nature microbiology

Article

https://doi.org/10.1038/s41564-022-01239-0

Short prokaryotic Argonautes provide defence against incoming mobile genetic elements through NAD$^+$ depletion

Received: 15 December 2021

Accepted: 26 August 2022

Published online: 3 October 2022

Check for updates

Mindaugas Zaremba[1,6] ✉, Donata Dakineviciene[1,6], Edvardas Golovinas[1,6], Evelina Zagorskaitė[1,6], Edvinas Stankunas[1,3], Anna Lopatina[2,4], Rotem Sorek[2], Elena Manakova[1], Audrone Ruksenaite[1], Arunas Silanskas[1], Simonas Asmontas[1], Algirdas Grybauskas[1], Ugne Tylenyte[1], Edvinas Jurgelaitis[1], Rokas Grigaitis[1,5], Kęstutis Timinskas[1], Česlovas Venclovas[1] and Virginijus Siksnys[1] ✉

Argonaute (Ago) proteins are found in all three domains of life. The so-called long Agos are composed of four major domains (N, PAZ, MID and PIWI) and contribute to RNA silencing in eukaryotes (eAgos) or defence against invading mobile genetic elements in prokaryotes (pAgos). The majority (~60%) of pAgos identified bioinformatically are shorter (comprising only MID and PIWI domains) and are typically associated with Sir2, Mrr or TIR domain-containing proteins. The cellular function and mechanism of short pAgos remain enigmatic. Here we show that *Geobacter sulfurreducens* short pAgo and the NAD$^+$-bound Sir2 protein form a stable heterodimeric complex. The GsSir2/Ago complex presumably recognizes invading plasmid or phage DNA and activates the Sir2 subunit, which triggers endogenous NAD$^+$ depletion and cell death, and prevents the propagation of invading DNA. We reconstituted NAD$^+$ depletion activity in vitro and showed that activated GsSir2/Ago complex functions as a NADase that hydrolyses NAD$^+$ to ADPR. Thus, short Sir2-associated pAgos provide defence against phages and plasmids, underscoring the diversity of mechanisms of prokaryotic Agos.

Being at the core of RNA interference, eukaryotic Argonaute (Ago) proteins (eAgos) are involved in the regulation of gene expression, silencing of mobile genome elements and defence against viruses[1,2]. The best-studied human Argonaute-2 uses small RNA molecules as guides for target RNA recognition, and eAgos are similar both structurally and mechanistically[2–5]. Monomeric eAgos are composed of four major domains: N, PAZ (Piwi/Argonaute/Zwille), MID (middle) and PIWI (P-element induced wimpy testis) (Fig. 1a) and share a bilobed structure, where the N- and C-terminal lobes are formed by conserved N/PAZ and MID/PIWI domains, respectively[4–7]. The N-domain acts as a wedge that separates guide and target strands[3,7], while the MID and PAZ domains respectively bind the 5′ and 3′ terminus of the guide (g)RNA, located between the N- and C-terminal lobes[4,6]. eAgos can slice target RNA through endonucleolytic cleavage by the PIWI domain or inhibit translation through RNA binding by the catalytically inactive eAgos that may also trigger RNA decay by auxiliary cellular nucleases[2,4,5].

[1]Institute of Biotechnology, Life Sciences Center, Vilnius University, Vilnius, Lithuania. [2]Department of Molecular Genetics, Weizmann Institute of Science, Rehovot, Israel. [3]Present address: Max Perutz Labs, Medical University of Vienna, Vienna Biocenter Campus (VBC), Vienna, Austria. [4]Present address: Division of Microbial Ecology, Center for Microbiology and Environmental Systems Science, University of Vienna, Vienna, Austria. [5]Present address: VU LSC-EMBL Partnership for Genome Editing Technologies, Life Sciences Center, Vilnius University, Vilnius, Lithuania. [6]These authors contributed equally: Mindaugas Zaremba, Donata Dakineviciene, Edvardas Golovinas, Evelina Zagorskaitė. ✉e-mail: zare@ibt.lt; siksnys@ibt.lt

https://doi.org/10.1038/s41564-022-01239-0

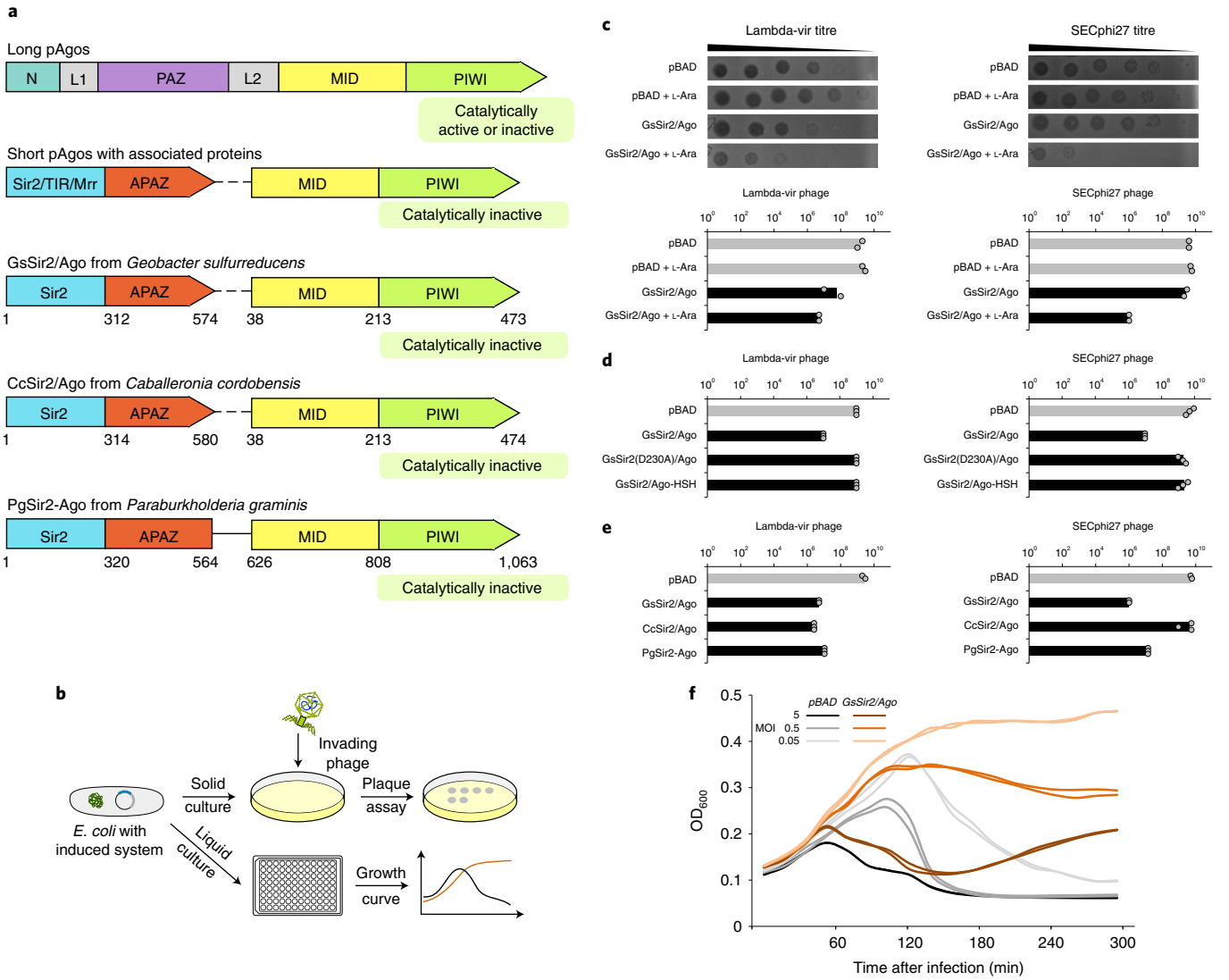


Fig. 1 | Sir2/Ago provides defence against phages. a, Schematic representation of the subunit/domain composition of different pAgo variants. Catalytically active pAgos contain a conserved catalytic DEDX tetrad that is mutated in inactive pAgos. Catalytically inactive short pAgos lack the canonical PAZ domain; however, an accessory APAZ domain is present in putative Sir2, TIR or Mrr proteins associated with short pAgos. MID, middle; L, linker domain; N,N-terminal domain. Short pAgos from *G. sulfurreducens*, *C. cordobensis* and *P. graminis* associated with Sir2 protein were studied in this work. **b**, Schematic diagram of phage restriction assays. **c**, Bottom: EOP of lambda-vir and SECphi27 phages infecting *E. coli* cells with and without the GsSir2/Ago system. The bar graphs show the number of plaque-forming units (p.f.u.) as arithmetic means of two replicates in the absence and the presence of the inducer L-arabinose (L-Ara), with individual data points superimposed. Grey and black bars respectively represent EOP on pAgo-lacking and pAgo-containing cells. Top: representative images of plaque assays. **d**, EOP of lambda-vir and SECphi27 phages infecting the wt and mutant GsSir2/Ago systems in the presence of L-Ara. GsSir2(D230A)/Ago and GsSir2/Ago-HSH are variants that contain a D230A mutation in the Sir2 domain or an HSH-tag on the C terminus of pAgo, respectively. The bar graphs show the number of p.f.u. as arithmetic means of three replicates, with individual data points superimposed. Grey and black bars respectively represent EOP on pAgo-lacking and pAgo-containing cells. **e**, EOP of lambda-vir and SECphi27 phages infecting pAgo-lacking cells and cells containing GsSir2/Ago, CcSir2/Ago and PgSir2-Ago in the presence of L-Ara. The bar graphs show the number of p.f.u. as arithmetic means of three replicates, with individual data points superimposed. Grey and black bars respectively represent EOP on pAgo-lacking and pAgo-containing cells. **f**, Lambda phage infection in liquid cultures of *E. coli* cells containing the GsSir2/Ago system. GsSir2/Ago-lacking (shown in grey) or GsSir2/Ago-containing (shown in orange) *E. coli* were infected at $t = 0$ at MOI of 0.05, 0.5 and 5. Each curve represents one individual replicate; two replicates for each MOI are shown.

pAgos are quite widespread and are present in 9% of sequenced bacterial and 32% of archaeal genomes[6]. So far, more than ~1,000 pAgos have been identified bioinformatically, revealing a striking diversity. Moreover, pAgos are often associated with additional putative nucleases, helicases and DNA binding proteins that are not linked to eAgos[8]. pAgos are divided into full-length or long pAgos (~40%) sharing conserved N, PAZ, MID and PIWI domain architecture with eAgos, and short pAgos (~60%) composed only of MID and PIWI domains (Fig. 1a)[8,9]. Long pAgos are relatively well-characterized both structurally and functionally[10,11] and, similar to eAgos, contain either a catalytically active or inactive PIWI domain. Some long pAgos with the catalytically active PIWI, as exemplified by CbAgo and TtAgo, use DNA guides to target and cleave DNA providing defence against invading phages or plasmids, or contributing to chromosome segregation after replication, respectively[12–14]. Meanwhile, a long RsAgo from *Rhodobacter spaeroides* with an inactive PIWI domain guided by small RNA is thought to mobilize an unknown cellular nuclease(s) for degradation of invading plasmids and mobile genetic elements[9,15]. Interestingly, long KmAgo from *Kurthia massiliensis* can use both DNA and RNA guides to target DNA and RNA in vitro, albeit with different efficiencies[16,17]. In

https://doi.org/10.1038/s41564-022-01239-0

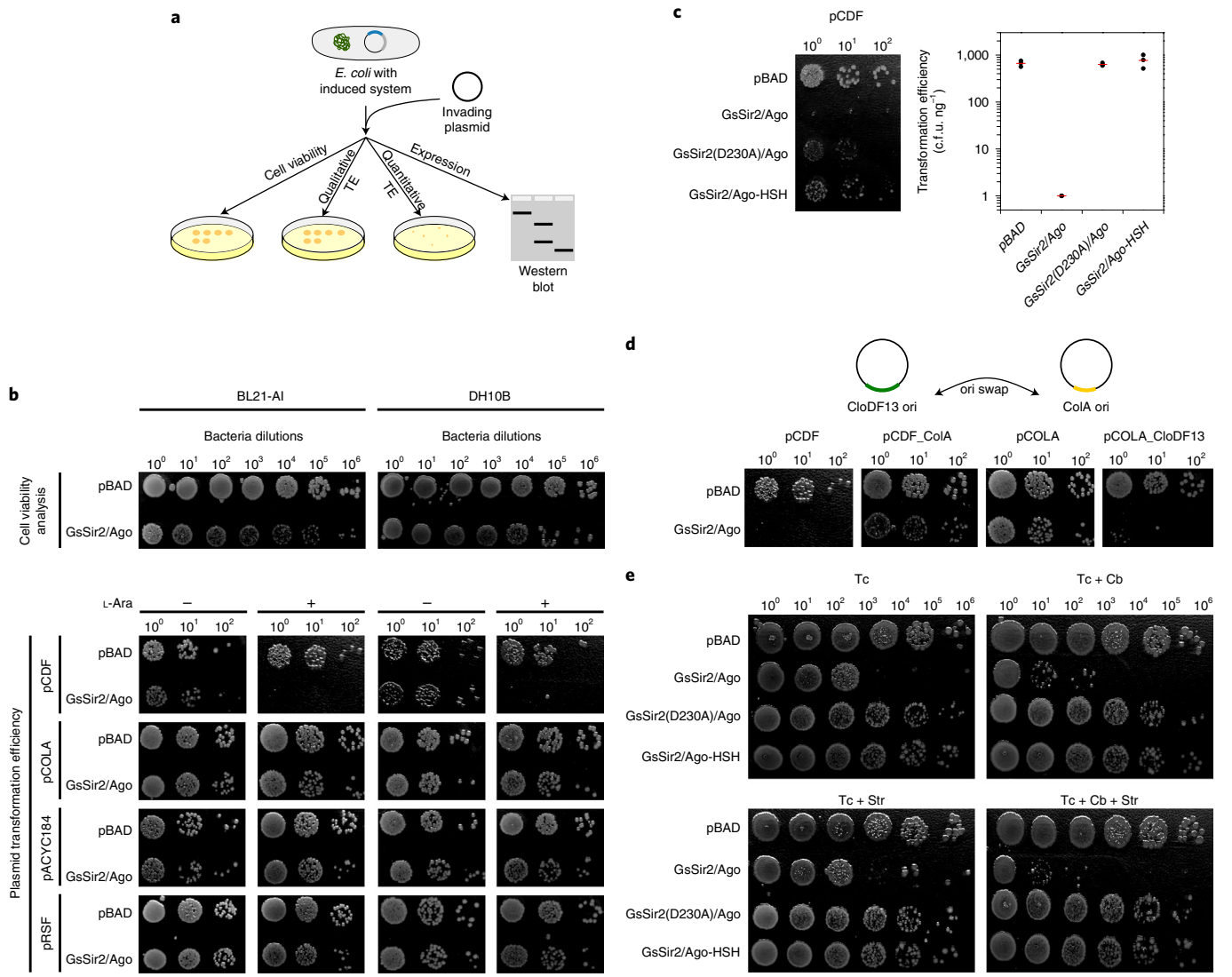


Fig. 2 | The GsSir2/Ago system interferes with plasmid transformation. **a**, Schematic representation of the experiment. TE, transformation evaluation. **b**, Qualitative evaluation of plasmid transformation efficiency in *E. coli* cells carrying the GsSir2/Ago system. Top: comparison of cell viability in the presence or absence of plasmid-borne GsSir2/Ago expression. Bottom: comparison of plasmid transformation efficiencies in the presence or absence of plasmid-borne GsSir2/Ago expression. **c**, Left: comparison of pCDF transformation efficiency between cells expressing wt and mutant GsSir2/Ago complexes. Right: quantification of transformation efficiencies (three independent replicates, the red line represents average transformation efficiency). **d**, Top: schematic representation of *ori* exchange between pCDF and pCOLA plasmids. Bottom: comparison of plasmid transformation efficiencies. pCDF, pCDF with CloDF13 *ori* exchanged with ColA *ori* (pCDF_ColA), pCOLA and pCOLA with ColA *ori* exchanged with CloDF13 *ori* (pCOLA_CloDF13) plasmids were used for transformation of *E. coli* cells carrying the GsSir2/Ago system. **e**, Cell viability in the absence of antibiotic selection. In the case of the wt GsSir2/Ago system, cell viability decreased on the plates even in the absence of Cb and Str antibiotics, suggesting that GsSir2/Ago in the presence of the pCDF plasmid triggers cell death.

contrast to long pAgos, all short pAgos possess a catalytically inactive PIWI domain and are typically associated with proteins containing a domain initially thought to be analogous to PAZ (APAZ)[18]. Subsequently, however, it was proposed that APAZ may actually be homologous to the N-terminal domains (N and L1) of Ago[11,19]. APAZ-containing proteins are often fused to Sir2 (Silent informator regulator 2), Mrr nucleases or TIR (Toll-interleukin-1 receptor) domains[8,18]. About half of all identified short pAgos are associated with or fused into a single-chain protein with Sir2-APAZ proteins[8]. The Sir2 domain-containing proteins are widely distributed in all domains of life and perform protein deacetylation or ADP-ribosylation functions using NAD^+ as a co-factor[20,21]. Particularly, bacterial Sir2 proteins are involved in many cellular processes including transcription, translation, carbon and nitrogen metabolism, virulence and resistance to stress[21]. Even though short pAgos, half of which are associated with Sir2 proteins, make up the majority of all pAgos, their function in the cell and in vitro remains to be established.

Results

In this study, we aimed to explore whether short pAgos can act as prokaryotic defence systems against viruses or plasmids. To this end, we selected two short pAgos, GsSir2/Ago from *Geobacter sulfurreducens* and CcSir2/Ago from *Caballeronia cordobensis*, each encoded in a putative operon together with a Sir2 domain protein, and PgSir2-Ago from *Paraburkholderia graminis*, representing a fusion of Sir2 and pAgo (Fig. 1a). The coding regions of the Sir2 and pAgo proteins in GsSir2/Ago and CcSir2/Ago systems overlap by 11 and 8 bp, respectively, indicating that they belong to the same operon (Supplementary Note). Next, we engineered heterologous *E. coli* cells by cloning GsSir2/Ago and CcSir2/

https://doi.org/10.1038/s41564-022-01239-0

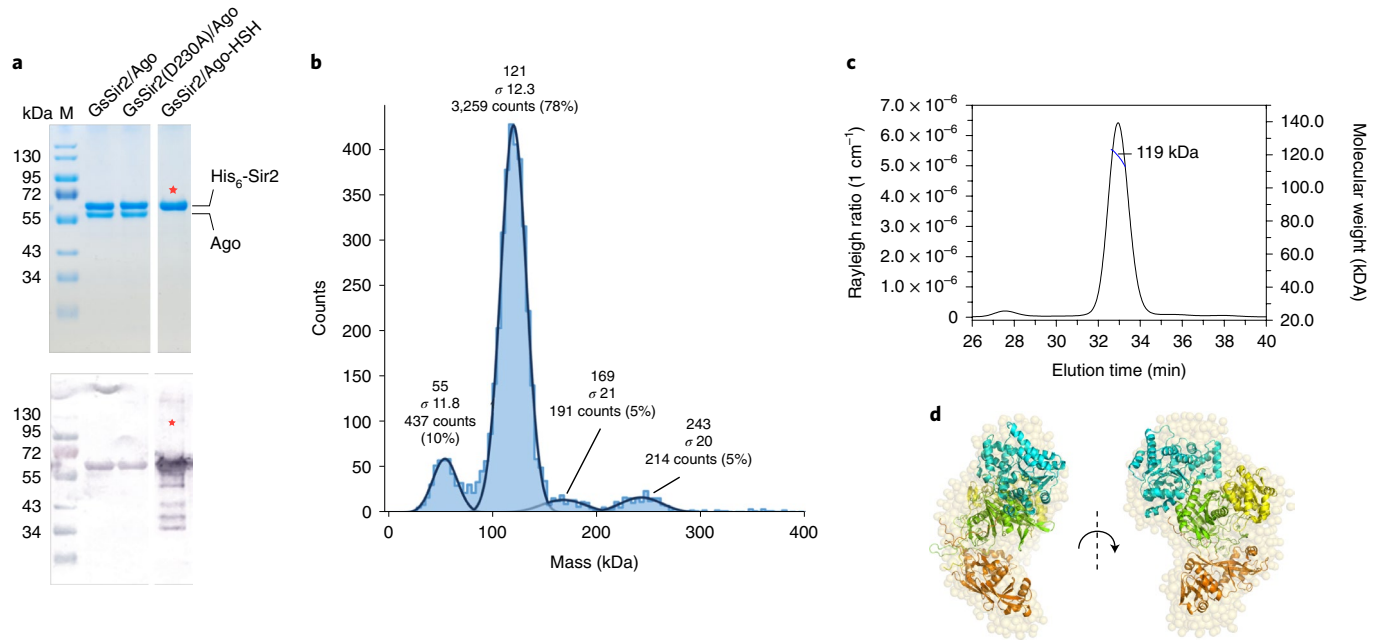


Fig. 3 | The GsSir2 and GsAgo proteins form a heterodimeric complex. a, Top: Sodium dodecyl sulfate–polyacrylamide gel electrophoresis (SDS–PAGE) of purified wt GsSir2/Ago, a D230A mutant, and C-terminal HSH-tag-bearing GsSir2/Ago. The red star in the HSH-tagged sample lane marks an overlap of bands in the gel due to similar mass. Bottom: anti-His-tag western blot of same samples. The red star shows the lane where the His-tag is on the C terminus of Ago, rather than the N terminus of Sir2. Two replicates. **b**, Mass photometry data of the GsSir2/Ago complex, with masses and respective particle population counts indicated. According to mass spectrometry of the purified GsSir2/Ago complex (Extended Data Fig. 3f), the molar mass of the GsSir2/Ago heterodimer is 121 kDa. **c**, SEC-MALS data of GsSir2/Ago showing the chromatography peak and molar weight of the Sir2/Ago heterodimer. **d**, Semitransparent space-filling ab initio model of GsSir2/Ago calculated from SAXS data with a fitted-in AlphaFold GsSir2/Ago model in cartoon representation. Colour coding: cyan, Sir2 domain; brown, APAZ; yellow, MID; green, PIWI.

Ago genes, or a single gene, encoding PgSir2-Ago into pBAD expression vectors under a P_{BAD} promoter (Supplementary Table 1) and challenged them with phages or plasmids.

Sir2/Ago systems provide defence against phages

To test whether the GsSir2/Ago system provides defence against phages, we challenged *E. coli* host carrying the GsSir2/Ago system with a set of six *E. coli* phages spanning four morphological families including *Podoviridae* (T7), *Siphoviridae* (lambda-vir, SECphi27, SECphi18), *Myoviridae* (T4) and *Microviridae* (SECphi17, a single-stranded (ss)DNA phage) (Supplementary Table 1). We measured the efficiency of plating (EOP) of these phages with and without L-arabinose induction of the GsSir2/Ago system. The system showed protection against two out of six phages: lambda-vir (~100-fold) and SECphi27 (~1,000-fold) (Fig. 1c).

To probe the role of individual Sir2 and Ago proteins in antiviral defence, we engineered *E. coli* cells carrying mutant variants of either the Sir2 or the Ago protein of the GsSir2/Ago system and performed small-drop plaque assays using the lambda-vir and SECphi27 phages. In Sir2 variants, the highly conserved D230 residue, presumably involved in NAD^+ binding, was replaced by alanine (Extended Data Fig. 1c)[18]. Phage challenge assay revealed that the D230A mutation completely abolished defence against both phages (Fig. 1d). The Ago protein is catalytically inactive due to the active site mutations in the PIWI domain; therefore, to obtain a binding-deficient Ago variant, we fused a bulky 29 amino acid **H**is$_6$-**S**trepII-**H**is$_6$-tag (HSH-tag) at the C terminus that is important for nucleic acid binding in other Agos[3,6,8,22]. We found that Ago C-terminal modification abolished protection from both phages (Fig. 1d). Thus, both the Sir2 and the Ago proteins are required for protection against phages by the GsSir2/Ago system.

Additionally, we probed the ability of homologous CcSir2/Ago and PgSir2-Ago systems to restrict lambda-vir and SECphi27 phages (Fig. 1e). After L-arabinose induction, the PgSir2-Ago system showed ~500-fold protection against lambda phage and ~400-fold protection against SECphi27, while the CcSir2/Ago system showed ~1,000-fold protection against lambda and no protection against SECphi27 phage. Thus, homologous Sir2/Ago systems from three phylogenetically distant bacteria showed protection against phage infection, albeit with different efficiency and specificity.

Next, *E. coli* MG1655 cells either lacking or containing the GsSir2/Ago system were infected in liquid culture with lambda phage at a multiplicity of infection (MOI) of 0.05, 0.5 and 5 (Fig. 1f). At high MOI where, on average, a single bacterial cell is infected by a single phage, the culture collapses, while at low MOI the culture survives. This phenotype implies that GsSir2/Ago-mediated defence triggers cell death at approximately the same time when the phage-induced lysis occurs.

Sir2/Ago systems interfere with plasmid transformation

Further, to test whether heterologous expression of the GsSir2/Ago, CcSir2/Ago and PgSir2-Ago systems in *E. coli* cells (strains BL21-AI and DH10B) provides a barrier for plasmid transformation, four plasmids (pCDF, pCOLA, pACYC184 and pRSF) with different *ori* regions and copy numbers were used in a plasmid interference assay (Supplementary Table 1, Fig. 2 and Extended Data Fig. 2). We found that the GsSir2/Ago system prevented only pCDF plasmid transformation, reducing its efficiency nearly ~100-fold (Fig. 2b). Next, we tested pCDF transformation efficiency in *E. coli* cells expressing the GsSir2/Ago mutants (Fig. 2c). Both Sir2 D230A mutation and pAgo HSH-tag modification that impaired phage restriction also abolished plasmid interference making cells permissive to pCDF plasmid transformation (Fig. 2c). Similar to GsSir2/Ago, the CcSir2/Ago system provided resistance only for pCDF plasmid transformation (Extended Data Fig. 2b), while cells carrying the single-chain PgSir2-Ago system were permissive to transformation of all four plasmids (Extended Data Fig. 2b).

https://doi.org/10.1038/s41564-022-01239-0

Table 1 | Nucleic acid binding by Sir2/Ago and their binary complexes pre-loaded with either ssRNA or ssDNA guides

Nucleic acid binding by Sir2/Ago		
Nucleic acid	**K_d, nM**	
	GsSir2/Ago	**CcSir2/Ago**
ssDNA	0.033±0.016	0.138±0.007
ssRNA	0.042±0.015	0.36±0.098
circular ssDNA	0.072±0.018	0.18±0.068
dsRNA/DNA	4.31±0.37	2.4±0.69
dsDNA	51±20	330±87
dsRNA	73±29	81±15

Nucleic acid binding by GsSir2/Ago-gRNA and GsSir2/Ago-gDNA complexes			
Guide	**Target**	**K_d, nM**	
		−Heparin	**+Heparin**
ssRNA	ssDNA	0.020±0.008	0.079±0.007
	ssRNA	0.205±0.058	0.567±0.112
ssDNA	ssDNA	0.022±0.005	n.b.[a]
	ssRNA	0.252±0.114	n.b.[a]
ssRNA	nsp-ssDNA[b]	0.420±0.231	4.29±2.11

ssDNA binding by GsSir2/Ago-gRNA variants		
Mutant	**K_d, nM**	
	−Heparin	**+Heparin**
GsSir2(D230A)/Ago	0.038±0.013	0.140±0.057
GsSir2/Ago-HSH	0.140±0.017	0.203±0.058
GsSir2^{APAZ}/Ago	n.b.[a]	n.b.[a]
GsSir2/Ago^{PIWI}	n.b.[a]	n.b.[a]

[a]No binding was detected under experimental conditions used.[b]A non-complementary ssDNA was used.K_d values (averages±s.d. of three independent replicates) were calculated from EMSA data. Some binding experiments were performed in the presence of heparin (indicated).

Interestingly, although the GsSir2/Ago system in *E. coli* interfered with the pCDF plasmid transformation, the pCOLA plasmid that differs mainly in the *ori* region and antibiotic resistance gene was permissive. To test whether the *ori* region determines differences in the transformation efficiency between pCDF and pCOLA plasmids, we swapped the *ori* sequences of pCDF and pCOLA (Fig. 2d). The pCDF plasmid with ColA *ori* instead of CloDF13 became permissive in *E. coli* cells expressing the GsSir2/Ago system, whereas transformation of pCOLA plasmid bearing CloDF13 *ori* instead of ColA *ori* was prevented. These results indicate that the CloDF13 *ori* is a key element that controls plasmid transformation efficiency in *E. coli* cells expressing the GsSir2/Ago system.

The plasmid interference by the GsSir2/Ago system could be due to either plasmid entry exclusion, replication inhibition or plasmid degradation. To eliminate the possible role of plasmid entry barriers on the pCDF plasmid transformation efficiency, we engineered heterologous *E. coli* cells carrying two plasmids: pBAD plasmid providing carbenicillin (Cb) resistance and expressing GsSir2/Ago (or its mutants) under the control of P_{BAD} inducible promoter and the pCDF plasmid providing streptomycin (Str) resistance, and tested cell viability in the presence or absence of the inducer. In this case, the pCDF plasmid is already in the cell and provides streptomycin resistance; however, antibiotic resistance should be lost if the plasmid is restricted after GsSir2/Ago expression. In the absence of induction, cell viability of *E. coli* cells carrying wild-type (wt) GsSir2/Ago (or its mutants) and an empty pBAD vector (Extended Data Fig. 2e) was identical. In the presence of the inducer, the viability of cells expressing the wt GsSir2/Ago system, but not its mutants, significantly decreased (Fig. 2e), indicating that GsSir2/Ago interferes with the pCDF plasmid already present in the cell. Notably, a decrease in cell viability is observed in *E. coli* BL21-AI cells (tetracycline-resistant) without the cell selection for Str and Cb resistance. It cannot be ruled out that upon recognition of pCDF, the GsSir2/Ago system becomes activated and triggers cell death. A similar cell death phenotype triggered by the GsSir2/Ago was observed during phage infection in liquid cultures (Fig. 1f). Taken together, these data show that the GsSir2/Ago system acts as a defence system against phages and plasmids via cell death or suicidal mechanism.

Short pAgo and Sir2 form a stable heterodimeric complex

To characterize the Sir2/Ago systems biochemically, we aimed to express individual Sir2 and Ago proteins in *E. coli*. The GsSir2 and CcSir2 proteins (but not PgSir2-Ago) were expressed and purified by chromatography (Fig. 3a and Extended Data Fig. 3). The N-terminal His_6-tagged GsSir2 and CcSir2 proteins co-expressed with Ago proteins co-purified on the Ni^{2+}-affinity column (Extended Data Fig. 3), indicating that Sir2 and pAgo proteins form a stable complex. We failed to express Sir2 and Ago proteins individually, suggesting that they form an obligatory Sir2/Ago complex. Functionally compromised GsSir2(D230A)/Ago and GsSir2/Ago-HSH variants also formed a complex, indicating that while the introduced mutations abolished the activity in vivo, they did not affect the protein complex structure (Fig. 2c and Extended Data Fig. 3c). Further analysis of the oligomeric state of GsSir2/Ago and CcSir2/Ago complexes in solution using multi-angle light scattering coupled with size exclusion chromatography (SEC-MALS), mass photometry and small-angle X-ray scattering (SAXS) showed that heterodimeric complexes are formed in a wide range of protein concentrations (from 20 nM to 6.5 μM) (Fig. 3b–d and Extended Data Fig. 4). According to the SAXS data, the heterodimeric GsSir2/Ago complex acquires a notably asymmetric shape (Fig. 3d, Extended Data Fig. 4 and Supplementary Table 3) that is consistent with a structural model of the heterodimer (Supplementary Note). In summary, the results show that pAgos and associated Sir2 proteins encoded by a single operon form a stable heterodimeric complex.

Sir2/Ago prefers an RNA guide to bind a DNA target

Long pAgos use ssDNA and/or ssRNA guides to recognize their complementary DNA and/or RNA targets[10]. To establish the guide preference of Gs and Cc Sir2/Ago, we analysed binding of single- or double-stranded (ds)DNA or RNA by electrophoretic mobility shift assay (EMSA) (Extended Data Fig. 6). Both Sir2/Ago heterodimers showed a strong preference for ssDNA and ssRNA binding. RNA/DNA heteroduplex was bound with an intermediate affinity, while dsRNA or dsDNA showed only weak binding (Extended Data Fig. 6a,b,e and Table 1). Interestingly, neither the 5′-terminal phosphate nor the 3′-OH end or Mg^{2+} ions were required for ssDNA binding as both Gs and Cc complexes bound the circular ssDNA and linear oligonucleotides with similar affinity (Extended Data Fig. 6a,b). As expected, Gs and Cc Sir2/Ago containing the inactivated PIWI domains showed no cleavage activity for any nucleic acid (NA) substrate tested (Extended Data Fig. 6f). In summary, EMSA experiments suggest that in vitro ssRNA or ssDNA are preferable GsSir2/Ago guides.

Next, we analysed ssDNA or ssRNA target binding by the binary GsSir2/Ago complexes pre-loaded with either ssRNA or ssDNA guides. In a separate set of experiments, reaction mixtures also contained heparin, a competitor of nucleic acid binding (Fig. 4a and Table 1). The binary GsSir2/Ago-ssRNA complex showed ~10-fold better binding to the ssDNA than to ssRNA, and heparin addition had only a little effect (~4-fold decrease) on binding affinity in this case. The GsSir2/Ago-ssDNA binary complex bound to the matching ssDNA target with affinity similar to the GsSir2/Ago-ssRNA binary complex; however, heparin addition abolished binding (Extended Data Fig. 6a,d). Furthermore, the binary

https://doi.org/10.1038/s41564-022-01239-0

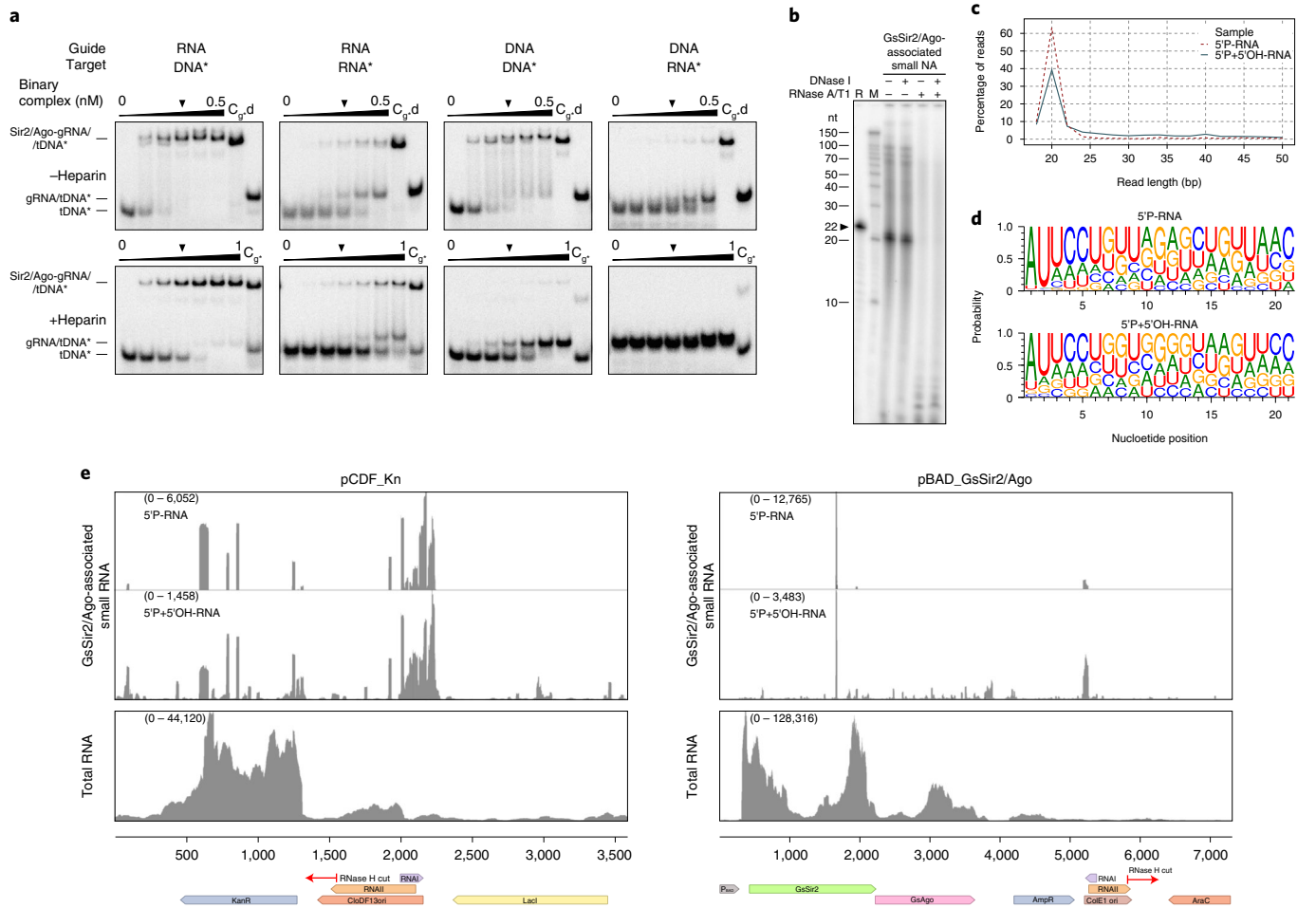


Fig. 4 | Nucleic acid binding by GsSir2/Ago in vitro and in vivo. a, Binding of RNA or DNA targets by GsSir2/Ago binary complexes pre-loaded with 5′P-RNA or 5′P-DNA guides. In EMSA experiments, the pre-formed GsSir2/Ago-gNA binary complex was mixed with a radiolabelled target strand indicated by the asterisk (see Materials and Methods for details). To show that no displacement of the guide by the target strand occurs under these experimental conditions, a control (C_g*) experiment was performed where the guide, rather than the target, was radioactively labelled. Only the pre-annealed RNA/DNA heteroduplex was loaded in the control lane 'd'. Three independent replicates were performed. **b**, GsSir2/Ago co-purifies with small RNAs. Nucleic acids that co-purified with GsSir2/Ago were first dephosphorylated, then [γ-^{32}P]-ATP radiolabelled and treated with DNase I or RNase A/T1, or both, and resolved on a denaturing polyacrylamide gel. R, control 22 nt RNA oligonucleotide; M, RNA ladder Decade Marker System (Ambion). Three independent replicates were performed. **c**, Length distribution of small RNA co-purified with GsSir2/Ago as determined by sequencing. In the '5′P-RNA' sample, only small RNAs containing 5′-phosphate were sequenced, while in the '5′P+5′OH-RNA' sample, both 5′-phosphate or 5′-OH-bearing small RNAs were sequenced. **d**, Small RNAs associated with GsSir2/Ago show 5′-AU preference. **e**, Top: distribution of small RNAs co-purified with GsSir2/Ago from the *E. coli* host carrying the pCDF_Kn target (left) and pBAD_GsSir2/Ago expression plasmids (right). Bottom: corresponding IGV viewer representations of total RNA extracted from *E. coli*. Cartoons indicate promoters (P_{BAD}), protein-coding genes (KanR, LacI, GsSir2, GsAgo, AmpR, AraC), plasmid *ori* (CloDF13, ColE1) and their RNAI and RNAII transcripts. Red arrow shows RNase H cleavage site in RNAII required for initiation of DNA strand synthesis during plasmid replication.

GsSir2/Ago-ssRNA complex bound to the complementary ssDNA target ~200-fold better than apo-GsSir2/Ago-bound pre-annealed gRNA/tDNA heteroduplex, indicating that GsSir2/Ago requires binding of the RNA guide first to interact with the DNA target (Fig. 4a and Table 1). The binding affinity of the functionally compromised in vivo GsSir2(D230A)/Ago mutant was similar to the wt, while the binding affinity of the GsSir2/Ago-HSH mutant was slightly (~7-fold) weaker (Fig. 4a, Extended Data Fig. 6c and Table 1). Charge reversal mutations of positively charged residues that are involved in interactions with the guide and target NAs according to the GsSir2/Ago model abolished both the target NA binding and the pCDF plasmid interference (Extended Data Figs. 2f,g, 3b and 6c). Taken together, our data suggest that GsSir2/Ago uses ssRNA as a guide for the recognition of an ssDNA target.

To identify NAs bound by GsSir2/Ago in vivo, we purified the GsSir2/Ago-NA complex from *E. coli* transformed with the pBAD_GsSir2/Ago expression vector and the pCDF target plasmid, extracted NAs and subjected them to sequencing. Subsequent analysis revealed that GsSir2/Ago is associated with small (predominantly 21 nt) RNAs with or without the 5′-phosphate (Fig. 4b,c). Different from other Argonaute proteins that show base selectivity for the first nucleotide at the 5′-end of the guide[13–15,23], GsSir2/Ago-associated small RNAs show preference for the 5′-AU dinucleotide (Fig. 4d). This preference is more pronounced for small RNAs containing 5′-phosphate, implying that GsSir2/Ago uses as a guide small RNAs containing the 5′-AU dinucleotide. Most co-purified small RNAs (~95%) matched the *E. coli* genome, whereas the smaller fraction (~5%) originated from the pBAD_GsSir2/Ago and pCDF plasmids (Supplementary Data 1). Interestingly, small plasmid-borne RNAs that matched CloDF13 and ColE1 *ori* regions of corresponding plasmids were noticeably enriched (Fig. 4e). Taken together, RNA-seq data suggest that GsSir2/Ago could use small 5′-AU-RNAs originating from the invader transcripts (for example, pCDF *ori* region) as guides to target the invaders' DNA.

https://doi.org/10.1038/s41564-022-01239-0

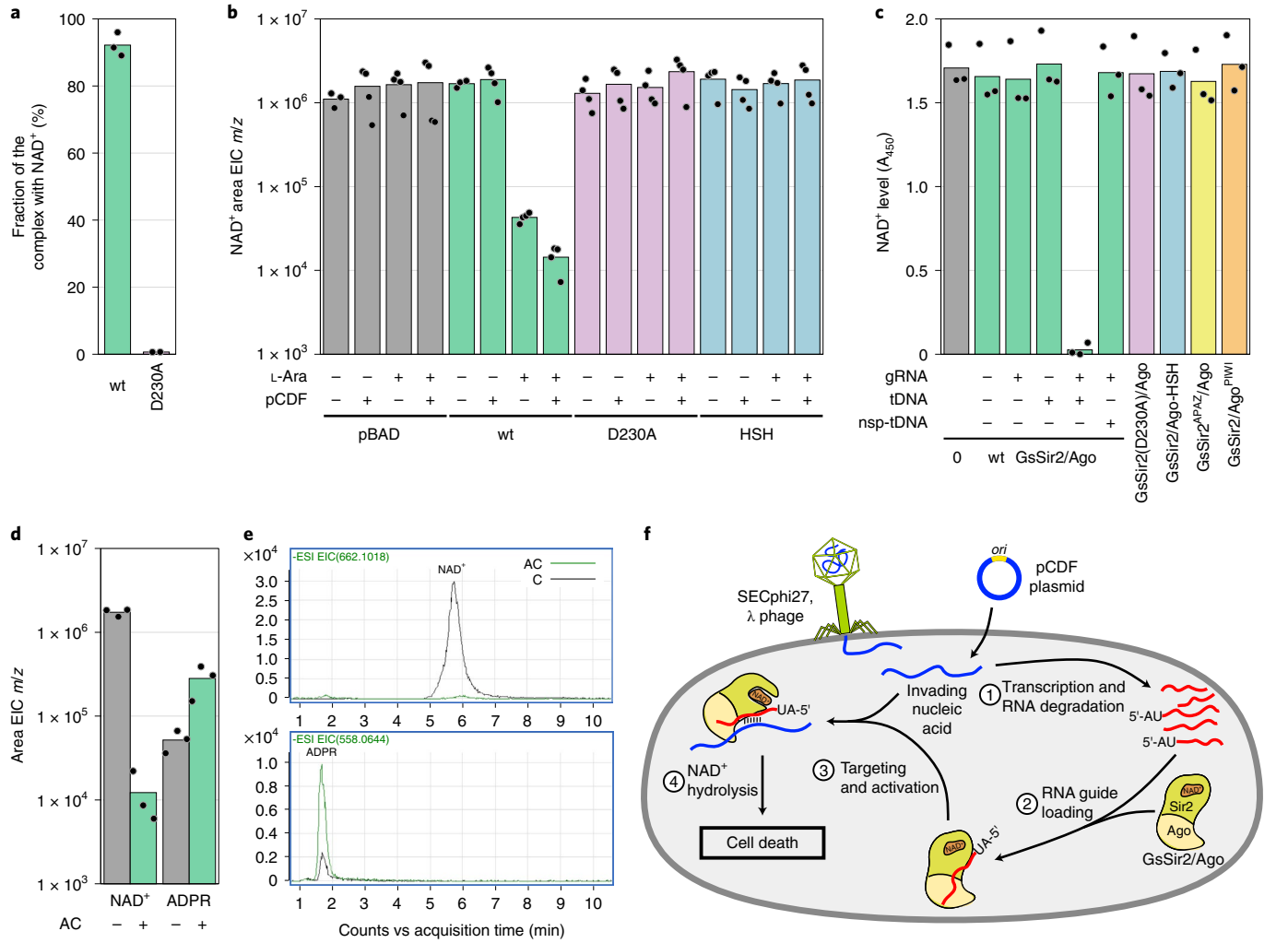


Fig. 5 | GsSir2/Ago binds and hydrolyses NAD$^+$. a, The D230A mutation within the Sir2 protein abolished NAD$^+$ binding. **b**, NAD$^+$ amounts in *E. coli* cells in the presence of the non-induced wt and mutant GsSir2/Ago systems, and in the presence and absence of the pCDF plasmids. NAD$^+$ amounts were estimated according to the extracted ion current (EIC) areas of NAD$^+$ (*m/z* 662.1018). pBAD, empty vector; wt, GsSir2/Ago; D230A, GsSir2(D230A)/Ago; HSH, GsSir2/Ago-HSH. **c**, NAD$^+$ depletion by GsSir2/Ago in vitro. The wt GsSir2/Ago or mutant complex (0.5 μM) was incubated with NAD$^+$ (50 μM) for 1 h at 37 °C (see experimental details in Materials and Methods). gRNA, 5′P-RNA guide; tDNA, target DNA complementary to the RNA guide; nsp-tDNA, ssDNA non-complementary to the RNA guide. **d**, NAD$^+$ hydrolysis by wt GsSir2/Ago in vitro. The binary GsSir2/Ago-gRNA complex (0.5 μM) was incubated with NAD$^+$ (50 μM) for 1 h at 37 °C in the presence of the complementary DNA target (0.5 μM). NAD$^+$ depletion and ADPR accumulation were analysed by MS according to the EIC of NAD$^+$ (*m/z* 662.1018) and ADPR (*m/z* 558.0644), respectively. AC, the activated wt GsSir2/Ago-gRNA/tDNA complex. **e**, Representative mass chromatograms of NAD$^+$ hydrolysis by wt GsSir2/Ago in vitro (as in **d**). C, a control sample without GsSir2/Ago. **f**, Putative model of GsSir2/Ago defence against mobile genetic elements. After lambda phage infection or pCDF plasmid transformation, GsSir2/Ago acquires small 5′-AU-RNAs originating from the invader transcripts (for example, from the pCDF *ori* region). The GsSir2/Ago binary complex, guided by small RNA, targets the invaders' complementary DNA, becomes activated and hydrolyses NAD$^+$ resulting in cell death.

The GsSir2/Ago complex binds NAD$^+$ and causes its depletion

Computational analysis of Sir2 domains showed that they possess a conserved NAD$^+$-binding pocket (Extended Data Figs. 1c and 5). To determine whether Sir2 domains can indeed bind endogenous NAD$^+$, purified GsSir2/Ago and CcSir2/Ago complexes were heat-treated, protein aggregates removed by centrifugation and the supernatant analysed by mass spectrometry–high-performance liquid chromatography (MS–HPLC) (Fig. 5 and Extended Data Fig. 7). The quantitative analysis showed that both the GsSir2/Ago and CcSir2/Ago complexes co-purified with bound endogenous NAD$^+$ in approximately 1:1 (Sir2:NAD$^+$) molar ratio (Fig. 5a, and Extended Data Figs. 3 and 7). However, in the case of the functionally inactive GsSir2(D230A)/Ago mutant, only 0.6% of all complexes were NAD$^+$-bound, indicating that the mutation severely compromised NAD$^+$ binding by the Sir2 domain (Fig. 5a). NAD$^+$ binding by the GsSir2 subunit and the similarity of GsSir2 to the N-terminal NADase of the ThsA protein from the anti-phage Thoeris system[24] prompted us to investigate the level of endogenous NAD$^+$ in the presence of the induced GsSir2/Ago system and its target pCDF plasmid. In these experiments, the corresponding *E. coli* cells were lysed, proteins were removed and the amount of NAD$^+$ in the supernatant was examined by MS–HPLC. When the wt GsSir2/Ago expression was induced in the presence of pCDF plasmid, NAD$^+$ was depleted (~120-fold decrease), whereas in the case of the functionally inactive GsSir2(D230A)/Ago and GsSir2/Ago-HSH mutants, the level of endogenous NAD$^+$ was similar to that of the empty pBAD vector (Fig. 5b and Extended Data Fig. 7). It should be noted that a significant ~30-fold decrease in the NAD$^+$ level was observed when the wt GsSir2/Ago expression was induced even in the absence of pCDF, suggesting that the heterologously expressed system may be toxic to the cells resulting in their slower growth (Fig. 5b and Extended Data Fig. 7). To test the hypothesis that the activated

GsSir2/Ago system, similar to the Thoeris anti-phage system, depletes the endogenous NAD^+ through hydrolysis or cyclization, we attempted to identify possible products (ADPR, cADPR, AMP, cAMP, ADP, cADP, nicotinamide and adenine) using MS–HPLC, albeit without success. It is possible that NAD^+ conversion products were not detected due to either being processed in the cell to other reaction intermediates or to ion suppression during MS analysis of the cell lysates.

Next, we investigated whether the NAD^+ depletion activity detected in cells could be reconstituted in vitro. To this aim we mixed the binary GsSir2/Ago-gRNA complex with ssDNA and monitored NAD^+ concentration using a commercial kit (Fig. 5c). We found that NAD^+ concentration decreased when wt GsSir2/Ago-gRNA complex was added to the complementary ssDNA target; however, no changes were observed in the case of non-complementary ssDNA. Mutations in the Sir2 domain (D230A) or the APAZ/Ago part (GsSir2/Ago-HSH, $GsSir2^{APAZ}$/Ago and GsSir2/Ago^{PIWI} mutants) compromised NAD^+ depletion (Fig. 5c). MS analysis revealed that the GsSir2/Ago-gRNA complex, in the presence of the complementary ssDNA, hydrolyses NAD^+ to ADPR (Fig. 5d,e), similar to the Thoeris anti-phage system[24]. Taken together, these results demonstrate that GsSir2/Ago functions as a NADase that becomes activated upon target DNA binding.

Discussion

The association of Sir2-like domains with short pAgo proteins has been identified bioinformatically in the pioneering Makarova et al. paper[18]. It has been speculated that Sir2-domain proteins can act as nucleases, but the structure and function of Sir2 proteins have so far not been elucidated. Here we show that the APAZ-containing Sir2 and short pAgo proteins form a heterodimeric complex (Fig. 3), similar to a short pAgo and a Mrr nuclease domain-containing protein[25]. Furthermore, our structure modelling results show that the APAZ region of Sir2 proteins shares similarity with the N, L1 and L2 domains of canonical Agos, substantiating previous sequence-based predictions[11,19] At the same time, Sir2 proteins entirely lack the PAZ domain (Fig. 1a and Extended Data Fig. 5a). Thus, apparently both split and single-chain Sir2/Ago systems evolved from long pAgos by the loss of the PAZ domain and acquisition of the Sir2 domain.

Next, we provide the experimental evidence that the Sir2/Ago complex functions as a defence system against invading phages and plasmids (Figs. 1 and 2). Intriguingly, plasmid interference assay using four plasmids with different replicons (Extended Data Fig. 2) revealed that the GsSir2/Ago system prevents transformation only of the pCDF plasmid that contains CloDF13 *ori* (Fig. 2), suggesting that GsSir2/Ago may recognize specific replicon elements or structures. Indeed, *ori* swap between pCDF and permissive pCOLA plasmid made the latter sensitive to GsSir2/Ago interference. We show here that GsSir2/Ago co-purifies from *E. coli* cells together with small (predominantly 21 nt long) RNAs that preferentially contain the 5′-AU dinucleotide (Fig. 4b–d). Interestingly, a fraction of small RNAs that originates from pCDF CloDF13 and pBAD ColE1 *ori* regions is enriched (Fig. 4e). ColE1-like origins, including CloDF13, use two small RNAs (RNAI and RNAII) for priming of the replication that involves the R-loop intermediate[26–28]. It is tempting to speculate that GsSir2/Ago is able to bind nucleic acids of different lengths (Fig. 4b,c) and preferentially binds *ori*-associated small RNAs that can be subjected to further processing by cellular RNases to produce ~21 nt gRNAs similarly to long RsAgo that shares an inactivated PIWI domain with GsSir2/Ago[15].

We further show that in vitro, the reconstituted wt GsSir2/Ago-gRNA complex becomes activated after binding the complementary DNA target and triggers NAD^+ hydrolysis generating ADPR (Fig. 5c–e). It is likely that in *E. coli* cells, the APAZ/Ago part of the GsSir2/Ago complex guided by the *ori*-associated RNA guides could bind to the complementary plasmid or phage DNA target, activating the Sir2 effector domain that depletes endogenous NAD^+ leading to cell death, thereby restricting plasmid and phage propagation (Fig. 5f, Supplementary Text and ref. [29]). A similar anti-phage defence mechanism based on NAD^+ exhaustion has been shown for the Thoeris and the Pycsar systems, CBASS (cyclic oligonucleotide-based signalling system) and DSR (defence-associated sirtuins)[29–31]. In the Thoeris system of *Bacillus cereus* MSX-D12, the Sir2 domain is similar to that of the GsSir2/Ago system and performs the hydrolysis of NAD^+ to ADPR and nicotinamide[31], similar to the GsSir2/Ago system. Further structural and biochemical studies are underway to establish the structure of the heterodimeric Sir2/Ago complex and the mechanism of Sir2 domain activation that triggers NAD^+ hydrolysis.

Methods

Oligonucleotides used in this work

All synthetic DNA oligonucleotides used for cloning and site-specific mutagenesis were purchased from Metabion and are listed in Supplementary Table 2.

Cloning and mutagenesis

A whole operon of the GsSir2/Ago system, composed of the Sir2 (GSU1360, NP_952413.1) and Ago (GSU1361, NP_952414.1) encoding genes, was amplified by PCR from the genomic DNA of *G. sulfurreducens* Caccavo (ATCC No. 51573, LGC Standards, 51573D-5) using the oligonucleotides MZ-239 and MZ-240 (Supplementary Table 2), respectively. The resulting DNA fragment was digested by Eco31I (ThermoFisher, FD0293) and XhoI (ThermoFisher, FD0694), and cloned into pBAD/HisA expression vector (ThermoFisher, V43001) pre-cleaved with NheI (ThermoFisher, FD0973) and XhoI, and dephosphorylated using FastAP (ThermoFisher, EF0651). The D230A mutant of the GsSir2 protein was produced by QuikChange site-directed mutagenesis[32] using respective mutagenic oligonucleotides (Supplementary Table 2). To generate the GsAgo protein containing a bulky His_6-StrepII-His_6-tag (HSH-tag, 29 amino acids: LEGHHHHHHSSWSHPQFEKGVEGHHHHHH), a whole operon of the GsSir2/Ago system was amplified by PCR from the genomic DNA using the oligonucleotides MZ-325 and MZ-326 (Supplementary Table 2), respectively. The resulting DNA fragment was digested by Eco31I and XhoI cloned into pBAD24 vector through NcoI/XhoI sites to generate pBAD24-HSH expression vector. The GsSir2/Ago mutants ($GsSir2^{APAZ}$/Ago, GsSir2/Ago^{MID} and GsSir2/Ago^{PIWI}) of the putative surface of the interaction with nucleic acids were designed on the basis of the GsSir2/Ago structural model (see Supplementary Methods). We changed positively charged residues that are structurally equivalent to the RsAgo residues involved in the interaction with the guide and the target to negatively charged residues in the corresponding domains (APAZ: R440E/R442E/R517E/R525E/R528E/R531E, MID: K155E/H166E/K170E/K207E, PIWI: R269E/K270E/H310E/K441E). The surface mutants were obtained using whole-gene synthesis and cloning service provided by Twist Bioscience.

A PgSir2-Ago gene (BgramDRAFT_6510, WP_052303232.1) was amplified by PCR from the genomic DNA of *P. graminis* C4D1M (ATCC No. 700544, LGC Standards) using the oligonucleotides MZ-915 and MZ-916 (Supplementary Table 2), respectively. The resulting DNA fragment was digested by BveI (ThermoFisher, FD1744) and HindIII (ThermoFisher, FD0504), ligated into pBAD24 expression vector pre-cleaved with NcoI and HindIII, and dephosphorylated using FastAP.

The *E. coli* codon optimized genes (IDT codon optimization tool) encoding the Sir2 (WP_053571900.1) and Ago (WP_053571899.1) of the CcSir2/Ago system (from *C. cordobensis*, NCBI taxon_id 1353886) were synthesized and cloned into pBAD/HisA expression vector by Twist Bioscience. The CcSir2 protein contains at its N terminus a His_6-tag that can be cleaved by TEV protease. For purification of the CcSir2/Ago complex, a TwinStrep-tag (35 amino acids: MGGSAWSHPQFEKGGGSGGGSGGSAWSHPQFEKGS) was additionally fused to the N terminus of the CcSir2 protein already containing a His_6-tag.

To swap *ori* regions between pCOLA and pCDF plasmids, the DNA fragments containing ColA and CloDF13 *ori* were amplified by PCR

using the oligonucleotides MZ-1217/MZ-1218 and MZ-1230/MZ-123, respectively (Supplementary Table 2). The resulting DNA fragments were digested by NheI and XbaI (ThermoFisher, FD0684), ligated into pCOLA and pCDF vectors pre-cleaved with NheI and XbaI, and dephosphorylated using FastAP.

To swap streptomycin resistance for kanamycin in the pCDF plasmid, plasmids pCDF and pCOLA were cleaved with NheI and Eco81I (ThermoFisher, FD0374), and isolated using a runView electrophoresis system (Cleaver Scientific). The purified fragments were then ligated into pCDF to yield a pCDF_Kn plasmid.

All gene sequences were confirmed by sequencing; links to DNA and protein sequences are presented in Supplementary Table 1.

Phage restriction assay

E. coli MG1655 (ATCC 47076) cells carrying pBAD plasmids expressing wt or mutated Sir2/Ago systems were used for phage infection assays as described below. Whole-plasmid sequencing was applied to all transformed *E. coli* clones to verify the integrity of the system and lack of mutations, as previously described[33].

E. coli phages (T4, T7, lambda-vir) were kindly provided by U. Qimron. Phages SECphi17, SECphi18 and SECphi27 were isolated by the Sorek lab as previously described[34]. Small-drop plaque assay was performed as previously described[35]. An overnight culture of *E. coli* bacteria was diluted 1:100 in MMB medium (LB + 0.1 mM $MnCl_2$ + 5 mM $MgCl_2$ + 5 mM $CaCl_2$) supplied with 0.1% L-arabinose for expression induction. Bacterial cultures were incubated at 37 °C until early log phase (optical density $(OD)_{600}$ = 0.3), and 500 µl of bacteria were mixed with 25 ml of MMB agar (LB + 0.1 mM $MnCl_2$ + 5 mM $MgCl_2$ + 5 mM $CaCl_2$ + 0.5% agar + 0.1% L-arabinose) and poured into a square Petri dish. Serial dilutions of phage lysate in MMB were dropped on top of the cell lawn. After the drops dried up, plates were incubated at room temperature for 24 h. EOP was determined via comparison of phage lysate titre on control bacteria and bacteria containing the Argonaute system with and without induction with L-arabinose.

Liquid culture phage infection experiments were performed as previously described[33]. After overnight incubation, the liquid suspensions of pAgo-lacking and pAgo-containing *E. coli* cells were diluted 1:100 in MMB medium supplied with 0.2% L-arabinose and dispensed (180 µl volume) into a 96-well plate. Plates were incubated at 37 °C until the early log phase (OD_{600} = 0.3), then 20 µl of phage lysate was added to each well at multiplicities of infection of 5, 0.5 or 0.05, and each experiment was performed in three replicates. Optical density measurements at a wavelength of 600 nm were taken every 15 min using a TECAN Infinite 200 plate reader.

Plasmid interference assay

E. coli (BL21-AI and DH10B strain) cells were pre-transformed with pBAD/HisA plasmid encoding a GsSir2/Ago (either wild-type or mutant variants), CcSir2/Ago or PgSir2-Ago under the control of araBAD promoter. After 2 h of induction with either 0.01% (w/v) (CcSir2/Ago) or 0.1% (GsSir2/Ago and PgSir2-Ago) L-arabinose at 37 °C and 200 r.p.m., cells were heat-shock transformed with pCDF (pCDF_Kn in the case of PgSir2-Ago in DH10B), pCOLA, pACYC184 or pRSF plasmids. After recovery, cells were either serially diluted and aliquots spotted on selection medium, or undiluted suspensions were spread on selection medium and colony-forming units counted manually. In parallel, viability and overexpression of pAgo-containing *E. coli* cells were monitored using serial dilutions and western blot, respectively.

In separate experiments, *E. coli* BL21-AI strain cells were heat-shock transformed with both GsSir2/Ago operon-encoding pBAD/HisA construct and pCDF plasmid. Protein expression in selected double transformants was then induced by the addition of L-Ara (final concentration of 0.1% (w/v)) into liquid LB culture. After a 2 h induction at 37 °C and 200 r.p.m., OD_{600} equalized, cultures were serially diluted and aliquots were spotted on a selection medium containing different antibiotics.

Expression and purification of GsSir2/Ago complexes

For GsSir2/Ago protein expression, *E. coli* DH10B strain was transformed with a corresponding plasmid (Supplementary Table 1). Cells were grown at 37 °C in LB medium in the presence of 50 µg ml^{-1} ampicillin until OD_{600} = 0.7 was reached. Then, the temperature was decreased to 16 °C and proteins were expressed for 16 h by adding 0.2% w/v L-arabinose. Next, collected cells were disrupted by sonication in buffer A (20 mM Tris-HCl (pH 8.0 at 25 °C), 500 mM NaCl, 2 mM phenylmethylsulfonyl fluoride, 5 mM 2-mercaptoethanol), and cell debris was removed by centrifugation. GsSir2/Ago complexes were purified to >90% homogeneity by chromatography through HisTrap HP chelating, HiTrap Heparin HP and HiLoad Superdex 200 columns (GE Healthcare). Purified proteins were stored at −20 °C in a buffer containing 20 mM Tris-HCl (pH 8.0 at 25 °C), 200 mM KCl, 1 mM dithiothreitol (DTT) and 50% v/v glycerol. The identity of the purified proteins was confirmed by mass spectrometry. Protein concentrations were determined from OD_{280} measurements using the theoretical extinction coefficients calculated with the ProtParam tool available at http://web.expasy.org/protparam/. GsSir2/Ago complex concentrations are expressed in terms of heterodimer. The GsSir2/Ago^{MID} surface mutant could not be purified due to its poor expression.

Antibodies used in this work

For western blot analysis of target proteins, the following antibodies were used: 6x-His-Tag monoclonal antibody (ThermoFisher, MA1-21315, RRID AB_557403, Lot no. WE323793, clone HIS.H8); goat anti-mouse IgG (H + L) secondary antibody, AP conjugated (ThermoFisher, 31320, RRID AB_228304, Lot no. VH311913, polyclonal).

SEC-MALS and mass photometry

Size-exclusion chromatography of GsSir2/Ago complexes was carried out at room temperature using a Superdex 200 10/300 GL column (GE Healthcare) pre-equilibrated with a buffer (20 mM Tris-HCl (pH 8.0 at 25 °C), 500 mM NaCl). A calibration curve was generated by measuring the elution volumes of a series of standard proteins of known molecular mass (Bio-Rad). The molecular masses of pAgos complexes were calculated by interpolating their elution volume onto the calibration curve. SEC-MALS of GsSir2/Ago and CcSir2/Ago complexes was performed at room temperature using a Superdex 200 10/300 GL column (GE Healthcare) pre-equilibrated with a buffer (20 mM Tris-HCl (pH 8.0 at 25 °C), 500 mM NaCl, 0.03% NaN_3, 1 mM DTT), at 0.4 ml min^{-1} flow rate. Sample concentrations were 6 µM and 6.5 µM for GsSir2/Ago and CcSir2/Ago, respectively. Light scattering signals were monitored on a miniDawn TREOS II detector, and concentrations of protein samples were measured using an Optilab T-rEX refractive index detector (Wyatt Technologies). Data were analysed in Astra software (Wyatt Technologies) using a specific refractive index increment (dn/dc) value of 0.185 ml g^{-1}. Mass photometry of the GsSir2/Ago complex was performed using a Refeyn OneMP system (Refeyn). The protein complex was diluted to 20 nM in a buffer containing 20 mM Tris-HCl (pH 8.0) and 500 mM NaCl before measurement.

SAXS analysis

The synchrotron SAXS data were collected at beamline P12 operated by EMBL Hamburg at the PETRA III storage ring (DESY, Hamburg, Germany)[36]. The GsSir2/Ago sample in the storage buffer was transferred into the sample buffer (20 mM Tris-HCl (pH 7.5), 200 mM NaCl, 5 mM $MgCl_2$, 2 mM β-mercaptoethanol) using gel-filtration NAP column (GE Healthcare) and concentrated by ultrafiltration to 1.2, 1.3, 1.6 and 5.5 mg ml^{-1} concentrations. The data were collected at 0.124 nm wavelength and the distance to the detector (Pilatus 2 M, Dectris) was set to 3 m. Samples in the sample changer were kept at 10 °C and capillary temperature was set to 20 °C. Twenty frames exposed for 0.045 s were averaged for each concentration. The *s*-range of collected

data was from 0.0133796 to 3.7925 nm^{-1}, where $s = 4\pi\sin\theta/\lambda$ is the momentum transfer (2θ is the scattering angle, λ is the wavelength of the X-rays). The data were analysed using programmes of the ATSAS 2.8.4 (r10552) suite[37]. Data were normalized to an absolute scale with water as standard. As the data collected for the sample with a concentration of 1.22 mg ml^{-1} were noisy at higher s (Extended Data Fig. 4), and higher-concentration data showed more aggregation at low s, we used a merged dataset produced with PRIMUS[38]. Scattering data were parameterized and indirectly Fourier transformed with GNOM5 (ref. [39]). Structural parameters of this dataset are summarized in Supplementary Table 3. The dimensionless Kratky plot in Extended Data Fig. 4 was calculated as described previously[40]. The ab initio models were calculated by GASBOR[41] software. Molecular mass estimations of the apo-GsSir2/Ago complex in solution, assessed by ATSAS tools (DATVC, DATMW) and server SAXSMoW (http://saxs.ifsc.usp.br/)[42], are presented in Supplementary Table 3.

Nucleic acid binding assay

The oligonucleotide substrates (Supplementary Table 2) were 5′-labelled with [γ-^{32}P]ATP (PerkinElmer) and T4 polynucleotidyl kinase (PNK) (ThermoFisher, EK0031). The 3′-labelled substrate was prepared with [α-^{32}P]cordycepin-5′-triphosphate (Hartmann Analytics) and terminal deoxynucleotidyl transferase. An aliquot of the 3′-labelled substrate was subsequently phosphorylated with cold ATP and T4 PNK to obtain a 3′[α-^{32}P]cordycepin-labelled oligonucleotide containing 5′-phosphate. Annealing was performed in the PNK reaction buffer supplemented with 50 mM EDTA at 2 μM total single-stranded oligonucleotide concentration. Circular ssDNA substrate was prepared by circularization of 5′-^{32}P-labelled TK-49 using CircLigase II (Lucigen, CL9021K) according to the manufacturer's recommendations, purified from a denaturing PAA gel (21% 29:1 acrylamide/bis-acrylamide in Tris-borate-EDTA buffer supplemented with 8 M urea) by phenol-chloroform extraction, precipitated in 96% ethanol with 0.45 M sodium acetate, washed with 75% ethanol, and resuspended in water.

For EMSA experiments, appropriate substrates and proteins were pre-diluted to two times the final binding reaction concentration in 40 mM Tris-acetate (pH 8.3 at 23 °C) and 1 mM EDTA (TAE buffer, Invitrogen, 24710-030), supplemented with 5 mM magnesium acetate, 0.1 mg ml^{-1} BSA, 1 mM DTT and 10% glycerol. The binding reactions were conducted by mixing equal volumes of enzyme and radiolabelled substrate. In all cases, final binding reactions contained 0.1 nM of radiolabelled substrate at 0–2 nM (0, 0.02, 0.05, 0.1, 0.2, 0.5, 1, 2) or 0–500 nM (0, 5, 10, 20, 50, 100, 200, 500) of GsSir2/Ago and CcSir2/Ago complexes. Three independent replicates were performed. For clarity, only EMSA gels obtained using low protein concentrations (0–2 nM) are shown, while for calculations of high dissociation constant (K_d) values, EMSA data obtained using high protein concentrations (0–500 nM) were used.

Binding experiments for the binary GsSir2/Ago complex were conducted by first pre-mixing 5′-phosphorylated ssRNA or ssDNA guide with the equimolar GsSir2/Ago complex in the same buffer as above. The binary GsSir2/Ago:NA guide complex was then diluted to two times the final reaction concentration (with respect to the guide) in the same buffer and mixed with a complementary 5′-^{32}P-target oligonucleotide in the presence or absence of 67 ng μl^{-1} heparin sodium salt (Sigma-Aldrich, H3149). The final reaction contained 10 pM target NA and 0, 0.02, 0.05, 0.1, 0.2, 0.5 and 1 nM of GsSir2/Ago:NA guide complex. A control (C_g*) contained 0.1 nM GsSir2/Ago:NA guide complex with the guide labelled with [γ-^{32}P]ATP, and 10 pM unlabelled target NA. Three independent replicates were performed.

The binding reaction mixtures were analysed by EMSA in a PAA gel (8% 29:1 acrylamide/bis-acrylamide in TAE). The electrophoresis TAE buffer was supplemented with 5 mM magnesium acetate. Radiolabelled substrates were detected and quantified using a phosphor imager. The results were analysed with OptiQuant and OriginPro software. The K_d was calculated from the following formula:

$$S_{NB} = A1 + \frac{\frac{100}{S_0}\left(S_0\frac{100-A1}{100} - E_0 - K_d + \sqrt{\left(S_0\frac{100-A1}{100} + E_0 + K_d\right)^2 - 4S_0E_0\frac{100-A1}{100}}\right)}{2}$$

where S_{NB} is the unbound substrate (nM), S_0 is the initial substrate concentration (nM), E_0 is the initial protein complex concentration (nM), and A1 is the non-binding fraction of substrate (%).

Nucleic acid extraction and analysis

To obtain GsSir2/Ago-bound nucleic acids, *E. coli* DH10B was transformed with pBAD/HisA_TwinStrep_TEV_GsSir2/Ago and pCDF_Kn plasmids (Supplementary Table 1). Cells were grown at 37 °C in LB medium in the presence of 50 μg ml^{-1} ampicillin and 25 μg ml^{-1} kanamycin until OD_{600} = 0.7 was reached. Then, expression was induced by adding 0.1% w/v L-arabinose, and cells were collected after 2 h. Cells were disrupted using B-PER bacterial protein extraction reagent (ThermoFisher, 78248) containing 6 mg ml^{-1} lysozyme. The GsSir2/Ago-NA complex was purified as described above, except that all buffer solutions contained 100 mM NaCl.

To extract nucleic acids co-purified with the GsSir2/Ago complex, 800 μl of Roti-phenol/chloroform/isoamyl alcohol (Carl-Roth, A156) was added to 800 μl of purified protein-NA fractions in 5PRIME phase lock gel tubes (Quantabio, 733-2477). The upper aqueous phase was isolated and 0.1 volume of 1 M sodium acetate, 3 volumes of 100% ethanol and 10 μl glycogen (ThermoFisher, R0561) were added. This mixture was vortexed briefly and incubated at −20 °C for 20 h. Samples were centrifuged for 20 min and the supernatant was removed from the pellet. The pellet was washed with cold (−20 °C) 70% ethanol. The pellets containing the co-purified nucleic acids were dried for 20 min at room temperature, and pellets were resuspended in 30 μl water (free of nucleases).

Co-purified nucleic acids were dephosphorylated with FastAP thermosensitive alkaline phosphatase (ThermoFisher, EF0651) and [γ-^{32}P]-ATP (PerkinElmer) labelled with T4 PNK (ThermoFisher, EK0031). Labelled nucleic acids were incubated with nucleases (ThermoFisher DNase I, 18047019, RNase A/T1, EN0551) for 30 min at 37 °C. After nuclease treatment, samples were mixed with RNA gel loading dye (ThermoFisher, R0641), heated for 5 min at 95 °C and resolved on 20% denaturing (8 M urea) polyacrylamide gels. The molecular weight marker used for RNA size identification was Decade Marker System (Ambion, AM7778) and 22-nt-long RNA oligonucleotide. Radioactivity was captured from gels using phosphor screens and imaged using a Typhoon FLA 7000 laser-scanner (GE Healthcare).

In a control sample, total RNA from induced cells was extracted using SPLIT RNA extraction kit (Lexogen, 008). Then ribosomal RNA was removed using RiboCop for Gram-negative bacteria (Lexogen, 126).

RNA sequencing and analysis

Half of extracted RNA was treated with T4 PNK (ThermoFisher, EK0031) according to the protocol of the manufacturer. Then T4 PNK treated and untreated RNA samples were converted to DNA libraries using the Small RNA-seq Library Prep kit (Lexogen, 052). Concentration and quality of libraries were measured with a Qubit fluorometer (ThermoFisher) and a 2100 Bioanalyzer (Agilent).

Both libraries were sequenced using Illumina MiniSeq sequencing with single-end reads and 75 bp read length. Single-end reads were processed by trimming adapters with AdapterRemoval v2.3.0 (ref. [43]). Then the processed reads were aligned to the *E. coli* strain K12 substrain DH10B genome (GenBank: CP000948.1) and the additional pBAD/HisA_TwinStrep_TEV_GsSir2/Ago, pCDF_Kn plasmids (Supplementary Table 1) using BWA-MEM v0.7.17 (ref. [44]). To avoid filtering out shorter reads during the alignment process, aligned reads with mapping quality values greater than or equal to 15 were chosen. FastQC

v0.11.8 (ref. [45]) was used for read quality control and SAMtools v1.7 (ref. [46]) for indexing, sorting and analysing alignment files. A custom script (fragmentation-bias.jl) in combination with Weblogo v3.7.4 (ref. [47]) was used to produce nucleotide frequency plots. The custom script had to be implemented to ensure that only aligned reads would be used for nucleotide frequency analysis. Gene enrichment analysis was performed with bedtools v2.26.0 (ref. [48]) and FPKM_count.py v4.0.0 of the RSeqQC package[49]. IGV v2.5.2 (ref. [50]) was mainly used to inspect and visualize read coverage along the genomes. A control DNA library of total RNA was prepared using CORALL Total RNA-seq Library Prep kit (Lexogen, 095). Concentration and quality of the library were measured with a Qubit Fluorometer (ThermoFisher) and a 2100 Bioanalyzer (Agilent) according to the protocol of the manufacturer.

The control DNA library was sequenced using Illumina NextSeq sequencing with paired-end reads and 75 bp read length. Read processing, alignment and alignment analysis were similar to those for Illumina MiniSeq sequencing.

Preparation of *E. coli* cells for NAD^+ quantification

Overnight cultures of single colonies of *E. coli* DH10B strain harbouring a pBAD-His construct with either wt GsSir2/Ago or mutant system (GsSir2/Ago-HSH or GsSir2(D230A)/Ago), or empty vector (negative control) were diluted and grown in LB broth (BD) supplemented with respective antibiotics (50 μg ml^{-1} ampicillin and 25 μg ml^{-1} streptomycin) at 37 °C until they reached OD_{600} = 0.4–0.5. Cell cultures were either induced to express the protein or not (control samples). L-Ara (0.1% final concentration) was added to induce protein expression. Induced and non-induced cultures were collected 2 h later. The cultures were normalized to OD_{600} of approximately 0.7 and the pellet from 1 ml of culture suspension was stored at −80 °C until further analysis. All cell pellets were lysed by adding B-PER solution (ThermoFisher) supplemented with 6 mg ml^{-1} lysozyme (62971, Fluka) for 20 min at room temperature while gently rocking (Multi Bio 3D Mini-Shaker, Biosan). Cell debris was removed by centrifugation and metabolites were isolated by phenol:chloroform:isoamyl alcohol (25:24:1, v/v/v) extraction. Metabolites were stored at −20 °C until MS–HPLC analysis. Additionally, the endogenous NAD^+ concentration was estimated using NAD/NADH quantitation kit (Sigma Aldrich, MAK037) from four independent measurements.

In vitro NADase assay

Reaction mixtures with a volume of 25 μl were prepared with the following final concentrations: 0.5 μM GsSir2/Ago or mutant complex, 50 μM NAD^+, 1× Tango buffer (33 mM Tris-acetate (pH 7.9 at 37 °C), 10 mM magnesium acetate, 66 mM potassium acetate, 0.1 mg ml^{-1} BSA; ThermoFisher, BY5), 1 mM DTT, 0.5 μM 5′P-RNA guide (TF-A) and/or 0.5 μM ssDNA (MZ-949 or MZ-589) (Supplementary Table 2). Reactions with RNA guide were pre-incubated for 15 min at 37 °C, then ssDNA was added and the mixture incubated for 1 h at 37 °C. A volume of 3 μl of each sample was used as input for the NAD/NADH quantitation kit (Sigma Aldrich, MAK037) according to the instructions provided by the manufacturer. All experiments were performed in triplicate. These samples were also used for mass spectrometry.

Mass spectrometry of NAD^+

To quantitate NAD^+ bound to GsSir2/Ago and CcSir2/Ago complexes, high-performance liquid chromatography-mass spectrometry/mass spectrometry (HPLC–MS/MS) analysis was used. First, purified pAgos complexes were diluted to 5 μM in a buffer containing 20 mM Tris-HCl (pH 8.0 at 25 °C) and 200 mM NaCl. Then 20 μl of the solution was incubated at 70 °C for 20 min and centrifuged for 30 min (16,100 *g* at 4 °C) to remove unfolded proteins. The supernatants and NAD^+ standards were analysed by electrospray ionization mass spectrometry (ESI–MS) using an integrated HPLC/ESI–MS system (1290 Infinity, Agilent Technologies/Triple Quadrupole 6410, Agilent Technologies) equipped with a Supelco Discovery HS C18 column (7.5 cm × 2.1 mm, 3 μm; Agilent Technologies). HPLC/ESI–MS/MS was performed using two ion transitions to detect NAD^+ in the samples: 662.1→540.1 and 662.1→426.0. Ion transition 662.1→540.1, being the most abundant, was used for the quantitative analysis. Mobile phase A was 5 mM ammonium acetate in water (pH 7.0) and mobile phase B was 5 mM ammonium acetate in methanol (pH 7.0). The HPLC parameters were as follows: flow 0.25 ml min^{-1}; column temperature 30 °C; 0–3 min, 0% B; 3–9 min, 0–40% B; 9–10 min, 40–100% B; 10–13 min, 100% B. The MS was operated using negative electrospray ionization at 2,500 V, the gas temperature was set to 300 °C and the fragmentor voltage was 135 V. Multiple reaction monitoring was used with a collision energy of 15 V to measure ion *m/z* 540.1 (ion transition 662.1 → 540.1) and also with a collision energy of 20 V to measure ion *m/z* 426.0 (ion transition 662.1 → 426.0).

To quantitate endogenous NAD^+, HPLC–MS analysis was performed by ESI–MS using an integrated HPLC/ESI–MS system (1290 Infinity, Agilent Technologies/Q-TOF 6520, Agilent Technologies) equipped with a Supelco Discovery HS C18 column (7.5 cm × 2.1 mm, 3 μm; Agilent Technologies). The samples were investigated in both negative and positive ionization modes. For negative ionization mode, solvents A (5 mM ammonium acetate in water, pH 7.0) and B (5 mM ammonium acetate in methanol, pH 7.0) were used. For positive ionization mode, solvents C (0.02% formic acid in water) and D (0.02% formic acid in acetonitrile) were used. In both cases, elution was performed with a linear gradient of solvents at a flow rate of 0.3 ml min^{-1} at 30 °C as follows: 0–5 min, 0% B; 5–18 min, 20% B; 18–22 min, 100% B, 22–27 min 100% B. Ionization capillary voltage was set to 2,500 V and fragmentor to 150 V. A list of compounds that could be expected to be products of NAD^+ hydrolysis and relative *m/z* values are as follows: ADPR [M-H]− *m/z* = 558.0644, cADPR [M-H]− *m/z* = 540.0538, AMP [M-H]− *m/z* = 346.0558, cAMP [M-H]− *m/z* = 328.0452, ADP [M-H]− *m/z* = 426.0221, cADP [M-H]− *m/z* = 408.0116, nicotinamide [M+H]+ *m/z* = 123.0553, adenine [M+H]+ *m/z* = 136.0618. Only traces of AMP and ADP were detected in all samples; other products from the list were absent.

NAD^+ hydrolysis products generated by GsSir2/Ago in vitro were analysed as above. Using negative ionization mode only accumulation of ADPR was detected.

Notes

During the revision of this manuscript, a paper was published that shows plasmid-induced degradation of NAD^+ in vivo by short pAgo-associated TIR-APAZ systems (named SPARTA)[51]. The paper also shows that expression of Sir2/Ago systems (named SPARSA) in *E. coli* triggers NAD depletion.

Reporting summary

Further information on research design is available in the Nature Research Reporting Summary linked to this article.

Data availability

All data are available in the paper and the supplementary material. In addition, small and total RNA sequencing data are available on the NCBI Sequence Read Archive under BioProject ID PRJNA851009. SAXS data are available in the Small Angle Scattering Biological Data Bank SASBDB under SASBDB ID SASDNH2: https://www.sasbdb.org/data/SASDNH2/. Plasmid sequences used in this work are available at https://www.benchling.com, with exact links for each plasmid provided in Supplementary Table 1.

Geobacter sulfurreducens, *Caballeronia cordobensis* and *Paraburkholderia graminis* genomes (respective GenBank accessions: GCA_000210155.1, GCA_001544575.2 and GCA_000172415.1) and all associated sequence and annotation data were obtained from NCBI (ftp://ftp.ncbi.nlm.nih.gov/genomes/Bacteria/). Searches through Pfam (http://pfam.xfam.org/), SwissProt (https://www.expasy.org/

resources/uniprotkb-swiss-prot) and PDB (https://www.rcsb.org/) databases were performed. PDB structures mentioned in this study: 5AWH, 4N41, 5UX0, 6LHX, 2H4F.

Code availability

The Julia script used to identify nucleotide frequency at the beginning of the aligned reads and prepare the input for the Weblogo programme is available in the GitHub repository at https://github.com/agrybauskas/argonaute-bound-rna-manuscript.

References

1. Kuhn, C. D. & Joshua-Tor, L. Eukaryotic Argonautes come into focus. *Trends Biochem. Sci.* **38**, 263–271 (2013).
2. Pratt, A. J. & MacRae, I. J. The RNA-induced silencing complex: a versatile gene-silencing machine. *J. Biol. Chem.* **284**, 17897–17901 (2009).
3. Sheu-Gruttadauria, J. & MacRae, I. J. Structural foundations of RNA silencing by argonaute. *J. Mol. Biol.* **429**, 2619–2639 (2017).
4. Olina, A. V., Kulbachinskiy, A. V., Aravin, A. A. & Esyunina, D. M. Argonaute proteins and mechanisms of RNA interference in eukaryotes and prokaryotes. *Biochemistry* **83**, 483–497 (2018).
5. Hutvagner, G. & Simard, M. J. Argonaute proteins: key players in RNA silencing. *Nat. Rev. Mol. Cell Biol.* **9**, 22–32 (2008).
6. Swarts, D. C. et al. The evolutionary journey of Argonaute proteins. *Nat. Struct. Mol. Biol.* **21**, 743–753 (2014).
7. Kwak, P. B. & Tomari, Y. The N domain of Argonaute drives duplex unwinding during RISC assembly. *Nat. Struct. Mol. Biol.* **19**, 145–151 (2012).
8. Ryazansky, S., Kulbachinskiy, A. & Aravin, A. A. The expanded universe of prokaryotic argonaute proteins. *mBio* **9**, e01935-18 (2018).
9. Hegge, J. W., Swarts, D. C. & Van Der Oost, J. Prokaryotic argonaute proteins: novel genome-editing tools? *Nat. Rev. Microbiol.* **16**, 5–11 (2018).
10. Lisitskaya, L., Aravin, A. A. & Kulbachinskiy, A. DNA interference and beyond: structure and functions of prokaryotic Argonaute proteins. *Nat. Commun.* **9**, 5165 (2018).
11. Willkomm, S., Makarova, K. S. & Grohmann, D. DNA silencing by prokaryotic Argonaute proteins adds a new layer of defense against invading nucleic acids. *FEMS Microbiol. Rev.* **42**, 376–387 (2018).
12. Kuzmenko, A. et al. DNA targeting and interference by a bacterial Argonaute nuclease. *Nature* https://doi.org/10.1038/s41586-020-2605-1 (2020).
13. Swarts, D. C. et al. DNA-guided DNA interference by a prokaryotic Argonaute. *Nature* **507**, 258–261 (2014).
14. Jolly, S. M. et al. *Thermus thermophilus* Argonaute functions in the completion of DNA replication. *Cell* **182**, 1545–1559.e18 (2020).
15. Olovnikov, I., Chan, K., Sachidanandam, R., Newman, D. K. & Aravin, A. A. Bacterial Argonaute samples the transcriptome to identify foreign DNA. *Mol. Cell* **51**, 594–605 (2013).
16. Liu, Y. et al. A programmable omnipotent Argonaute nuclease from mesophilic bacteria *Kurthia massiliensis*. *Nucleic Acids Res.* https://doi.org/10.1093/nar/gkaa1278 (2021).
17. Kropocheva, E., Kuzmenko, A., Aravin, A. A., Esyunina, D. & Kulbachinskiy, A. A programmable pAgo nuclease with universal guide and target specificity from the mesophilic bacterium *Kurthia massiliensis*. *Nucleic Acids Res.* https://doi.org/10.1093/nar/gkab182 (2021).
18. Makarova, K. S., Wolf, Y. I., van der Oost, J. & Koonin, E. V. Prokaryotic homologs of Argonaute proteins are predicted to function as key components of a novel system of defense against mobile genetic elements. *Biol. Direct* **4**, 29 (2009).
19. Burroughs, A. M., Ando, Y. & Aravind, L. 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 (2014).
20. North, B. J. & Verdin, E. Sirtuins: Sir2-related NAD-dependent protein deacetylases. *Genome Biol.* **5**, 224 (2004).
21. Gallego-Jara, J. et al. Bacterial sirtuins overview: an open niche to explore. *Front. Microbiol.* **12**, 744416 (2021).
22. Wang, Y., Sheng, G., Juranek, S., Tuschl, T. & Patel, D. J. Structure of the guide-strand-containing argonaute silencing complex. *Nature* **456**, 209–213 (2008).
23. Frank, F., Sonenberg, N. & Nagar, B. Structural basis for 5′-nucleotide base-specific recognition of guide RNA by human AGO2. *Nature* **465**, 818–822 (2010).
24. Ka, D., Oh, H., Park, E., Kim, J. H. & Bae, E. Structural and functional evidence of bacterial antiphage protection by Thoeris defense system via NAD+ degradation. *Nat. Commun.* **11**, 2816 (2020).
25. Kim, S., Jung, Y. & Lim, D. Argonaute system of *Kordia jejudonensis* is a heterodimeric nucleic acid-guided nuclease. *Biochem. Biophys. Res. Commun.* **525**, 755–758 (2020).
26. Dasgupta, S., Masukata, H. & Tomizawa, J. Multiple mechanisms for initiation of ColE1 DNA replication: DNA synthesis in the presence and absence of ribonuclease H. *Cell* **51**, 1113–1122 (1987).
27. del Solar, G., Giraldo, R., Ruiz-Echevarria, M. J., Espinosa, M. & Diaz-Orejas, R. Replication and control of circular bacterial plasmids. *Microbiol. Mol. Biol. Rev.* **62**, 434–464 (1998).
28. Selzer, G., Som, T., Itoh, T. & Tomizawa, J. The origin of replication of plasmid p15A and comparative studies on the nucleotide sequences around the origin of related plasmids. *Cell* **32**, 119–129 (1983).
29. Garb, J. et al. Multiple phage resistance systems inhibit infection via SIR2-dependent NAD+ depletion. Preprint at *bioRxiv* https://doi.org/10.1101/2021.12.14.472415 (2021).
30. Tal, N. et al. Cyclic CMP and cyclic UMP mediate bacterial immunity against phages. *Cell* **184**, 5728–5739.e16 (2021).
31. Ofir, G. et al. Antiviral activity of bacterial TIR domains via immune signalling molecules. *Nature* **600**, 116–120 (2021).
32. Zheng, L., Baumann, U. & Reymond, J. L. An efficient one-step site-directed and site-saturation mutagenesis protocol. *Nucleic Acids Res.* **32**, e115 (2004).
33. Ofir, G. et al. DISARM is a widespread bacterial defence system with broad anti-phage activities. *Nat. Microbiol.* **3**, 90–98 (2018).
34. Doron, S. et al. Systematic discovery of antiphage defense systems in the microbial pangenome. *Science* **359**, eaar4120 (2018).
35. Mazzocco, A., Waddell, T. E., Lingohr, E. & Johnson, R. P. Enumeration of bacteriophages using the small drop plaque assay system. *Methods Mol. Biol.* https://doi.org/10.1007/978-1-60327-164-6_9 (2009).
36. Blanchet, C. E. et al. Versatile sample environments and automation for biological solution X-ray scattering experiments at the P12 beamline (PETRA III, DESY). *J. Appl. Crystallogr.* **48**, 431–443 (2015).
37. Franke, D., Petoukhov, M. V., Konarev, P. V. & Panjkovich, A. ATSAS 2.8: a comprehensive data analysis suite for small-angle scattering from macromolecular solutions. *J. Appl. Crystallogr.* https://doi.org/10.1107/S1600576717007786 (2017).
38. Konarev, P. V., Volkov, V. V., Sokolova, A. V., Koch, M. H. J. & Svergun, D. I. PRIMUS - a Windows-PC based system for small-angle scattering data analysis. *J. Appl. Crystallogr.* **36**, 1277–1282 (2003).
39. Svergun, D. I. Determination of the regularization parameter in indirect-transform methods using perceptual criteria. *J. Appl. Crystallogr.* **25**, 495–503 (1992).

40. Durand, D. et al. NADPH oxidase activator p67phox behaves in solution as a multidomain protein with semi-flexible linkers. *J. Struct. Biol.* **169**, 45–53 (2010).
41. Svergun, D. I., Petoukhov, M. V. & Koch, M. H. J. Determination of domain structure of proteins from x-ray solution scattering. *Biophys. J.* **80**, 2946–2953 (2001).
42. Fischer, H., de Oliveira Neto, M., Napolitano, H. B., Polikarpov, I. & Craievich, A. F. Determination of the molecular weight of proteins in solution from a single small-angle X-ray scattering measurement on a relative scale. *J. Appl. Crystallogr.* **43**, 101–109 (2010).
43. Schubert, M., Lindgreen, S. & Orlando, L. AdapterRemoval v2: rapid adapter trimming, identification, and read merging. *BMC Res. Notes* **9**, 88 (2016).
44. Li, H. & Durbin, R. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* **25**, 1754–1760 (2009).
45. Andrews, S. et al. *FastQC: A Quality Control Tool for High-Throughput Sequence Data* (2010); https://www.bioinformatics.babraham.ac.uk/projects/fastqc/
46. Li, H. et al. The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **25**, 2078–2079 (2009).
47. Crooks, G. E., Hon, G., Chandonia, J.-M. & Brenner, S. E. WebLogo: a sequence logo generator. *Genome Res.* **14**, 1188–1190 (2004).
48. Quinlan, A. R. BEDTools: the swiss-army tool for genome feature analysis. *Curr. Protoc. Bioinform.* **47**, 11.12.1–11.12.34 (2014).
49. Wang, L., Wang, S. & Li, W. RSeQC: quality control of RNA-seq experiments. *Bioinformatics* **28**, 2184–2185 (2012).
50. Thorvaldsdottir, H., Robinson, J. T. & Mesirov, J. P. Integrative Genomics Viewer (IGV): high-performance genomics data visualization and exploration. *Brief. Bioinform.* **14**, 178–192 (2013).
51. Koopal, B. et al. Short prokaryotic Argonaute systems trigger cell death upon detection of invading DNA. *Cell* **185**, 1471–1486.e19 (2022).
52. Kaya, E. et al. A bacterial Argonaute with noncanonical guide RNA specificity. *Proc. Natl Acad. Sci. USA* **113**, 4057–4062 (2016).

Acknowledgements

We thank the Siksnys Laboratory members for their comments on the manuscript and fruitful discussion, and T. de Garay at Refeyn Ltd. for the mass photometry experiments. This work was supported by the European Social Fund (09.3.3-LMT-K-712-01-0126 to V.S.) under a grant agreement with the Research Council of Lithuania (LMTLT), the Israel Science Foundation (grant ISF 296/21 to R.S.), and the Deutsche Forschungsgemeinschaft (SPP 2330, grant 464312965). Funding for open access was provided by Vilnius University.

Author contributions

V.S. and M.Z. designed the study; K.T. and C.V. performed bioinformatics and structural modelling; A.L. performed the phage restriction experiments; E.S., D.D., E.G., S.A. and U.T. performed the plasmid transformation experiments; A.S. purified the proteins and performed the SEC-(MALS) experiments; R.G. performed the SEC experiments; E.M. performed the SAXS experiments; E.Z., E.G., D.D. and R.G. performed the EMSA experiments; E.Z. and E.J. reconstituted the GsSir2/Ago-RNA complex and performed the biochemical analysis; A.G. performed the RNA-seq analysis; A.R. performed the mass spectrometry analysis; D.D. performed the NAD^+ determination experiments in vivo; E.Z. performed the NAD^+ hydrolysis experiments in vitro; R.S., V.S. and M.Z. analysed the data; M.Z. wrote the initial manuscript with input from E.G., R.S., V.S. and other authors. All authors approved the final version.

Competing interests

V.S. is the chairman of CasZyme. R.S. is the scientific founder of BiomX and Ecophage. The remaining authors declare no competing interests.

Additional information

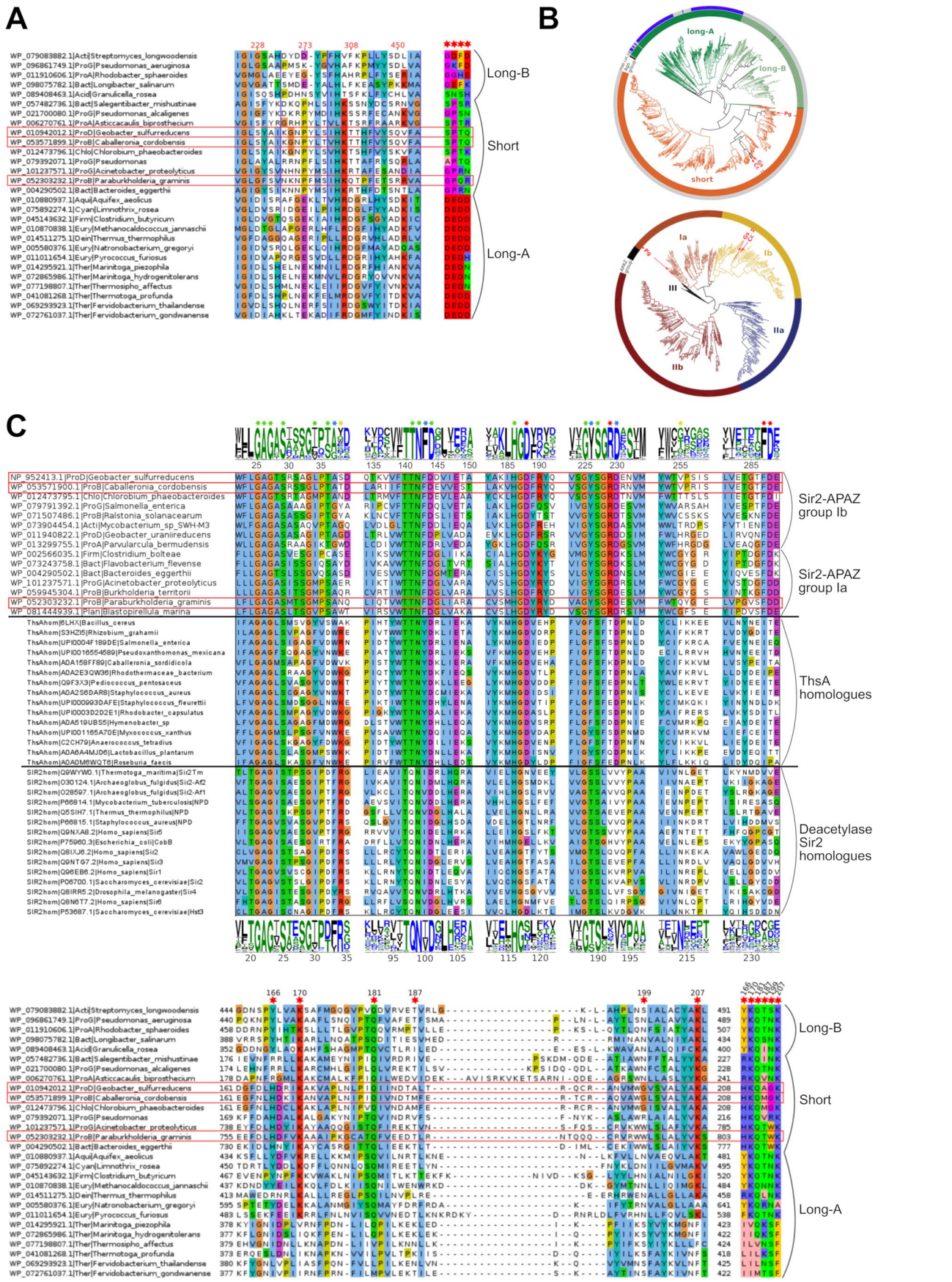
Extended data is available for this paper at https://doi.org/10.1038/s41564-022-01239-0.

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41564-022-01239-0.

Correspondence and requests for materials should be addressed to Mindaugas Zaremba or Virginijus Siksnys.

Peer review information *Nature Microbiology* thanks Malcolm White and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. Peer reviewer reports are available.

https://doi.org/10.1038/s41564-022-01239-0

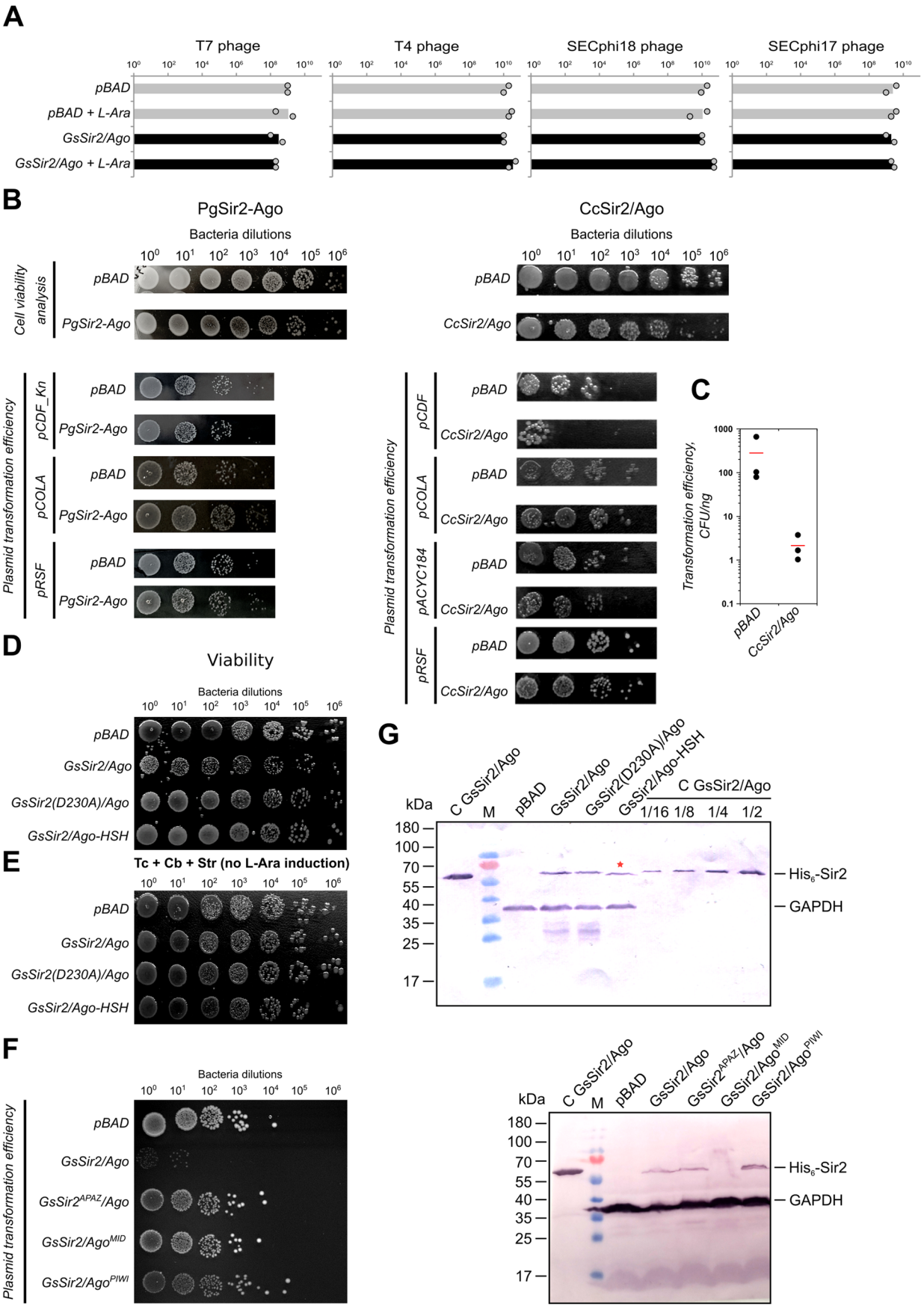


Extended Data Fig. 1 | See next page for caption.

https://doi.org/10.1038/s41564-022-01239-0

Extended Data Fig. 1 | Bioinformatic analysis. a, PIWI catalytic tetrad DEDX alignment. The 4 catalytic residues (red numbers indicate positions of corresponding GsAgo positions) are shown in 4 motifs of ±3 positions. The motifs are separated by vertical blue lines. Sequence names consist of the following: NCBI sequence ID, abbreviated phylum (for example, 'ProG' – gamma-proteobacteria) and organism name. **b**, Top - a circular phylogenetic tree was generated according to supplementary data provided with Ryazansky et al.[8] Long-A pAgo variants are coloured in green (truncated variants without the PAZ domain, light green), long-B pAgo proteins are light green (truncated variants without PAZ, green), and short pAgo proteins are orange. pAgo proteins containing the catalytic tetrad DEDX in their PIWI domain are indicated in blue on the outer circle; pAgos with inactivated PIWI domain are indicated in light grey on the outer circle. pAgo proteins of the GsSir2/Ago, CcSir2/Ago and PgSir2-Ago systems are indicated by 'Gs', 'Cc' and 'Pg', respectively. Bottom - Circular phylogenetic tree of APAZ domains. The circular phylogenetic tree of the five groups of APAZ domains was generated using APAZ domain alignments from Ryazansky et al.[8] supplementary file 7. **c**, Top - Combined alignment of Sir2 domains. Alignment consists of 3 parts, separated by horizontal black lines. In the top part, the Sir2 domain sequences of the GsSir2, CcSir2, PgSir2-Ago and homologues are shown. Logos above depict the conservation of Sir2 domains of Ia and Ib groups. The indicated position numbers correspond to the GsSir2 sequence. In the bottom part, homologues (sirtuins) of catalytically active *Thermotoga maritima* Sir2 (TmSir2) deacetylase are shown. Logos below indicate the conservation of these homologues. The position numbers correspond to the TmSir2 sequence. Sequences of six motifs that include all positions that form the NAD^+-binding pocket, as seen in the TmSir2 structure (PDB ID 2H4F) are shown. Sequence names for the top alignment consist of sequence ID, abbreviated phylum and organism name. Sequence names for bottom alignment all start with 'Sir2hom' followed by sequence ID, organism name and short protein name (based on annotation). Stars above the logos indicate residues in the NAD^+-binding pocket of canonical sirtuins (for example, TmSir2) that are also conserved. Star colours indicate conservation between the two groups: green – conserved in both canonical sirtuins and GsSir2-like; blue – conserved in both groups, but different; yellow – conserved only in canonical sirtuins; red – conserved only in GsSir2-like proteins. In the middle, alignment of ThsA homologues with Sir2 domains. Bottom - MID domain alignment. Red stars indicate positions of amino acids involved in the binding of the 5′-P end of the guide nucleic acid. The numbering above corresponds to the GsAgo sequence. Additionally, concatenated alignment of just the 6 indicated positions is shown on the right. The three sequences of interest are indicated with red rectangles. Numbers on the left and right of the alignment indicate the first and last positions in the alignment for each sequence.

https://doi.org/10.1038/s41564-022-01239-0

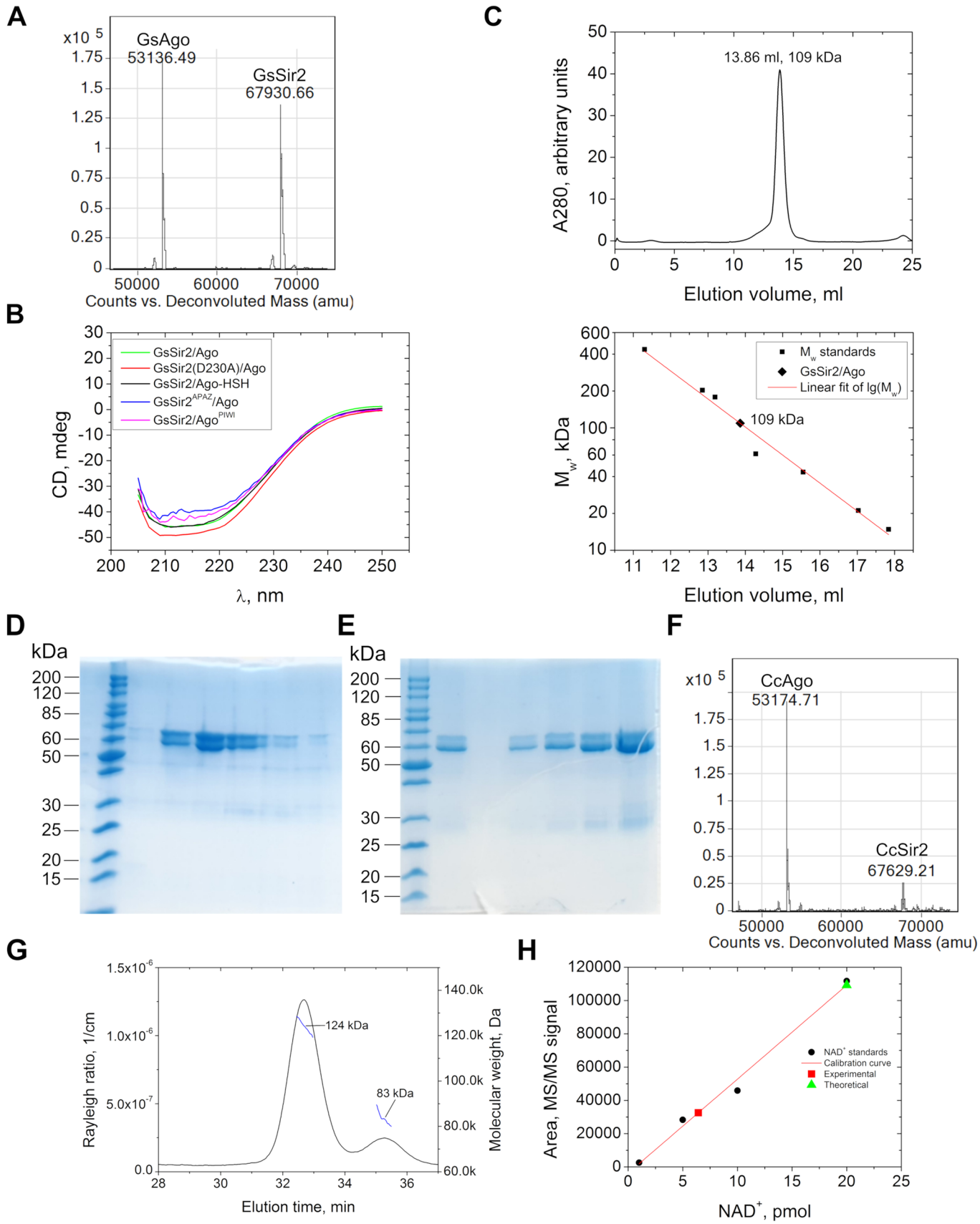


Extended Data Fig. 2 | See next page for caption.

https://doi.org/10.1038/s41564-022-01239-0

Extended Data Fig. 2 | *In vivo* characterization of Sir2/Ago systems. a, Efficiency of plating (EOP) of 4 phages infecting *E. coli* cells with and without the GsSir2/Ago system from *Geobacter sulfurreducens*, where the GsSir2/Ago system exhibits no defence activity. The x-axis represents the number of p.f.u. Shown are the means of two replicates in the absence and in the presence of the inducer L-arabinose (L-Ara), with individual data points superimposed. Grey bars represent efficiency of plating (EOP) on pAgo-lacking cells and black bars are EOP in pAgo-containing cells. **b**, Left - qualitative characterization of plasmid restriction capabilities of PgSir2-Ago system in *E. coli* strain DH10B. Top: comparison of cell viability in the presence or absence of plasmid-borne PgSir2-Ago expression. Bottom: comparison of plasmid transformation efficiencies in the presence or absence of plasmid-borne PgSir2-Ago expression. Right - qualitative characterization of plasmid restriction capabilities of CcSir2/Ago system in *E. coli* strain BL21-AI: top - comparison of cell viability in the presence or absence of plasmid-borne CcSir2/Ago expression. Bottom - comparison of plasmid transformation efficiencies in the presence or absence of plasmid-borne CcSir2/Ago expression. **c**, Quantification of transformation efficiencies for pCDF plasmid with CcSir2/Ago system (three independent replicates, the red line represents average transformation efficiency). **d**, Cell viability control of BL21-AI, expressing GsSir2/Ago and mutants. **e**, Control for Fig. 2e – cells contain the pCDF plasmid, however, expression of the GsSir2/Ago system is not induced. **f**, Qualitative evaluation of pCDF plasmid transformation efficiency in *E. coli* cells carrying GsSir2/Ago mutants ($GsSir2^{APAZ}$/Ago, $GsSir2/Ago^{MID}$ and $GsSir2/Ago^{PIWI}$) of the putative surface of the interaction with nucleic acids. **g**, Expression analysis of GsSir2/Ago assayed by Western blot. Top: semiquantitative Western blot of the wt GsSir2/Ago complex and its mutants. Numbers above the lanes indicate which part of control protein amount is loaded. The red star shows the lane where the His-tag is on the C-terminus of Ago, rather than the N-terminus of Sir2. GAPDH, loading control. Bottom: Expression analysis of GsSir2/Ago mutants of the putative surface of the interaction with nucleic acids. Three replicates.

https://doi.org/10.1038/s41564-022-01239-0

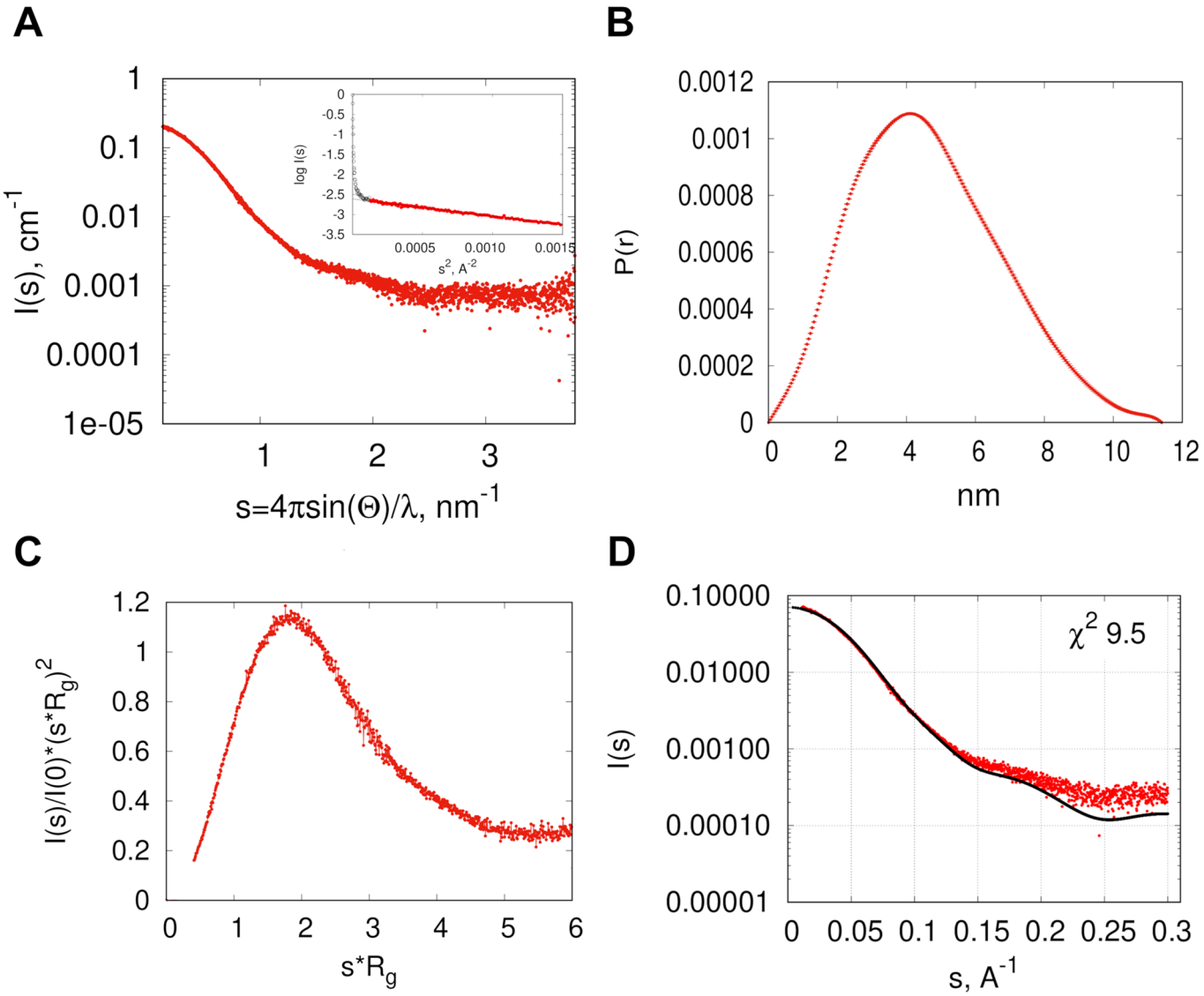


Extended Data Fig. 3 | See next page for caption.

https://doi.org/10.1038/s41564-022-01239-0

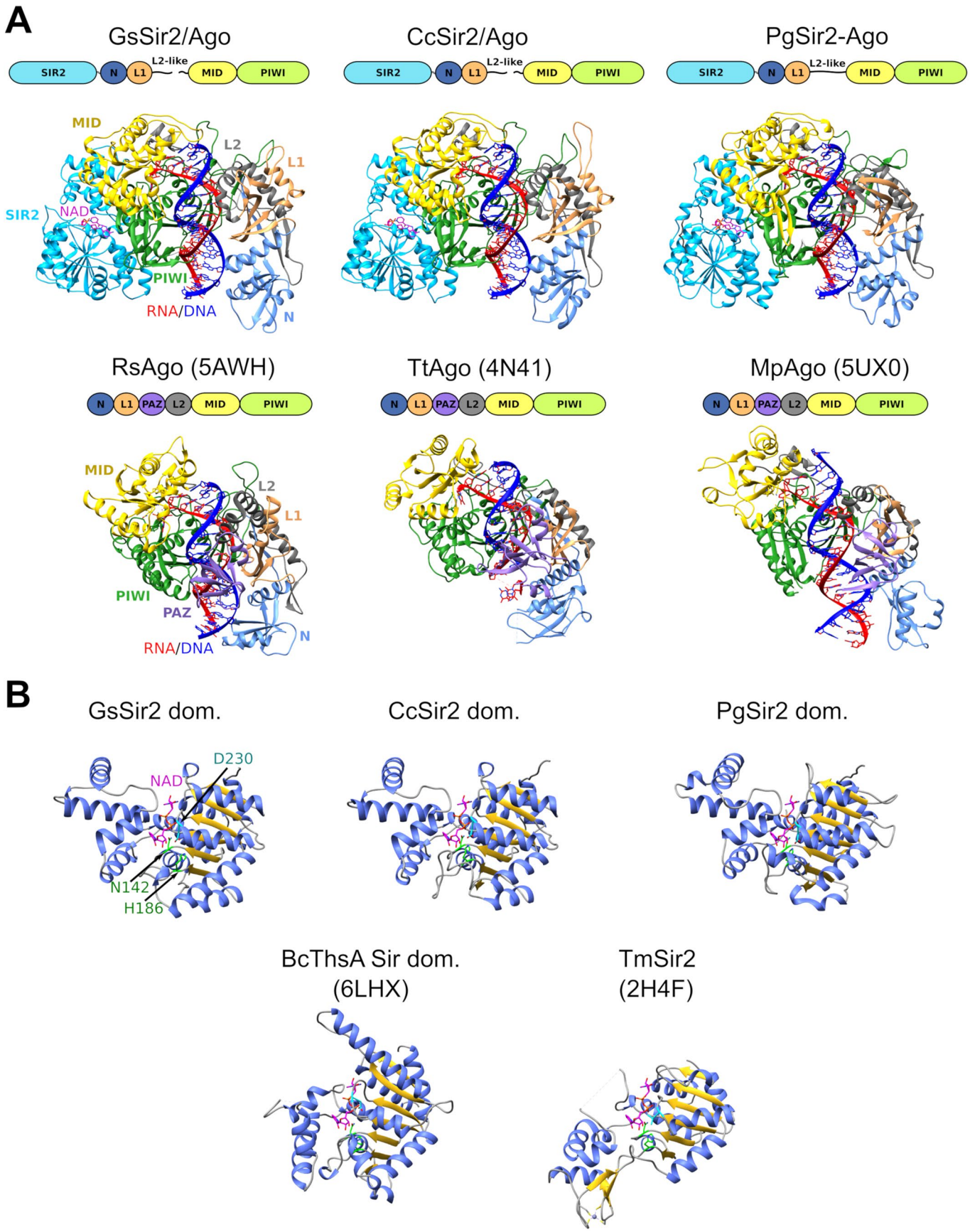
Extended Data Fig. 3 | Purification and characterization of Sir2/Ago complexes. a, MS analysis of wt GsSir2/Ago complex. The theoretical Mw of the GsSir2 and GsAgo proteins (without 1st Met) are 67929.51 Da and 53135.61 Da, respectively. **b**, CD spectra of wt GsSir2/Ago and mutants. Mutant spectra are similar to that of a natively folded protein. **c**, Size-exclusion chromatography of wt GsSir2/Ago, showing elution volume and comparing to mass standards. According to mass spectrometry of the purified GsSir2/Ago complex, the molar mass of the GsSir2/Ago heterodimer is 121 kDa. **d**, SDS–PAGE analysis of fractions containing the CcSir2/Ago complex eluted from Heparin column. Densitometric inspection shows that Sir2 and Ago proteins are in the ratio ~1:1. Single replicate. **e**, SDS–PAGE analysis of the CcSir2/Ago stock after dialysis against a storage buffer. Various amounts of the stock solution were loaded on the gel. Densitometric inspection shows that Sir2 and Ago proteins are in the ratio ~0.3:1. Single replicate. **f**, MS analysis of the CcSir2/Ago complex. The experimental masses (53174.71 Da and 67629.21 Da) are close to the theoretical molecular masses of the Ago protein (53173.91 Da) and the Sir2 protein with the truncated tag at the N terminus (67626.24 Da). **g**, SEC-MALS analysis of the CcSir2/Ago complex. The experimental mass of 124 kDa is close to the theoretical molecular mass of the CcSir2/Ago heterodimer (121 kDa). **h**, MS/MS calibration curve of NAD^+ standard (marked in black) and the observed amount of NAD^+ (marked in red) in the CcSir2/Ago complex (20 pmol according to the Ago protein). The discrepancy between the expected amount of NAD^+ (20 pmol, marked in green) and the actual amount (6.45 pmol, marked in red) was due to the decrease of the Sir2 protein in the CcSir2/Ago preparation (see **e**).

https://doi.org/10.1038/s41564-022-01239-0



Extended Data Fig. 4 | SAXS data. a, Scattering data on an absolute scale. Linear Guinier plot of the initial part of the scattering curve is in the insert. Points cut from the further processing are shown with empty black symbols. **b**, Kratky plot, normalized by Rg and I(0) parameters. **c**, Pair distance distribution function. **d**, CRYSOL Fit of the scattering curve calculated from the GsSir2/Ago AlphaFold model (black curve) with SAXS data (red points).

https://doi.org/10.1038/s41564-022-01239-0

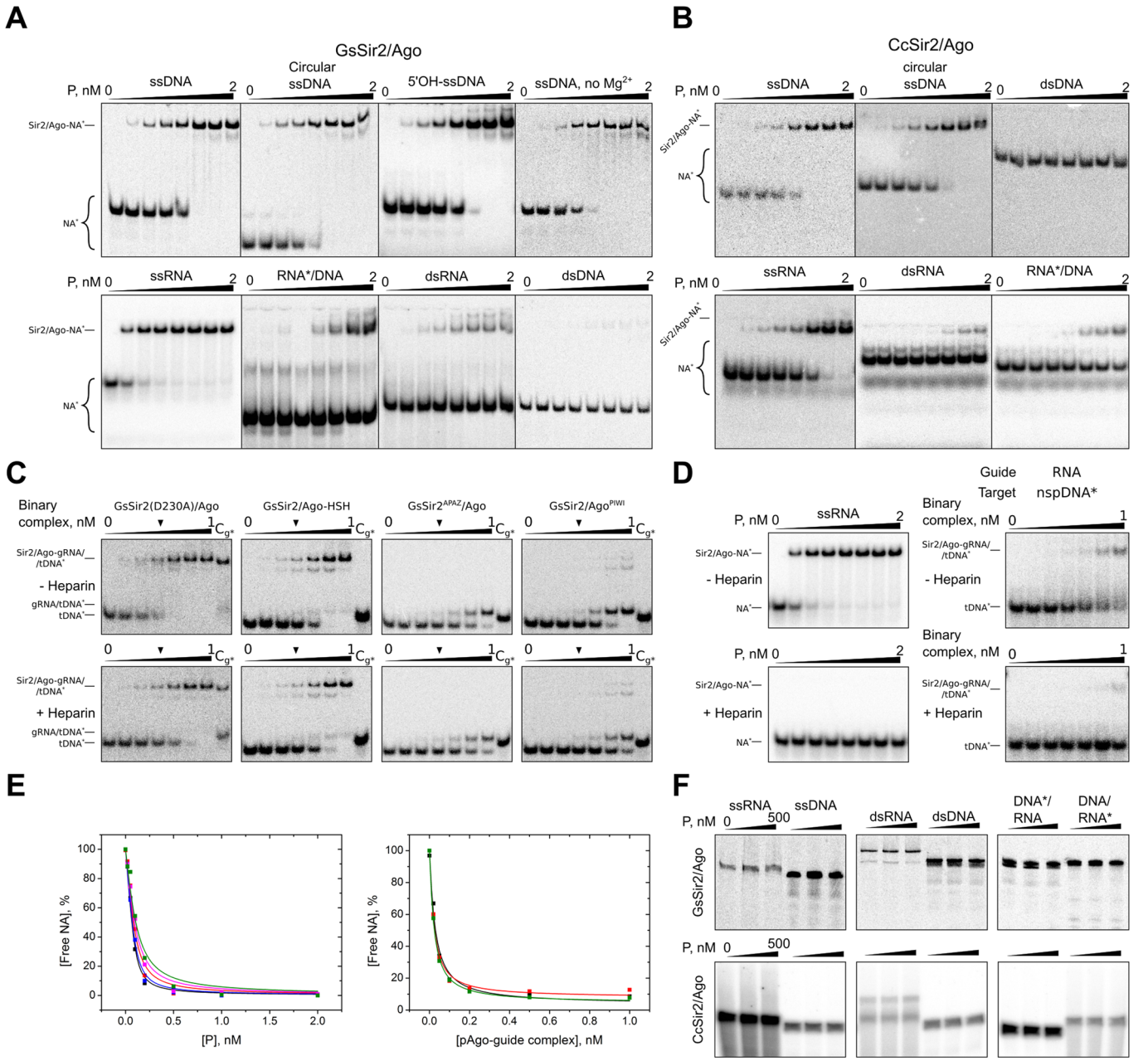


Extended Data Fig. 5 | See next page for caption.

https://doi.org/10.1038/s41564-022-01239-0

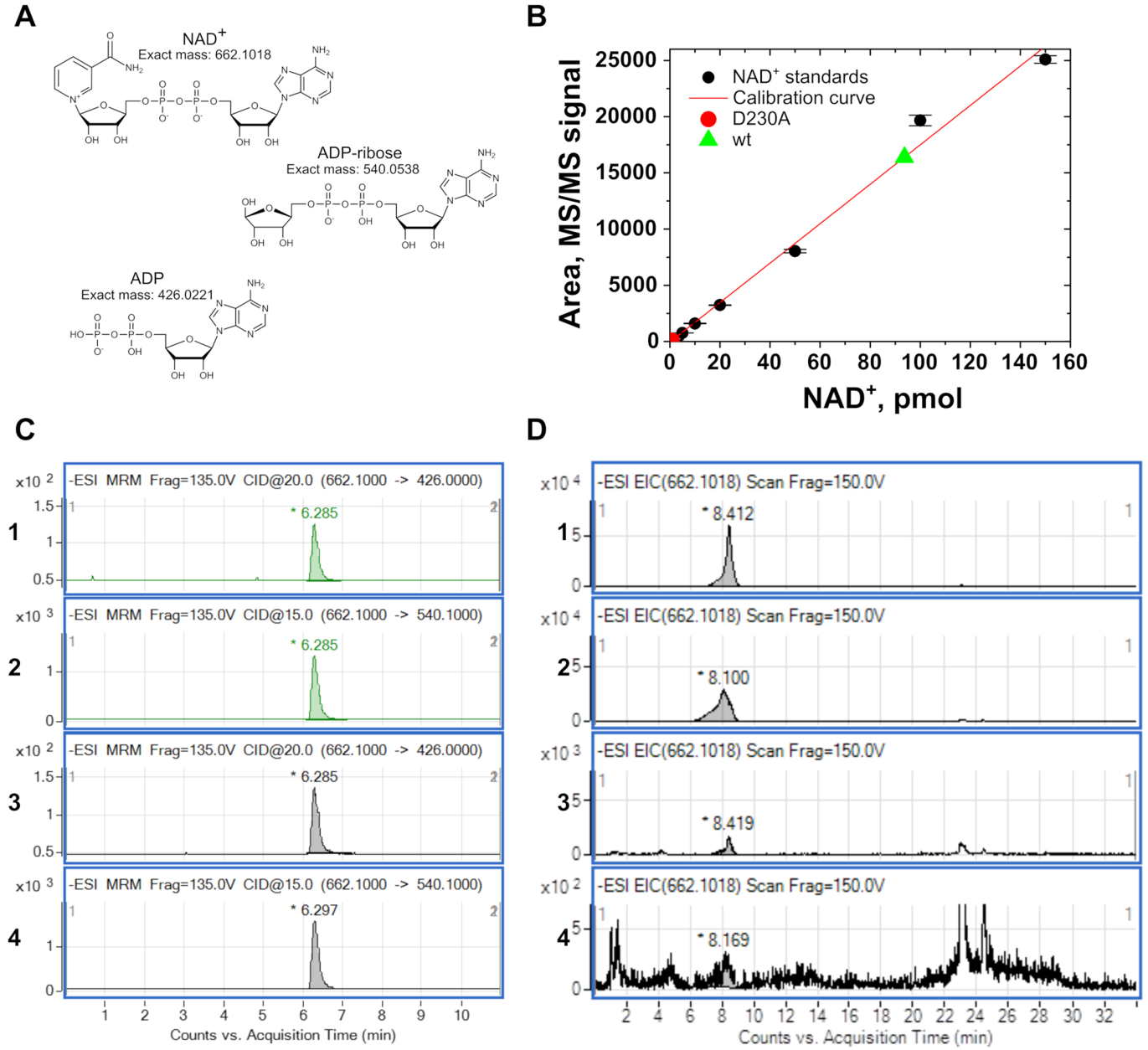
Extended Data Fig. 5 | Structural analysis. a, Comparison of GsSir2/Ago, CcSir2/Ago and PgSir2-Ago AlphaFold models with the X-ray structures of long pAgos. Structures are coloured by domains, schematic domain architecture is given above each structure. Guide RNA and target DNA strands are coloured red and blue, respectively. PDB ID codes for long pAgo structures are given in parentheses. RsAgo represents long-B group, TtAgo – longA group of long pAgos (based on Ryazansky et al. classification[8]). MpAgo has a distinct OH-type MID domain that is specific to 5′-OH instead of phosphate (MID domain classification – Ryazansky et al.[8], MpAgo MID domain biochemical assay - Kaya et al.[52]). In Sir2/Ago models the N, L1 and L2-like domains previously identified as the APAZ domain correspond to the analogous domains of long pAgos. **b**, Gs, Cc and Pg Sir2 structural models (cut from full-length models) compared to canonical Sir2 deacetylase TmSir2 and the Sir2 domain of Thoeris defence system protein ThsA. Structures are coloured based on secondary structure. Positions corresponding to ThsA Sir2 N112 and H152 are indicated in green. These residues have been shown to be critical for NAD^+ hydrolysis in ThsA[24]. GsSir2 D230 and corresponding positions in other structures are indicated in cyan. NAD^+ was also superimposed on the ThsA Sir2 structure from TmSir2.

https://doi.org/10.1038/s41564-022-01239-0



Extended Data Fig. 6 | EMSA and nucleic acid cleavage experiments. a, **b**, Binding of single- and double-stranded oligonucleotides by wt GsSir2/Ago and wt CcSir2/Ago, respectively. A radiolabelled strand indicated by the asterisk. **c**, Binding of complementary DNA targets by GsSir2/Ago binary complexes pre-loaded with RNA guide containing 5'-phosphate terminus in the presence or absence of heparin. To show that no displacement of the radiolabelled guide by the target strand is observed, a control (Cg*) equivalent to the experimental lane marked by a black triangle, but with the guide, rather than the target, bearing the radioactive label, was performed. **d**, Control EMSA experiments of ssRNA guide binding by wt GsSir2/Ago (left) and non-complementary DNA target binding by the binary wt GsSir2/Ago-gRNA complex in the presence and absence of heparin. **e**, Representative binding fit curves of several independent replicates used to calculate K_d of ssRNA guide binding by wt GsSir2/Ago (left) and of target DNA binding by wt GsSir2/Ago-gRNA complex (right). **f**, (No) cleavage activity of various DNA and RNA oligonucleotides by wt GsSir2/Ago and wt CcSir2/Ago. Reaction products were resolved on a 21% denaturing polyacrylamide gel. In heteroduplexes, the asterisk indicates the radiolabelled strand. For panels **A-D** and **F**, at least three independent replicates were performed for each experiment.

https://doi.org/10.1038/s41564-022-01239-0



Extended Data Fig. 7 | The GsSir2/Ago complex binds NAD^+ and causes its depletion. a, Two ion transitions were used to detect NAD^+ in the analysed samples: 662.1→540.1 and 662.1→426.0. **b**, MS/MS calibration curve of NAD^+ standard (marked in black, two replicates) and the observed amount of NAD^+ in two samples: 93.7 pmol in the wt GsSir2/Ago sample (marked in green), 0.7 pmol in the sample D230A (marked in red). Black dots represent the means of two replicates and error bars are the standard deviation. **c**, Detection of NAD^+. Comparison of the extracted LC-MS/MS chromatograms: ion transition 662.1→426.0 of wt GsSir2/Ago sample (panel 1) and NAD^+ standard (panel 3); ion transition 662.1→540.1 of wt GsSir2/Ago sample (panel 2) and NAD^+ standard (panel 4). Green curves - wt GsSir2/Ago sample, grey curves – NAD^+ standard. **d**, Mass chromatogram. For NAD^+ detection, an extracted ion current (EIC) for [M-H]- m/z = 662.1018 was used. The comparison of EIC signals shows that the amount of NAD^+ in the samples of the non-induced GsSir2/Ago system in the absence (panel 1) and presence (panel 2) of pCDF plasmid is almost the same, while a significant decrease is observed in the sample of the induced GsSir2/Ago system in the absence of pCDF plasmid (panel 3) and only traces of NAD^+ are detected in the presence of pCDF plasmid (panel 4).

nature portfolio

Corresponding author(s): Mindaugas Zaremba, Virginijus Siksnys

Last updated by author(s): Jul 28, 2022

Reporting Summary

Nature Portfolio wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Portfolio policies, see our Editorial Policies and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed	
☐	☒	The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒	A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☒	☐	The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☒	☐	A description of all covariates tested
☒	☐	A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒	A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☒	☐	For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☒	☐	For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☒	☐	For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☒	☐	Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection	SAXS data were collected using SASFLOW software. SEC-MALS data were collected using Astra software. Mass photometry data collection was performed using Refeyn AcquireMP. Mass spectrometry data were collected using Agilent MassHunter Workstation software: Data Acquisition for Triple Quad B.04.01; LC/MS Data Acquisition B.02.01 for 6200 series TOF & 6500 series Q-TOF.
Data analysis	All software and code used in this study has been described in published literature: ATSAS 2.8.4 (r10552) suite; PRIMUS; GNOM5; GASBOR; DATVC and DATMW from ATSAS suite; server SAXSMoW; Astra; BLAST; CLANS; MAFFT; Jalview; WebLogo 3; HHpred; FastTree; iTOL; trimAl; AlphaFold2 implemented as ColabFold; VoroMQA; Dali; UCSF Chimera; BWA-MEM v0.7.17; FastQC v0.11.8; SAMtools v1.7; bedtools v2.26.0; FPKM_count.py v4.0.0 of RSeqQC package; IGV v2.5.2. Mass spectrometry data analysis was performed using Agilent MassHunter Qualitative Analysis B.05.00. Mass photometry data analysis was performed using Refeyn DiscoverMP. Other data analysis was performed using Excel for Microsoft 360 and OriginPro 8.1.

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Portfolio guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A description of any restrictions on data availability
- For clinical datasets or third party data, please ensure that the statement adheres to our policy

All data are available in the manuscript and the supplementary material. Additionally, plasmid and protein sequences are deposited on Benchling.com with appropriate links provided. SAXS data are available on the Small Angle Scattering Biological Data Bank (SASBDB) under SASBDB ID SASDNH2. RNA sequencing data are available on the NCBI Sequence Read Archive under BioProject ID PRJNA851009.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	Sample size (n=3) was based on the work and experience of other groups in the field who generate reproducible results in experiments of similar design (no statistical methods were used to determine the sample size).
Data exclusions	No data was excluded from the analysis
Replication	Reproducibility was ensured by repeating most of the experiments independently at least 3 times.
Randomization	Samples were not randomized as it is not applicable for the current design of the study.
Blinding	Investigators were not blinded as it is not applicable for the current design of the study.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☐	☒ Antibodies
☒	☐ Eukaryotic cell lines
☒	☐ Palaeontology and archaeology
☒	☐ Animals and other organisms
☒	☐ Human research participants
☒	☐ Clinical data
☒	☐ Dual use research of concern

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☒	☐ Flow cytometry
☒	☐ MRI-based neuroimaging

Antibodies

Antibodies used	6x-His Tag monoclonal antibody (ThermoFisher, cat. #MA1-21315, RRID AB_557403, Lot # WE323793, clone HIS.H8); Goat anti-Mouse IgG (H+L) Secondary Antibody, AP conjugated (ThermoFisher, cat. #31320, RRID AB_228304, Lot # VH311913, polyclonal).
Validation	All commercial antibodies were validated by manufacturers for indicated species and application and can be accessed on the manufacturers' websites for the respective antibodies.