (Asp and Glu) residues, while a C-terminal extension is absent in clade S1 Sir2-APAZ proteins (Supplementary Figure S7H). This suggests that the activity of S2-clade pAgo systems, but not that of S1 clade pAgo systems, might be regulated by C-terminal tails in the APAZ domain-containing protein.

Molecular architecture of the activated tetrameric SPARTA complex

Guide RNA-mediated target DNA binding results in tetramerization of SPARTA complexes, which catalytically activates their TIR domains (13). However, the mechanisms underlying tetramerization and catalytic activation of SPARTA systems are unclear. To gain structural insights into these processes, we reconstituted the active tetrameric form of BabSPARTA in the presence of a 21-nucleotide guide RNA and a complementary 21-nucleotide ssDNA, and determined its cryo-EM structure at a resolution of 3.1 Å (Figure 3A, B, Supplementary Figure S8, and Table S1). The reconstruction reveals a butterfly-shaped architecture formed by four BabAgo:BabTIR-APAZ SPARTA complexes, each bound to a guide RNA-target DNA duplex (Figure 3A, B).

The guide RNA-target DNA duplex is bound in the positively charged groove formed between the APAZ domain, and the BabAgo MID and PIWI domains, in which the negatively charged backbone phosphate groups of both the guide RNA and the target DNA are coordinated by salt-bridge and hydrogen-bonding interactions with conserved positively charged and polar residues (Figure 3C). The Ct of the TIR-APAZ subunit is displaced from the binding groove and structurally disordered (Supplementary Figure S9A), indicating conformational flexibility, in agreement with its presumed modulatory role in SPARTA activation.

As observed for other pAgo-guide RNA complexes, the 5′-terminal phosphate group of the guide RNA is sequestered in a highly conserved MID domain pocket and coordinated by a Mg^{2+} ion, which precludes base pairing with a complementary nucleotide in the DNA target strand (Figure 3C and Supplementary Figure S9B). There are no base-specific interactions with the 5′-terminal nucleotide, in agreement with the observation that guide RNAs with any 5′-terminal nucleotide mediate SPARTA activation (Supplementary Figure S9C-E). Unlike in other pAgos (20,48–50), the guide RNA nucleotide in the second position also remains unpaired, and base pairing with target DNA is precluded by residue His251^{pAgo}, which is conserved in closely related SPARTA systems (Figure 3C and Supplementary Figure S4). In line with these observations, mismatches with the target at guide nucleotide positions 1 and 2 do not impair SPARTA activation (13).

The tetrameric assembly of the active BabSPARTA complex is a dimer of asymmetric dimers. Within each asymmetric dimer (protomers A1-B1 and A2-B2, Figure 3A, B), the MID, PIWI and APAZ domains are related by nearly perfect two-fold rotational symmetry, while the TIR domains adopt distinct orientations with respect to the other domains (Figure 3D). The orientation of one TIR domain resembles that observed in the inactive monomeric BabSPARTA complex, while the other TIR domain is repositioned by a ∼180° rotation about a hinge segment spanning residues 143–164$^{TIR-APAZ}$ (Figure 3D). The release and repositioning of the TIR domain promotes TIR-TIR interactions within each asymmetric dimer. In turn, the dimerized TIR domains of two asymmetric SPARTA-SPARTA dimers interact to form the tetrameric active SPARTA complex (Figure 3A, B), as detailed below.

RNA-guided target DNA binding induces pAgo-pAgo interactions

Comparisons of the inactive monomeric and active tetrameric BabSPARTA complexes reveal conformational changes in the BabAgo subunit brought about by RNA-guided target DNA binding. In the monomeric BabSPARTA complex, a loop spanning residues Glu324-His329^{pAgo} in the PIWI domain projects into the nucleic acid binding groove (Figure 2C and Figure 4A, B). This 'sensor loop' is displaced by the guide RNA–target DNA duplex, which is accompanied by allosteric restructuring of the upstream beta hairpin (residues Ala300-Gln323^{pAgo}) and an adjacent alpha helix with a downstream loop (residues Gly475-Leu501^{pAgo}) in the PIWI domain (Figure 4B and Supplementary Figure S10). These structural rearrangements reshape the surface of the pAgo subunit, which facilitates pAgo-pAgo interactions and thereby SPARTA–SPARTA dimer formation. Restructuring of the Pro499-Leu501^{pAgo} segment enables BabAgo-BabAgo dimerization via the formation of hydrogen bonding contacts with the backbone carbonyl and side chain amide groups of Asn131^{pAgo} in the MID domain of the other BabAgo protomer (Figure 4C and Supplementary Figure S10). This interaction is further reinforced by additional MID-MID contacts involving salt bridges between Lys135^{pAgo} and Asp137^{pAgo}, and MID-PIWI contacts involving the Asn129-Asp133^{pAgo} loop in one BabAgo protomer with residues Tyr262^{pAgo} and Lys504-Tyr505^{pAgo} in the other protomer (Figure 4D, E).

To verify the importance of the BabAgo-BabAgo dimerization interface residues, we investigated the effects of mutations in the corresponding residues in MapSPARTA on intracellular NAD levels and cell growth upon expression in *E. coli* in the presence of a target plasmid (pUC-mRFP$^{\Delta RBS}$) (Figure 4F and Supplementary Figure S11). MapSPARTA systems harboring mutations corresponding to Y37E^{pAgo} or D137K^{pAgo} exhibit lower NADase activities compared to WT MapSPARTA (72% and 22% reduction in NAD levels, compared to 99% reduction for WT). In turn, mutants corresponding to point mutations D133K^{pAgo} or Y262E^{pAgo}, the double mutation K504A/Y505A^{pAgo}, or the insertion mutation Glu132-Gly-Ser-Gly-Asp133^{pAgo} completely abolished NADase activity *in vivo* (Figure 4F and Supplementary Figure S11). For all mu-

tants, cell growth was comparable as for cells expressing the catalytically inactive E77A$^{TIR\text{-}APAZ}$ mutant (Supplementary Figure S11). These results corroborate the importance of key pAgo-pAgo interacting residues in the catalytic activation of SPARTA systems.

The BabAgo PIWI domain residue Lys395^{pAgo} interacts directly with both the backbone phosphate of guide RNA nucleotide 6 and BabTIR-APAZ residue Asn174$^{TIR\text{-}APAZ}$ (Supplementary Figure S12). SPARTA activity is highly sensitive to mismatches at this position (13). Despite the observation that these residues assume the same conformations in the apo and activated SPARTA complexes, we hypothesized that this interaction might be important for facilitating guide-target duplex interactions and/or contribute to SPARTA activation. Mutational analyses revealed that MapSPARTA systems harboring a mutation corresponding to K395A^{pAgo} exhibited lower NADase activities as compared to WT MapSPARTA (45% reduction in NAD levels, compared to 99% reduction for WT), while a mutation corresponding to N174A$^{TIR\text{-}APAZ}$ completely abolished NADase activity (Supplementary Figure S12). For both mutants, cell growth was similar to cells expressing the catalytically inactive E77A$^{TIR\text{-}APAZ}$ mutant. Both Lys395^{pAgo} and Asn174$^{TIR\text{-}APAZ}$, as well as the majority of other pAgo-pAgo interface residues, are conserved in SPARTA orthologs and to some extent also in other short pAgo systems (Supplementary Figures S4 and S5). This suggests that target DNA-dependent pAgo-pAgo dimerization is not only important for the activation mechanism of SPARTA, but is likely conserved in other short pAgo systems.

While the molecular architecture of the target-activated SPARTA tetramer suggests that the monomeric SPARTA complex initially dimerizes via pAgo-pAgo interactions upon binding of the target DNA, dimeric SPARTA assemblies are not readily detectable during size-exclusion chromatography (13). This suggests that pAgo-pAgo dimers are unstable. SPARTA tetramers are observed during size-exclusion chromatography (13), which implies that subsequent cooperative SPARTA tetramerization stabilizes the active conformation. The architecture of the SPARTA tetramer suggests that the target sites need to be located ∼90 Å apart so as to be simultaneously engaged by two SPARTA protomers within a SPARTA dimer, which would correspond to approximately 15 nucleotides. The target sites need to be located ∼150 Å apart to be simultaneously engaged by two SPARTA protomers from different SPARTA dimers, which would correspond to approximately 25 nucleotides. We hypothesized that SPARTA dimerization is enhanced upon recognition of two independent target sites located in the same DNA molecule. To test this, we designed 'chained target' DNAs in which two identical target sites are linked (Figure 4G).

We compared the NADase activity of BabSPARTA upon activation by a chained target DNA containing a 28-nucleotide linker, chained targets in which either the 5′ or the 3′ target site was scrambled, or an unchained target containing a single target site, keeping the effective concentration of target sites the same for all targets (Figure 4G). The chained target DNA elicited higher NADase activity as compared to the unchained target (Figure 4H). When either of the target sites in the chained targets was scrambled, SPARTA activity remained at a level similar to that observed when it was incubated with the unchained target (Figure 4H). Next, we analyzed SPARTA activation using chained targets with variable linker lengths. Targets with linker lengths ≤ 8 nucleotides activate SPARTA at similar levels as unchained targets, suggesting that the linker is too short to stabilize SPARTA dimer formation (Figure 4I). In contrast, chained targets with a linker length between 12–24 nucleotides elicited higher levels of SPARTA NADase activity, and chained targets with a linker length between 32–40 nucleotides elicited even higher SPARTA NADase activity (Figure 4I). Corroborating our structure-based predictions, these findings suggest that targets with 12–24 nucleotide-long linkers stabilize the interactions between two SPARTA protomers within a SPARTA dimer, while targets with 32–40 nucleotide-long linkers could stabilize the interactions of SPARTA protomers within and across SPARTA dimers. Together, these results confirm that catalytic activation of SPARTA is enhanced by target sites localized in *cis*, suggesting that stabilization of SPARTA dimers and tetramers by co-localized targets is an important aspect of the activation mechanism of SPARTA. While in our experiments these target sites have an identical sequence, we speculate that, *in vivo*, SPARTA complexes programmed with distinct guide RNAs could target DNA at adjacent target sites to promote SPARTA tetramerization.

TIR domain repositioning facilitates tetramerization and catalytic activation

The TIR domains in the active BabSPARTA tetramer adopt two distinct orientations with respect to the corresponding BabAgo subunits (Figure 5A, B). In protomers B1 and B2, the orientation is nearly identical to that observed in the inactive monomeric BabSPARTA complex. In protomers A1 and A2, however, the TIR domain and an alpha-helical linker segment connecting the TIR and APAZ domains (residues 148–168$^{TIR\text{-}APAZ}$) are restructured such that the TIR domain undergoes a 177° rotation with respect to the linker, centered on the hinge residue Pro147$^{TIR\text{-}APAZ}$ (Figure 5B). The TIR-MID domain interactions, as formed in the monomeric SPARTA complex (Figure 1D) and in the tetramer protomers B1/B2, are disrupted and the TIR domain is released from the pAgo subunit. The TIR-APAZ linker makes contacts with a loop (residues 413–416$^{TIR\text{-}APAZ}$) connecting the APAZ domain to the Ct, suggesting a possible mechanism for allosterically coupling TIR domain restructuring to the displacement of the Ct upon RNA-guided target DNA binding.

The reorientation of the TIR domain in protomer A1 consequently enables the formation of interface interactions with the TIR domain in protomer B1 within the asymmetric dimer. This results in the insertion of the BB-loop (51) (residues Asp40-Trp46$^{TIR\text{-}APAZ}$) of the rotated A1 TIR domain into a pocket in the B1 TIR domain, forming a composite NADase catalytic site capable of NAD binding (Figure 5C, D). The composite active site is homologous to those of other TIR domain NADases such as eukaryotic SARM1 (52) (Supplementary Figure S13A). Furthermore, the dimerization of the TIR domains generates an interface comprising TIR domain helices αB, αC (from the TIR domain of protomer A1) and αD (from the TIR domain of protomer B1). This interface mediates interactions with the TIR domain of protomer B2 (Figure 5C, D). Overall, this results in cooperative tetramerization of the TIR domains by the formation of a parallel dimer of dimers (Figure 5C, D).

To verify the functional significance of TIR-TIR interfaces for the activity of SPARTA systems, we tested the effects of corresponding mutations in MapSPARTA upon expression in *E. coli* in the presence of the pUC-mRFP$^{\Delta RBS}$ tar-

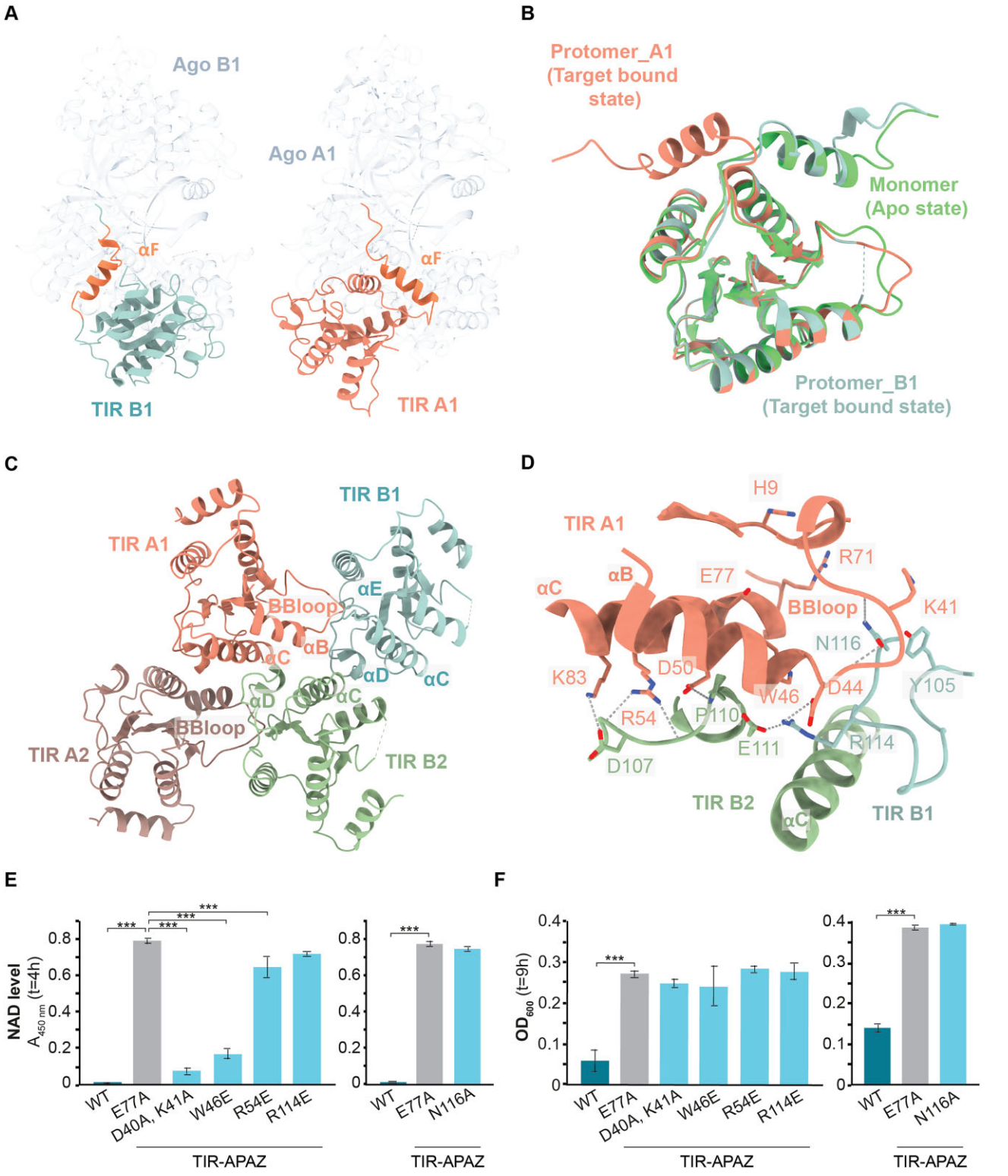


Figure 5. TIR domain repositioning facilitates tetramerization and catalytic activation**. (A)** Comparison of the TIR domain orientation between two SPARTA protomers interacting after tetramerization. The TIR domain in protomer A1 (in pink) is rotated 177° with respect to the TIR domain in protomer B1 (teal), with helix αF (bright orange) functioning as a hinge. Alignment performed with respect to the pAgo-APAZ moieties (grey). (**B**) Alignment of the TIR domains of the inactive apo monomeric SPARTA complex structure (green) with the TIR domains of two asymmetric SPARTA protomers from the activated tetrameric SPARTA structure (orange and teal). The overall TIR domain structure is maintained, except for the rotation of the αF helix (bright orange) in protomer A1 (orange) of the activated tetrameric SPARTA structure. (**C**) Overview of the assembly of the four TIR domains (TIR_A1/B1/A2/B2) in the activated tetrameric SPARTA structure. (**D**) TIR-TIR interfaces and key residues involved in TIR-TIR dimerization, TIR tetramerization, and TIR catalytic activity. (E, F) TIR-TIR interface residues are crucial for catalytic activation of SPARTA. The total NAD (NAD^+ + NADH) level (**E**) or $OD_{600\ nm}$ (**F**) was determined in *E. coli* cultures expressing MapSPARTA, catalytic mutant $E77A^{TIR\text{-}APAZ}$, or MapSPARTA with mutations at the TIR-TIR interfaces, in presence of a highly transcribed high copy number plasmid ($pUC\text{-}mRFP^{\Delta RBS}$). The MapSPARTA and the $E77A^{TIR\text{-}APAZ}$ mutant controls in the right panel are identical to those in Figure 1 (same experiment). The averages of three biological replicates are shown, error bars indicate standard deviations. $^*P < 0.05$, $^{**}P < 0.01$; $^{***}P < 0.001$. See also Supplementary Figure S13.

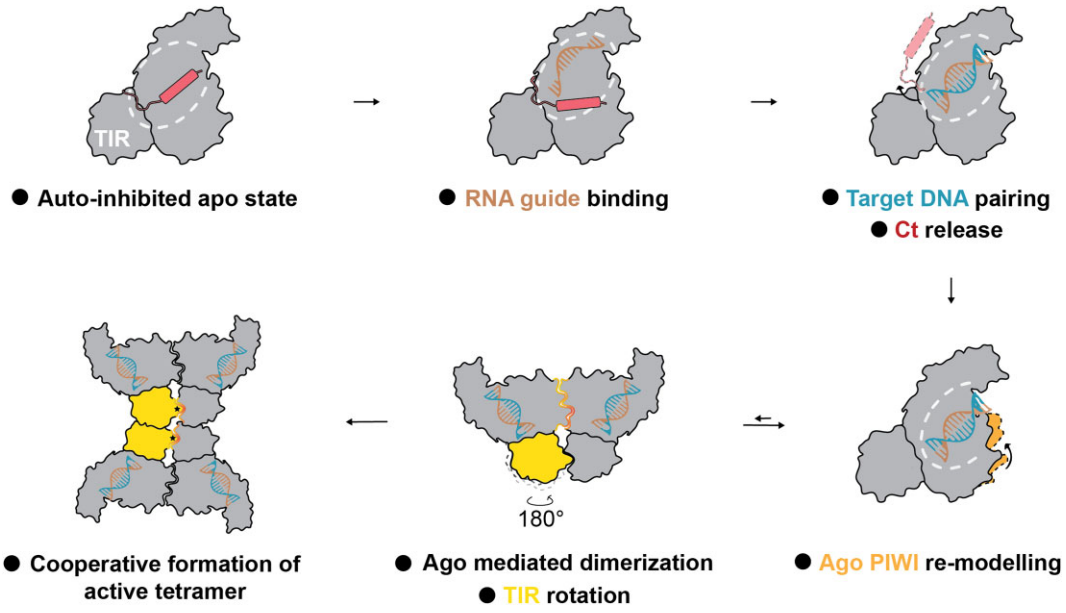


Figure 6. Schematic model of guide RNA-mediated target binding and catalytic activation of SPARTA.

get plasmid (Figure 5E, F, and Supplementary Figure S13B, C). BB loop mutations D40A/K41A$^{TIR\text{-}APAZ}$ or W46E$^{TIR\text{-}APAZ}$ showed substantially reduced NADase activity *in vivo* (91% and 78% reductions in NAD^+ levels, respectively, as compared to 99% reduction for WT MapSPARTA). Mutations of TIR tetramerization interface residues TIR-APAZ^{R54E} and TIR-APAZ^{R114E} abolished NADase activity completely (Figure 5E, F, and Supplementary Figure S13B, C). For all mutants, cell growth was comparable to that of cells expressing the catalytically inactive E77A$^{TIR\text{-}APAZ}$ mutant, indicating loss of interference activity. Collectively, these results confirm the importance of both the TIR-TIR dimerization and the tetramerization interfaces for catalytic activation of SPARTA, and indicate that cooperative tetramerization of target-bound SPARTA complexes, mediated by TIR domain interactions, is essential for the activity of SPARTA systems.

Discussion

SPARTA systems couple short pAgo-mediated, RNA-guided target DNA recognition with TIR domain tetramerization to elicit catalytic activation that triggers NADase activity. This in turn depletes the cellular pool of NAD, leading to growth arrest and/or cell death, thereby providing population-level immunity against genetic parasites such as plasmids (13). As such, the target sensing and activation of SPARTA systems needs to be tightly regulated to avoid spurious activation. To shed light on the target-dependent activation mechanism of SPARTA, we have performed structural analysis of Bab-SPARTA, determining atomic structure of the monomeric inactive and tetrameric nucleic acid-bound SPARTA complexes.

The monomeric SPARTA structure reveals that the TIR-APAZ and its associated short pAgo subunit form a heterodimeric complex through the mostly hydrophobic surfaces of their APAZ and MID-PIWI domains. As these hydrophobic interaction surfaces are generally conserved in short pAgo systems, we hypothesize that short pAgo systems from all clades adopt similar structural architectures with their respective interaction proteins. Homologous to the bilobed structure of long pAgos and eAgos (48,49), pAgo and TIR-APAZ together form a positively charged binding channel that accommodates a guide-target duplex. Interactions between the TIR and MID domains in the apo-SPARTA complex prevent productive dimerization of the TIR domains requiring the insertion of the BB-loop from one TIR domain into the complementary pocket in the other TIR domain, thus keeping apo-SPARTA in a catalytically inactive state.

Furthermore, in the absence of a guide RNA-target DNA duplex, the negatively charged Ct of TIR-APAZ resides in the guide RNA-target DNA binding channel. As Ct and guide-target duplex binding are mutually exclusive for steric reasons, we infer that the Ct acts as a negative regulator of target binding, thereby providing an autoinhibitory mechanism within SPARTA systems to avoid non-specific activation. This is in agreement with our observation that deletion of the Ct results in hyperactivation of SPARTA *in vivo*. A negatively charged C-terminal extension is conserved in the APAZ-containing proteins that belong to S2A and S2B clade short pAgo systems, but not in S1B clade short pAgo systems. Although the effector domains of the S2A (TIR) and S2B (DUF4365, Mrr-like, RecG/DHS-like) clade pAgo systems vary, this suggests that their activity is modulated through similar mechanisms. Possibly, the Ct contributes to short pAgo system fidelity and thereby prevents uncontrolled activation of these abortive infection systems.

In contrast to previously reported long pAgo and eAgo structures (53–56), we do not observe pre-ordering of the guide RNA by SPARTA in the absence of target DNA. This suggests that guide RNAs are only transiently or flexibly bound, or that SPARTA systems recognize preformed guide-target duplexes. Alternatively, guide loading might require additional factors. For example, certain long-A pAgos load ds-DNA substrates more efficiently than single stranded guide DNAs (49), and eAGOs are typically loaded with dsRNA guide-passenger duplexes by Dicer (4). Guide generation and loading mechanisms are currently poorly understood not only for SPARTA systems, but for pAgos in general. As loading of invader-specific guides is crucial for determining self (genome)

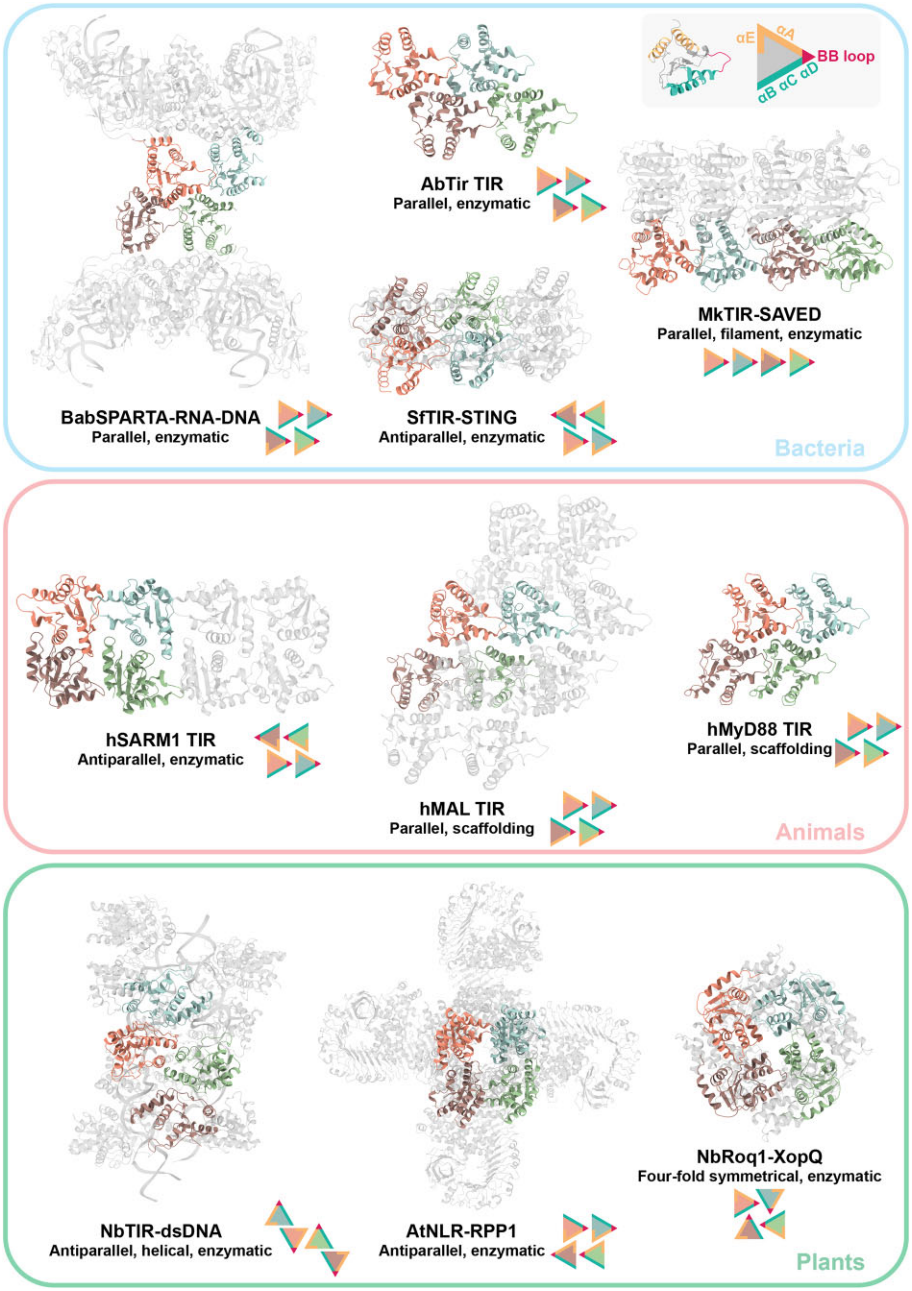


Figure 7. Comparison of oligomerization modes of enzymatic and scaffolding TIR domains found in bacteria, animals, and plants. Representative structures of TIR containing proteins with different origin, assembly geometry and function. Schematics of the main interaction interfaces are represented using triangles of different colors for different TIR-helices positions. (AbTir TIR, PDB: 7UXU; SfTIR-STING, PDB: 7UN8; MkTIR-SAVED, PDB: 7QQK; hSARM1 TIR, PDB: 7NAK; hMAL TIR, PDB: 5UZB; hMyD88 TIR, PDB: 6I3N; NbTIR-dsDNA, PDB: 7 × 5K; AtNLR-RPP1, PDB: 7CRC; NbRoq1-XopQ, PDB: 7JLV).

vs. non-self (invading DNA) targeting, future research should focus on this aspect of pAgo mechanism.

The structure of the target-activated tetrameric SPARTA assembly reveals that the guide RNA becomes ordered upon binding of the target DNA strand. Akin to other pAgos, strictly conserved MID domain residues lock the phosphorylated 5′ end of the guide RNA in a pocket that precludes base pairing between guide nucleotide 1 and the target. In contrast to other Agos, also base pairing between guide nucleotide 2 and the target is precluded in SPARTA. Corroborating this observation, target DNA mismatches to guide RNA nucleotides 1 and 2 do not reduce but instead appear to increase SPARTA activity (13). To our knowledge, precluding base pairing of the second guide nucleotide is unique for SPARTA systems and any potential functional relevance remains to be determined.

Based on our structural, biochemical and functional data, we propose a model for the activation mechanism of SPARTA (Figure 6). Guide RNA-dependent target DNA binding results in the displacement of the TIR-APAZ C-terminal tail. Given direct contacts between the linker segment connecting the TIR and APAZ domains and the loop connecting the APAZ domain and the Ct, we speculate that displacement of the Ct might result in allosteric perturbation of interactions that keep the TIR domain in an inhibited conforma-

tion, thereby 'unlocking' the TIR domain to enable TIR-TIR dimerization. Concurrent with displacement of the Ct, binding of the guide RNA-target DNA duplex additionally triggers the displacement of a 'sensor loop' located in the PIWI domain. Restructuring of the sensor loop in turn allosterically induces further changes in the PIWI domain that result in the formation of an interface that facilitates pAgo-pAgo dimerization. In the SPARTA-SPARTA dimer, the TIR domain of one SPARTA protomer is rotated and forms a BE helix interface with the TIR domain of the other SPARTA protomer which remains in the 'inactive' conformation. This TIR-TIR dimerization results in the formation of a composite catalytic site by insertion of one BB-loop into the other TIR domain. However, through formation of a BCD interface, the TIR dimer also interacts with the TIR-TIR dimer of another SPARTA dimer, forming a tetrameric assembly in a parallel configuration (Figure 6). Consequently, SPARTA activation requires a total of four activated SPARTA complexes forming a complex with a 4:4:4:4 stoichiometry (TIR-APAZ:pAgo:RNA:DNA). As formation of stable dimeric SPARTA complexes is not observed upon target DNA binding (13), we conclude that TIR tetramerization is a fast, cooperative process that is essential to stabilize the initially formed SPARTA-SPARTA dimers, while the initial dimerization may be the rate limiting step. In this way, tetramerization contributes to the sensitivity, fidelity, and spatiotemporal control of SPARTA activity.

TIR domain oligomerization is a general activation mechanism concept in immune systems that use TIR as an effector domain for NADase activity or signaling (24,26,29,32,34,35). However, TIR domains have distinct oligomerization configurations in diverse immune systems (Figure 7). For example, enzymatic TIR domains from animal SARM1, plant NLR-RPP1, and bacterial TIR-STING immune proteins arrange a double antiparallel TIR filament configuration (34,35,52) in which a catalytically active filament interacts though the BE-interfaces, while the other filament forms distinct interfaces and mostly acts to stabilize the catalytic filament through AE-interfaces. On the other hand, bacterial TIR-SAVED proteins form a single enzymatic TIR filament stabilized by cyclic oligoadenylate-dependent polymerization of the SAVED domains (29). In contrast to these TIR domains, SPARTA TIR dimers assemble in a parallel, head-to-tail orientation through a BCD interface analogous to the enzymatic TIR domain from the bacterial virulence factor Tir and scaffolding TIR oligomers formed by animal MAL and MyD88 proteins (57–59). In SPARTA, further scaffolding-like polymerization of TIR is not possible due to steric hindrance by the pAgo subunit, which keeps the second TIR in an inhibited conformation. Overall, these structural comparisons highlight the versatility of TIR oligomerization which, despite highly variable underlying interactions, is essential for TIR functions in innate immunity in all domains of life.

While this manuscript was in preparation, additional studies reporting structures of inactive and target-activated SPARTA complexes from *Crenotalea thermophila* (82% sequence ID with BabSPARTA) and MapSPARTA (80% sequence ID with BabSPARTA) were published (60–67). Beyond insights derived in our study, these studies reveal structures of SPARTA bound to the guide RNA-target DNA duplex prior to oligomerization (66) and in the dimeric form prior to tetramerization (60,64,67). Furthermore, mutational analyses in these studies extend insights into the structural determinants of guide RNA and target DNA binding (67), TIR-TIR interactions (60,61,66,67), and interactions of the TIR domains with the NAD substrate (61,62). Together with our structural analysis, these studies provide a comprehensive understanding of SPARTA mechanisms.

While the findings described in the structural studies of SPARTA complexes generally support each other, there are notable discrepancies in the effect of Ct truncation. Our results indicate that apo-SPARTA remains inactive when the Ct is truncated, but that the SPARTA Ct-truncation mutant has a higher affinity for guide RNA, and a higher NADase activity, as compared to WT SPARTA (Figure 2). It is possible that the Ct also influences the affinity for the target DNA, or that the Ct additionally plays a role in conformational activation of SPARTA. In other studies, Ct truncation has been reported to increase (60) or decrease (65) affinity for target DNA, and it either increases (67) or has no effect on NADase activity (60). Furthermore, one study shows that Ct truncation results in activation of SPARTA in the absence of guide and target DNA (67). Further studies are thus required to expose the mechanism by which the Ct modulates SPARTA activity.

Finally, as structure-guided mutants have improved the activity and fidelity of CRISPR-Cas-based nucleic acid detection and genome editing tools (68–71), we expect that this work provides a framework for further development of SPARTA-based nucleic acid detection tools (13). In conclusion, this study provides critical insights into the structural architecture of SPARTA systems and the molecular mechanisms that control the catalytic activation of their TIR domains. It contributes to our understanding of the general mechanisms underlying the biological function of short pAgo systems, and enables their further exploitation as programmable molecular tools.

Data availability

Plasmid and oligonucleotide sequences, multiple sequence alignments, C-terminal tail sequence analysis, and raw data underlying bar graphs are available via Mendeley Data (https://doi.org/10.17632/w94r9mgx2d.1).

The protein expression plasmid encoding MapSPARTA (pBK086) and BabSPARTA (pAP007) have been deposited on Addgene (Plasmid #183145 and #206809). Other plasmids are available upon request.

Atomic coordinates and cryo-EM maps have been deposited in the protein data bank (PDB entry ID 8QLO and 8QLP) and Electron Microscopy Data Bank (EMDB, entry ID EMD-18486 and EMD-18487).

Supplementary data

Supplementary Data are available at NAR Online.

Acknowledgements

We thank members of the Swarts and Jinek labs for discussion.

Author contributions: G.F., B.K., A.H.W., M.J. and D.C.S. designed experiments. G.F. and M.J. determined and analyzed SPARTA structures. B.K., C.H. and A.P., purified proteins and B.K., A.P., C.H., A.H.W. and D.C.S. performed and/or analyzed biochemical and *in vivo* experiments. G.F., B.K., M.J. and D.C.S. wrote the manuscript and

prepared figures. All authors have read and approved the manuscript.

Funding

European Research Council (ERC) [Consolidator Grant CRISPR2.0, ERC-CoG-820152 to M.J.]; Starter Grant COMPASS [ERC-2020-STG 948783 to D.C.S.]; Dutch Research Council [NWO; Veni Grant (Grant no. 016.Veni.192.072) to D.C.S.]. Funding for open access charge: European Research Council (ERC) starting grant [ERC-2020-STG 948783 to D.C.S.].

Conflict of interest statement

D.C.S., B.K. and A.P. are named inventors on a patent regarding the utilization of short pAgo systems for nucleic acid detection.

References

1. Bobadilla Ugarte,P., Barendse,P. and Swarts,D.C. (2023) Argonaute proteins confer immunity in all domains of life. *Curr. Opin. Microbiol.*, **74**, 102313.
2. Swarts,D.C., Makarova,K., Wang,Y., Nakanishi,K., Ketting,R.F., Koonin,E.V., Patel,D.J. and Van Der Oost,J. (2014) The evolutionary journey of argonaute proteins. *Nat. Struct. Mol. Biol.*, **21**, 743–753.
3. Burroughs,A.M., Ando,Y. and Aravind,L. (2014) New perspectives on the diversification of the RNA interference system: insights from comparative genomics and small RNA sequencing. *Wiley Interdiscip. Rev. RNA*, **5**, 141–181.
4. Meister,G. (2013) Argonaute proteins: functional insights and emerging roles. *Nature Publishing Group*, **14**, 447.
5. Jolly,S.M., Gainetdinov,I., Jouravleva,K., Zhang,H., Strittmatter,L., Bailey,S.M., Hendricks,G.M., Dhabaria,A., Ueberheide,B. and Zamore,P.D. (2020) Thermus thermophilus argonaute functions in the completion of DNA replication. *Cell*, **182**, 1545–1559.
6. Olina,A., Agapov,A., Yudin,D., Sutormin,D., Galivondzhyan,A., Kuzmenko,A., Severinov,K., Aravin,A.A. and Kulbachinskiy,A. (2023) Bacterial argonaute proteins aid cell division in the presence of topoisomerase inhibitors in Escherichia coli. *Microbiol. Spectr.*, **11**, e0414622.
7. Fu,L., Han,L., Xiang,Y., Xie,C. and Jin,M. (2019) The Prokaryotic argonaute proteins enhance homology sequence-directed recombination in bacteria. *Nucleic Acids Res.*, **47**, 3568–3579.
8. Swarts,D.C., Jore,M.M., Westra,E.R., Zhu,Y., Janssen,J.H., Snijders,A.P., Wang,Y., Patel,D.J., Berenguer,J., Brouns,S.J.J., *et al.* (2014) DNA-Guided DNA interference by a prokaryotic argonaute. *Nature*, **507**, 258–261.
9. Swarts,D.C., Hegge,J.W., Hinojo,I., Shiimori,M., Ellis,M.A., Dumrongkulraksa,J., Terns,R.M., Terns,M.P. and Van Der Oost,J. (2015) Argonaute of the archaeon pyrococcus furiosus is a DNA-guided nuclease that targets cognate DNA. *Nucleic Acids Res.*, **43**, 5120–5129.
10. Kuzmenko,A., Oguienko,A., Esyunina,D., Yudin,D., Petrova,M., Kudinova,A., Maslova,O., Ninova,M., Ryazansky,S., Leach,D., *et al.* (2020) DNA targeting and interference by a bacterial argonaute nuclease. *Nature*, **587**, 632–637.
11. Lisitskaya,L., Kropocheva,E., Agapov,A., Prostova,M., Panteleev,V., Yudin,D., Ryazansky,S., Kuzmenko,A., Aravin,A.A., Esyunina,D., *et al.* (2023) Bacterial argonaute nucleases reveal different modes of DNA targeting in vitro and in vivo. *Nucleic Acids Res.*, **1**, 13–14.
12. Prostova,M., Kanevskaya,A., Panteleev,V., Lisitskaya,L., Tugaeva,K.V., Sluchanko,N.N., Esyunina,D. and Kulbachinskiy,A. (2023) DNA-targeting short argonaute triggers effector nuclease to protect bacteria from invaders. bioRxiv doi: https://doi.org/10.1101/2023.06.08.544223, 08 June 2023, preprint: not peer reviewed.
13. Koopal,B., Potocnik,A., Mutte,S.K., Aparicio-Maldonado,C., Lindhoud,S., Vervoort,J.J.M., Brouns,S.J.J. and Swarts,D.C. (2022) Short prokaryotic argonaute systems trigger cell death upon detection of invading DNA. *Cell*, **185**, 1471–1486.
14. Zaremba,M., Dakineviciene,D., Golovinas,E., Zagorskaitė,E., Stankunas,E., Lopatina,A., Sorek,R., Manakova,E., Ruksenaite,A., Silanskas,A., *et al.* (2022) Short prokaryotic argonautes provide defence against incoming mobile genetic elements through NAD+ depletion. *Nat. Microbiol.*, **7**, 1857–1869.
15. Ryazansky,S., Kulbachinskiy,A. and Aravin,A.A. (2018) The expanded universe of prokaryotic argonaute proteins. *mBio*, **9**, e01935-18.
16. Koopal,B., Mutte,S.K. and Swarts,D.C. (2022) A long look at short prokaryotic argonautes. *Trends Cell Biol.*, **33**, 605–618.
17. Song,J.J., Smith,S.K., Hannon,G.J. and Joshua-Tor,L. (2004) Crystal structure of argonaute and its implications for RISC slicer activity. *Science*, **305**, 1434–1437.
18. Jin,S., Zhan,J. and Zhou,Y. (2021) Argonaute proteins: structures and their endonuclease activity. *Mol. Biol. Rep.*, **48**, 4837–4849.
19. Dayeh,D.M., Kruithoff,B.C. and Nakanishi,K. (2018) Structural and functional analyses reveal the contributions of the C- and N-lobes of argonaute protein to selectivity of RNA target cleavage. *J. Biol. Chem.*, **293**, 6308–6325.
20. Ma,J.-B., Yuan,Y.-R., Meister,G., Pei,Y., Tuschl,T. and Patel,D.J. (2005) Structural basis for 5′-end-specific recognition of guide RNA by the A. fulgidus piwi protein. *Nature*, **434**, 666–670.
21. Hur,J.K., Zinchenko,M.K., Djuranovic,S. and Green,R. (2013) Regulation of argonaute slicer activity by guide RNA 3′ end interactions with the N-terminal lobe. *J. Biol. Chem.*, **288**, 7829–7840.
22. Kwak,P.B. and Tomari,Y. (2012) The N domain of argonaute drives duplex unwinding during RISC assembly. *Nat. Struct. Mol. Biol.*, **19**, 145–151.
23. Kawamata,T., Seitz,H. and Tomari,Y. (2009) Structural determinants of MiRNAs for RISC loading and slicer-independent unwinding. *Nat. Struct. Mol. Biol.*, **16**, 953–960.
24. Horsefield,S., Burdett,H., Zhang,X., Manik,M.K., Shi,Y., Chen,J., Qi,T., Gilley,J., Lai,J.-S., Rank,M.X., *et al.* (2019) NAD + cleavage activity by animal and plant TIR domains in cell death pathways. *Science*, **365**, 793–799.
25. Wan,L., Essuman,K., Anderson,R.G., Sasaki,Y., Monteiro,F., Chung,E.H., Nishimura,E.O., DiAntonio,A., Milbrandt,J., Dangl,J.L., *et al.* (2019) TIR domains of plant immune receptors are NAD+-cleaving enzymes that promote cell death. *Science*, **365**, 799–803.
26. Yu,D., Song,W., Yong,E., Tan,J., Wu,B., Schulze-lefert,P., Chai,J., Yu,D., Song,W., Yong,E., *et al.* (2022) TIR domains of plant immune receptors are 2',3'-CAMP/CGMP synthetases mediating cell death. *Cell*, **185**, 2370–2386.
27. Ofir,G., Herbst,E., Baroz,M., Cohen,D., Millman,A., Doron,S., Tal,N., Malheiro,D.B.A., Malitsky,S., Amitai,G., *et al.* (2021) Antiviral activity of bacterial TIR domains via immune signalling molecules. *Nature*, **600**, 116–120.
28. Tal,N., Morehouse,B.B., Millman,A., Stokar-avihail,A., Avraham,C., Fedorenko,T., Yirmiya,E., Herbst,E., Brandis,A., Mehlman,T., *et al.* (2021) Cyclic CMP and cyclic UMP mediate bacterial immunity against phages. *Cell*, **184**, 5728–5739.
29. Hogrel,G., Guild,A., Graham,S., Rickman,H., Grüschow,S., Bertrand,Q., Spagnolo,L. and White,M.F. (2022) Cyclic nucleotide-induced superhelical structure activates a bacterial TIR immune effector. *Nature*, **608**, 808–812.
30. Morehouse,B.R., Govande,A.A., Apurva,A., Millman,A., Keszei,A.F.A., Duncan-lowey,B., Ofir,G., Shao,S., Sorek,R. and Kranzusch,P.J. (2020) STING cyclic dinucleotide sensing originated in bacteria. *Nature*, **586**, 429–433.
31. Gao,L., Altae-Tran,H., Böhning,F., Makarova,K.S., Segel,M., Schmid-Burgk,J.L., Koob,J., Wolf,Y.I., Koonin,E.V. and Zhang,F. (2020) Diverse enzymatic activities mediate antiviral immunity in prokaryotes. *Science*, **369**, 1077–1084.

32. Ma,S., Lapin,D., Liu,L., Sun,Y., Song,W., Zhang,X. and Logemann,E. (2020) Direct pathogen-induced assembly of an NLR immune receptor complex to form a holoenzyme. *Science*, **370** , eabe3069.
33. Ve,T., Williams,S.J. and Kobe,B. (2015) Structure and function of Toll/Interleukin-1 Receptor/Resistance protein (TIR) domains. *Apoptosis*, **20**, 250–261.
34. Martin,R., Qi,T., Zhang,H., Liu,F., King,M., Toth,C., Nogales,E. and Staskawicz,B.J. (2020) Structure of the activated ROQ1 resistosome directly recognizing the pathogen effector XopQ. *Science*, **370**, eabd9993.
35. Morehouse,B.R., Yip,M.C.J., Keszei,A.F.A., McNamara-Bordewick,N.K., Shao,S. and Kranzusch,P.J. (2022) Cryo-EM structure of an active bacterial TIR-STING filament complex. *Nature*, **608**, 803–807.
36. Nimma,S., Gu,W., Maruta,N., Li,Y., Pan,M., Saikot,F.K., Lim,B.Y.J., McGuinness,H.Y., Zaoti,Z.F., Li,S., *et al.* (2021) Structural evolution of TIR-domain signalosomes. *Front. Immunol.*, **12**, 784484.
37. Punjani,A., Rubinstein,J.L., Fleet,D.J. and Brubaker,M.A. (2017) CryoSPARC: algorithms for rapid unsupervised Cryo-EM structure determination. *Nat. Methods*, **14**, 290–296.
38. Mirdita,M., Schütze,K., Moriwaki,Y., Heo,L., Ovchinnikov,S. and Steinegger,M. (2022) ColabFold: making protein folding accessible to all. *Nat. Methods*, **19**, 679–682.
39. Pettersen,E.F., Goddard,T.D., Huang,C.C., Meng,E.C., Couch,G.S., Croll,T.I., Morris,J.H. and Ferrin,T.E. (2021) UCSF chimera X: structure visualization for researchers, educators, and developers. *Protein Sci.*, **30**, 70–82.
40. Emsley,P., Lohkamp,B., Scott,W.G. and Cowtan,K. (2010) Features and development of coot. *Acta. Crystallogr. D Biol. Crystallogr.*, **66**, 486–501.
41. Adams,P.D., Afonine,P.V., Bunkóczi,G., Chen,V.B., Davis,I.W., Echols,N., Headd,J.J., Hung,L.W., Kapral,G.J., Grosse-Kunstleve,R.W., *et al.* (2010) PHENIX: a comprehensive python-based system for macromolecular structure solution. *Acta. Crystallogr. D Biol. Crystallogr.*, **66**, 213–221.
42. Brown,A., Long,F., Nicholls,R.A., Toots,J., Emsley,P. and Murshudov,G. (2015) Tools for macromolecular model building and refinement into electron cryo-microscopy reconstructions. *Acta. Crystallogr. D Biol. Crystallogr.*, **71**, 136–153.
43. Krissinel,E. and Henrick,K. (2007) Inference of macromolecular assemblies from crystalline state. *J. Mol. Biol.*, **372**, 774–797.
44. The PyMOL Molecular Graphics System, Version 2.0, Schrödinger, LLC.
45. Kim,S., Jung,Y. and Lim,D. (2020) Argonaute system of kordia jejudonensis is a heterodimeric nucleic acid-guided nuclease. *Biochem. Biophys. Res. Commun.*, **525**, 755–758.
46. Katoh,K. and Standley,D.M. (2013) MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol. Biol. Evol.*, **30**, 772–780.
47. Roehrl,M.H.A., Wang,J.Y. and Wagner,G. (2004) A general framework for development and data analysis of competitive high-throughput screens for small-molecule inhibitors of protein-protein interactions by fluorescence polarization. *Biochemistry*, **43**, 16056–16066.
48. Elkayam,E., Kuhn,C.D., Tocilj,A., Haase,A.D., Greene,E.M., Hannon,G.J. and Joshua-Tor,L. (2012) The structure of human argonaute-2 in complex with MiR-20a. *Cell*, **150**, 100–110.
49. Swarts,D.C., Szczepaniak,M., Sheng,G., Chandradoss,S.D., Zhu,Y., Timmers,E.M., Zhang,Y., Zhao,H., Lou,J., Wang,Y., *et al.* (2017) Autonomous generation and loading of DNA guides by bacterial argonaute. *Mol. Cell*, **65**, 985–998.
50. Hegge,J.W., Swarts,D.C., Chandradoss,S.D., Cui,T.J., Kneppers,J., Jinek,M., Joo,C. and Van Der Oost,J. (2019) DNA-guided DNA cleavage at moderate temperatures by clostridium butyricum argonaute. *Nucleic Acids Res.*, **47**, 5809–5821.
51. Xu,Y., Tao,X., Shen,B., Horng,T., Medzhitov,R., Manley,J.L. and Tong,L. (2000) Structural basis for signal transduction by the Toll/Interleukin-1 receptor domains. *Nature*, **408**, 111–115.
52. Shi,Y., Kerry,P.S., Nanson,J.D., Bosanac,T., Sasaki,Y., Krauss,R., Saikot,F.K., Adams,S.E., Mosaiab,T., Masic,V., *et al.* (2022) Structural basis of SARM1 activation, substrate recognition, and inhibition by small molecules. *Mol. Cell*, **82**, 1643–1659.
53. 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.
54. Anzelon,R.T.A., Chowdhury,S., Hughes,S.M., Xiao,Y., Lander,G.C., Macrae,I.J., Structural,I., Biology,C., Jolla,L., Biology,C., *et al.* (2020) Structural basis for PiRNA-targeting. *Nature*, **597**, 285–289.
55. Nakanishi,K., Weinberg,D.E., Bartel,D.P. and Patel,D.J. (2012) Structure of yeast argonaute with guide RNA. *Nature*, **486**, 368–374.
56. Schirle,N.T. and MacRae,I.J. (2012) The crystal structure of human Argonaute2. *Science*, **336**, 1033–1037.
57. Manik,M.K., Shi,Y., Li,S., Zaydman,M.A., Damaraju,N., Eastman,S., Smith,T.G., Gu,W., Masic,V., Mosaiab,T., *et al.* (2022) Cyclic ADP ribose isomers: production, chemical structures, and immune signaling. *Science*, **377**, eadc8969.
58. Ve,T., Vajjhala,P.R., Hedger,A., Croll,T., Dimaio,F., Horsefield,S., Yu,X., Lavrencic,P., Hassan,Z., Morgan,G.P., *et al.* (2017) Structural basis of TIR-domain-assembly formation in MAL- and MyD88-dependent TLR4 signaling. *Nat. Struct. Mol. Biol.*, **24**, 743–751.
59. Clabbers,M.T.B., Holmes,S., Muusse,T.W., Vajjhala,P.R., Thygesen,S.J., Malde,A.K., Hunter,D.J.B., Croll,T.I., Flueckiger,L., Nanson,J.D., *et al.* (2021) MyD88 TIR domain higher-order assembly interactions revealed by microcrystal electron diffraction and serial femtosecond crystallography. *Nat. Commun.*, **12**, 2578.
60. Shen,Z., Yang,X.-Y., Xia,S., Huang,W., Taylor,D.J., Nakanishi,K. and Fu,T.-M. (2023) Oligomerization-mediated activation of a short prokaryotic Argonaute. *Nature*, **621**, 154–161.
61. Wang,X., Li,X., Yu,G., Zhang,L., Zhang,C., Wang,Y., Liao,F. and Wen,Y. (2023) Structural insights into mechanisms of Argonaute protein-associated NADase activation in bacterial immunity. *Cell Res.*, **33**, 699–711.
62. Ni,D., Lu,X., Stahlberg,H. and Ekundayo,B. (2023) Activation mechanism of a short Argonaute-TIR prokaryotic immune system. *Sci. Adv.*, **9**, eadh9002.
63. Guo,M., Zhu,Y., Lin,Z., Yang,D., Zhang,A., Guo,C. and Huang,Z. (2023) Cryo-EM structure of the ssDNA-activated SPARTA complex. *Cell Res.*, **33**, 731–734.
64. Kottur,J., Malik,R. and Aggarwal,A.K. (2023) Nucleic acid mediated activation of a short prokaryotic Argonaute immune system. bioRxiv doi: https://doi.org/10.1101/2023.09.17.558117, 17 September 2023, preprint: not peer reviewed.
65. Gao,X., Shang,K., Zhu,K., Wang,L., Mu,Z., Fu,X., Yu,X., Qin,B., Zhu,H., Ding,W., *et al.* (2023) Nucleic acid-triggered NADase activation of a short prokaryotic Argonaute. *Nature*, https://doi.org/10.1038/s41586-023-06665-6.
66. Zhang,J.-T., Wei,X.-Y., Cui,N., Tian,R. and Jia,N. (2023) Target ssDNA activates the NADase activity of prokaryotic SPARTA immune system. *Nat. Chem. Biol.*, https://doi.org/10.1038/s41589-023-01479-z.
67. Guo,L., Huang,P., Li,Z., Shin,Y.C., Yan,P., Lu,M., Chen,M. and Xiao,Y. (2023) Auto-inhibition and activation of a short Argonaute-associated TIR-APAZ defense system. *Nat. Chem. Biol.*, https://doi.org/10.1038/s41589-023-01478-0.
68. Slaymaker,I.M., Gao,L., Zetsche,B., Scott,D.A., Yan,W.X. and Zhang,F. (2016) Rationally engineered Cas9 nucleases with improved specificity. *Science (1979)*, **351**, 84–88.
69. Nishimasu,H., Shi,X., Ishiguro,S., Gao,L., Hirano,S., Okazaki,S., Noda,T., Abudayyeh,O.O., Gootenberg,J.S., Mori,H., *et al.* (2018) Engineered CRISPR-Cas9 nuclease with expanded targeting space. *Science*, **361**, 1259–1262.

70. Ma,E., Chen,K., Shi,H., Stahl,E.C., Adler,B., Trinidad,M., Liu,J., Zhou,K., Ye,J. and Doudna,J.A. (2022) Improved genome editing by an engineered CRISPR-Cas12a. *Nucleic Acids Res.*, **50**, 12689–12701.

71. Hu,J.H., Miller,S.M., Geurts,M.H., Tang,W., Chen,L., Sun,N., Zeina,C.M., Gao,X., Rees,H.A., Lin,Z., *et al.* (2018) Evolved Cas9 variants with broad PAM compatibility and high DNA specificity. *Nature*, **556**, 57–63.

Received: September 25, 2023. **Revised:** December 14, 2023. **Editorial Decision:** December 15, 2023. **Accepted:** December 19, 2023