Characterization of the self-targeting Type IV CRISPR interference system in *Pseudomonas oleovorans*

Bacterial Type IV CRISPR-Cas systems are thought to rely on multi-subunit ribonucleoprotein complexes to interfere with mobile genetic elements, but the substrate requirements and potential DNA nuclease activities for many systems within this type are uncharacterized. Here we show that the native *Pseudomonas oleovorans* Type IV-A CRISPR-Cas system targets DNA in a PAM-dependent manner and elicits interference without showing DNA nuclease activity. We found that the first crRNA of *P. oleovorans* contains a perfect match in the host gene coding for the Type IV pilus biogenesis protein PilN. Deletion of the native Type IV CRISPR array resulted in upregulation of pilN operon transcription in the absence of genome cleavage, indicating that Type IV-A CRISPR-Cas systems can function in host gene regulation. These systems resemble CRISPR interference (CRISPRi) methodology but represent a natural CRISPRi-like system that is found in many *Pseudomonas* and *Klebsiella* species and allows for gene silencing using engineered crRNAs.

Clustered regularly interspaced short palindromic repeats (CRISPR) and CRISPR-associated (Cas) proteins elicit adaptive immunity in many bacterial species. Foreign DNA fragments, termed protospacers, with short protospacer-adjacent motifs (PAM) are captured by Cas1-Cas2 adaptation complexes and inserted into an extending CRISPR locus. Transcription and processing of the CRISPR array results in mature CRISPR RNAs (crRNAs) that guide CRISPR ribonucleoprotein complexes (crRNPs) towards complementary nucleic acid target sequences. Two classes and six types of CRISPR-Cas systems have been classified and Type IV CRISPR-Cas remains the only type without description of its endogenous activity. Subtype IV-A CRISPR-Cas was first discovered in the genome of *Acidithiobacillus ferrooxidans* and later shown to contain the signature proteins CsF1 and the crRNP backbone subunits CsF2. Type IV-A systems are usually present on large plasmids (>200 kb) and lack adaptation modules and apparent DNA nucleases such as Cas3 or Cas10. Identification of plasmid-borne protospacers targeting conjugative elements suggested that Type IV systems are mainly involved in inter-plasmid competition. The characterization of recombinant Cas proteins of the Type IV-A system of *Aromatoleum aromaticum* indicated the presence of crRNPs that resemble Type I effectors. Heterologous anti-plasmid interference activity was reported for a Type IV-A system of *Pseudomonas aeruginosa* and shown to require DInG activity. However, it is not known how its crRNPs can combat plasmids without an apparent target DNA nuclease. To pinpoint the native activity of a Type IV-A CRISPR-Cas system, we identified and characterized a complete cas gene module and a neighbouring CRISPR array on a megaplasmid of *Pseudomonas oleovorans*. 

Received: 28 January 2022
Accepted: 4 August 2022
Published online: 29 September 2022

Xiaohan Guo, Mariana Sanchez-Londono, José Vicente Gomes-Filho, Rogelio Hernandez-Tamayo, Selina Rust, Leah M. Immelmann, Pascal Schäfer, Julia Wiegel, Peter L. Graumann and Lennart Randau.

1Department of Biology, Philipps Universität Marburg, Marburg, Germany. 2Department of Chemistry, Philipps-Universität Marburg, Marburg, Germany. 3SYNMIKRO, LOEWE Center for Synthetic Microbiology, Marburg, Germany. 4These authors contributed equally: Xiaohan Guo, Mariana Sanchez-Londono. e-mail: lennart.randau@staff.uni-marburg.de
adaptation modules. Small RNA libraries of cas1-cas2 I-E and a Type I-F CRISPR-Cas system are present in the host genome P. oleovorans strain DSM 1045 (Fig. 1a,b). In addition, a Type plasmid of

The Type IV-A CRISPR-Cas system was identified on an IncP-9 family

The CRISPR arrays of P. oleovorans contain 8-nt-long 5′-terminal repeat tags (Type IV: 5′-GUGAGCGG-3′, Type I-F: 5′-CUCAGAAA-3′) that indicate processing of the CRISPR array transcript at the base of hairpin structures

Results
The CRISPR arrays of P. oleovorans
The Type IV-A CRISPR-Cas system was identified on an IncP-9 family plasmid of P. oleovorans strain DSM 1045 (Fig. 1a,b). In addition, a Type I-E and a Type I-F CRISPR-Cas system are present in the host genome and contain cas1-cas2 adaptation modules. Small RNA libraries of P. oleovorans DSM1045 were sequenced using Illumina RNA-seq methodology and sequencing reads were found to map to all three CRISPR arrays (Fig. 1a and Extended Data Fig. 1). Mature crRNAs were identified to contain 8-nt-long 5′-terminal repeat tags (Type IV: 5′-GUGAGCGG-3′, Type I-E: 5′-AUGAACCG-3′, Type I-F: 5′-CUCAGAAA-3′) that indicate processing of the CRISPR array transcript at the base of hairpin structures

Article https://doi.org/10.1038/s41564-022-01229-2
of the individual repeats. Each CRISPR-Cas system was found to contain a Cas6-family crRNA endonuclease that enables subtype-specific crRNA maturation. Abundance of individual crRNAs was highly variable for different CRISPR arrays, but the large numbers of mature crRNA sequences suggests that all three native systems are active.

The Type IV-A CRISPR locus contains 20 spacers, generating 20 crRNAs with different targeting potential. BLAST search analyses identified targets for the spacers of the Type IV-A CRISPR array in transposon and plasmid elements. Alignment of the genomic context of these target protospacers unveiled a consensus 5′-AAG-3′ PAM motif at the 5′ end of the non-targeting strand (Fig. 1c). In addition, we identified two targets of Type I-E crRNAs and both protospacers also exhibited a 5′-AAG-3′ PAM (Supplementary Table 1). Therefore, it is plausible that the Type I-E adaptation module defines the PAM requirements for both systems and compensates for the absence of cas1-cas2 genes in the Type IV-A CRISPR-Cas system.

PAM-dependent Type IV-A CRISPR-Cas activity

We investigated *P. oleovorans* Type IV-A CRISPR-Cas activity using a plasmid expressing superfolder green fluorescent protein (sfGFP) and different versions of a second pUC18 plasmid coding for crRNAs that target selected *gfp* protospacers (Supplementary Table 2). The GFP signal was quantified using flow cytometry and revealed GFP signal reduction for protospacer targets in the *gfp* coding and non-coding strand in the presence of a 5′-AAG-3′ PAM (Fig. 1d). A protospacer target in the *gfp* coding strand with 5′-CGG-3′ PAM and a non-targeting crRNA did not result in significant GFP signal reduction, suggesting PAM-dependent CRISPR interference. To follow plasmid curing in this system, *P. oleovorans* was transformed with target and crRNA-production plasmids and cultivated for 12 h without antibiotic selection for the target plasmid (Supplementary Table 2). The cells were then transferred into a medium with antibiotics to select for the presence of both plasmids. PAM-dependent plasmid curing was observed, indicating that the native Type IV-A crRNPs can interfere with plasmid replication (Fig. 1e).

To analyse this targeting mechanism in greater detail, we designed a heterologous *E. coli* BL21-AI system to produce recombinant *P. oleovorans* Type IV-A crRNPs. The genes *csf1* and a minimal CRISPR array consisting of a single spacer-repeat-spacer unit were provided on two plasmids (Supplementary Table 2). Protospacer-containing target plasmids (pCDF-Duet1) were then transformed into *E. coli* using electroporation and the efficiency of transformation (EoT) was calculated. We observed reduced transformation efficiency only for plasmids carrying a protospacer with a 5′-AAG-3′ PAM, indicating that the recombinant Type IV-A system facilitates PAM-dependent interference (Fig. 1f). Base pairing of PAM sequences and crRNA repeat is used to identify self-targets in Type III CRISPR-Cas systems and to prevent DNA cleavage. In our system, the 1-position of the 5′-AAG-3′ PAM can base pair with the Type IV-A crRNA repeat. To obtain insights into the PAM specificity in this interaction, we mutated the 2- and 3-PAM nucleotides in the target plasmid and followed changes in the EoT for the different constructs (Supplementary Table 2). We showed that the presence of 5′-GTG-3′ and 5′-AAN-3′ PAM sequences results in significant EoT reduction (Fig. 1g). However, many other PAM sequences without complementarity to the 5′-crRNA repeat tag prevented anti-plasmid activity, suggesting the presence of a PAM selection mechanism that resembles Type I crRNP-mediated protospacer recognition. Larger EoT error bars correlate with highly variable sizes of obtained colonies and might be an indication of competition between plasmid replication and interference in these assays.

Next, we analysed Cas protein mutants in the recombinant system. The putative helicase DinG is found in all Type IV-A CRISPR-Cas systems, but its function is not known. A multiple sequence alignment of several Type IV-A associated DinG enzymes revealed a conserved Walker-A motif indicative of ATP-binding (Extended Data Fig. 2). EoT assays showed that a null mutation in the walker-A motif of DinG (K136A) abolished interference, suggesting that ATP-dependent helicase activity of DinG is crucial for CRISPR interference (Fig. 1f,h and Supplementary Table 2). These results agree with the requirement for DinG activity in EoT assays using recombinant *P. aeruginosa* crRNPs. The function of the signature protein Csf1 is elusive, but it was suggested that this protein might fulfill the crRNP large subunit’s role of PAM recognition and R-loop stabilization. A conserved cysteine-rich motif of Csf1 was identified and could be part of a zinc-finger structure involved in target recognition. EoT assays highlight the functional importance of this motif, as single mutations in each of the cysteine residues abolished CRISPR interference (Fig. 1b).

As some Type IV CRISPR arrays also contain spacers that match viral targets, we evaluated their phage-targeting potential. Therefore, plasmids coding for a crRNA that targets lambda phage gene *E* next to a 5′-AAG-3′ or a 5′-CGG-3′ PAM were transferred into *E. coli* producing recombinant Type IV-A crRNPs. Cells were infected with a virulent lambda phage (λvir) and plaque formation was quantified. We observed PAM-dependent immunity against lambda phages (Fig. 1f). The deletion of the dinG gene or the introduction of a null mutation (K136A) in *DinG* abolished plaque reduction, indicating that DinG is required for this activity.

Self-targeting activity of recombinant Type IV-A crRNPs

It remains unclear whether the observed activities against phages and plasmids require DNA degradation. An earlier report showed that Type I-C CRISPR-Cas activity can be evaluated using *lacZ* blue-white screening. In this system, white colonies correlate with large genomic deletions around the *lacZ* gene due to the activity of the DNA nuclease Cas3 (ref. ). We used this Type I-C system as a control and designed a comparable set-up to monitor genomic *lacZ* targeting of our Type IV-A system. Plasmids coding for a crRNA with a spacer base pairing with protospacers in either the coding or non-coding strand of *lacZ* were transferred into *E. coli* cells producing Type IV crRNPs (Supplementary Table 2). The presence of a non-targeting crRNA resulted in all blue colonies. For both Type I-C and Type IV CRISPR-Cas systems, the presence of crRNAs with *lacZ* protospacers targets resulted in comparable amounts of white colonies (Fig. 2a and Extended Data Fig. 3). Protospacers in both strands and adjacent to the *lacZ* gene region generated white colonies, providing further support for a DNA targeting mechanism (Extended Data Fig. 4). A *DinG* null mutant (K136A) did not induce white colony formation (Extended Data Fig. 3a). Next, we PCR-amplified the *lacZ* gene of selected white and blue colonies. All blue colonies contained intact *lacZ*. Strikingly, the PCR amplification of 14 white colonies indicated no bands for the Type I-C system and wild-type bands for the Type IV-A system (Fig. 2b,c). We sequenced these PCR amplicons and obtained wild-type sequences, confirming that Type IV-A-mediated CRISPR interference of *lacZ* did not involve DNA degradation or mutation. In agreement, individual blue and white colonies were picked and shown to be able to revert their phenotype when grown on new plates (Extended Data Fig. 3b).

Self-targeting activity of the native system

We aimed to evaluate the biological relevance of such Type IV-A self-targeting mechanism. The native Type IV-A CRISPR array in *P. oleovorans* contains with spacer 1a a perfect match for the host chromosome gene *pilN*, which encodes for a Type IV pilus assembly protein. Therefore, crRNA1 should be self-targeting, which is further supported by the presence of a correct 5′-AAG-3′ PAM sequence next to the *pilN* target (Fig. 1c). To visualize the scanning of self-targets in the *P. oleovorans* genome, we fused the mNeonGreen fluorescent protein to Cas6 in both a wild-type (WT) and a Type IV CRISPR array deletion (ΔCRISPR) strain, allowing us to quantify changes in the crRNP dynamics at the single-molecule level. Diffusion patterns of Cas6-mNeonGreen unveiled...
two molecule populations mainly distributed in the bacterial nucleoid (Fig. 3a and Extended Data Fig. 5a,b). The distribution of molecules over the nucleoid and the fact that most of the confined molecules (dwell events) in the WT strain are located there, suggest the existence of a self-target region (that is, pilN) (Fig. 3b,c). Similarly, among the mobile molecules, there is an increase of molecules with low diffusion rate in the WT strain (60.2%) compared with 24.8% of molecules in the ΔCRISPR strain, which again might be biologically associated with a longer scanning time of the crRNPs attached to the fluorophore in the WT strain (Fig. 3d). SMTracker 2.0 software was used to estimate the number of crRNPs and we observed an average of 26 crRNPs in the cells correlates well with the observed RNA-seq data of mature crRNAs (Fig. 1a).

Type IV-A self-targeting at the pilN gene should influence pilN transcript levels. We used RNA-seq to compare the transcriptomes of wild-type P. oleovorans cells and the ΔCRISPR strain. The effective absence of sequence reads mapping to the CRISPR locus verified its successful deletion in the ΔCRISPR strain. We calculated the changes in transcript abundance and observed the strongest increase for transcripts of pilN and neighbouring genes in the ΔCRISPR strain (Fig. 3e,g). Quantitative PCR with reverse transcription (RT-qPCR) verified significantly increased pilN transcript levels in this strain (Fig. 3f). These results suggest that the native pilN operon is downregulated by crRNA1 of the P. oleovorans Type IV-A system. It is possible that the self-targeting of pilN is a result of neutral evolution that is tolerated in the cell as the target is not degraded. In agreement, we did not observe a clear growth advantage for the ΔCRISPR strain. Alternatively, P. oleovorans might even benefit from a reduction of type IV pilus formation in an environment where its presence is not essential and the bacterium might save energy. Type IV pili are involved in the uptake of DNA (natural transformation and bacteriophage infection) and P. aeruginosa type IV pili were shown to bind DNA. Thus, it is intriguing to speculate that a CRISPR-Cas system that evolved to target foreign DNA might benefit from downregulating a cellular component that is involved in the uptake of such foreign genetic material.

Finally, we introduced synthetic crRNAs into wild-type P. oleovorans cells to analyse the efficiency of gene silencing applications based on endogenous Type IV-A CRISPR-Cas activity. A first target was the gene trpE designed to create a tryptophan auxotroph strain (Fig. 4a). Successful trpE silencing was evidenced by nearly four orders of magnitude lower colony numbers after crRNA-trpE induction in minimal medium without tryptophan (Extended Data Fig. 6). As a second target, a crRNA against gene hmgA was shown to stimulate production of the pigment pyomelanin (Fig. 4b). These phenotypes observed upon production of synthetic crRNAs indicate that they can compete with endogenous crRNAs for incorporation into wild-type Type IV-A effectors and enable sequence-specific inactivation of targeted pathways.

**Discussion**

In conclusion, we showed that the Type IV-A CRISPR-Cas system of P. oleovorans targets DNA in a PAM-dependent manner and can be used to downregulate gene expression in the absence of DNA nuclease activity. The mechanism of this system resembles CRISPRi methodology that uses CRISPR-Cas proteins with inactivated catalytic sites (for example, dead Cas9 (dCas9)) or missing nucleases (Cascade without
**Fig. 3 | Self-targeting Type IV-A CRISPR interference in** P. oleovorans **cells.**

**a.** Single-molecule tracking of crRNPs via Cas6-mNeonGreen in WT and ΔCRISPR array (ΔCRISPR) P. oleovorans strains. Gaussian mixture model (GMM) analyses of frame-to-frame displacements in x- and y-directions were performed: the black line represents the sum of the two Gaussian distributions; dashed red and dotted blue lines represent the single Gaussian distributions corresponding to the static and mobile fractions, respectively. Inset heat maps show the spatial distribution of fluorophores mainly located in the nucleoid of the cell. A darker colour indicates that more tracks with longer scanning time are distributed in that area. **b.** Projection of 98,300 tracks analysed from 60 cells in WT and 92,400 tracks analysed from 62 cells in ΔCRISPR array and represented into a standardized cell model. A trajectory is considered to present confinement (red) when it has at least one dwell event. Molecules changing between confinement and mobility are termed ‘transition’ (mixed behaviour) and are shown in turquoise. Freely diffusive molecules lacking considerable parts of confinement are shown in blue. **c.** Representatives cells from the WT and ΔCRISPR array P. oleovorans strains. Scale bar, 2 µm. Tracks show confinement (red), transition (turquoise) and freely diffusive (blue) molecules. Similar results were observed for the 60 and 62 cells from WT and ΔCRISPR, respectively. **d.** Bubble plots showing fraction sizes (size of the bubble) and diffusion constants (y-axis) in a P. oleovorans strain expressing free-diffused mNeonGreen, and WT and ΔCRISPR array P. oleovorans strains. Step size distributions reveal two populations: a mobile (upper circle) and a static (lower circle) fraction. **e.** Illumina RNA-seq coverage plots of the pilN operon region highlight specific loss of pilN reads for the wild-type strain in comparison with the ΔCRISPR (KO) strain. **f.** RT–qPCR data support upregulation of pilN in the ΔCRISPR array strain. Similar expression patterns are observed in early logarithmic phase (EP) as well as during late logarithmic phase (LP). Experiments were performed with n = 4 biologically independent colonies. Statistical analysis was performed using an unpaired F-test. Data are presented as mean ± s.d., ****P < 0.0001. **g.** Scatterplot of the genes upregulated in the ΔCRISPR array strain showing significant expression of genes in the pil operon including pilN, pilO, pilP and pilQ. A list of all genes with significant P value is provided in Supplementary Table 3.
CRISPRi activity in *P. oleovorans* using engineered crRNAs. Plasmids with a Type IV-A repeat-spacer-repeat construct were conjugated into wild-type *P. oleovorans* cells. The crRNA spacers (Supplementary Table 2) are complementary to the non-coding strand of the indicated genes in the host genomes, and CRISPRi phenotypes were analysed after the induction of crRNA expression with arabinose (ara).

**Methods**

**Strains and growth conditions**

*E. coli* BL21-AI and *P. oleovorans* DSM 1045 cells were cultivated in LB media at 37 °C. Gene expression was induced in *E. coli* BL21-AI by the addition of 1 mM isopropyl-β-D-thiogalactopyranoside (IPTG) and 0.2% arabinose. *E. coli* WM3064 cultures were grown at 37 °C in LB media supplemented with dianimopimelic acid (DAP).

**Illumina RNA sequencing**

*P. oleovorans* DSM 1045 cells were grown until optical density (OD)₆₀₀ = 0.4. To analyse crRNA processing, small RNAs were extracted and enriched with mirVana isolation kit, treated with DNase I (NEB), end-repaired with T4 polynucleotide kinase (NEB), and submitted for library preparation using a NEBNext Ultra RNA library prep kit for Illumina, following the manufacturer’s instructions. For the transcriptomic analysis of the *P. oleovorans* DSM 1045 WT and ΔCRISPR array strains, total RNA was extracted with a mirVana isolation kit (Ambion) from cells exponentially growing at two different phases: early phase (EP) until OD₆₀₀ = 0.3 and late phase (LP) OD₆₀₀ = 1. The extracted RNA was treated with DNase I (NEB), RNA depleted using a NEBNext tRNA depletition kit (Bacteria) and submitted for library preparation using a NEBNext Ultra II directional RNA library prep kit for Illumina, following the manufacturer’s instructions. In both experiments, sequencing was performed using an Illumina Miniseq in single-end mode generating 150 nt reads. Data quality was analysed using FastQC 0.11.9, reads were trimmed with Cutadapt 2.8, and aligned to the genome of *P. oleovorans* DSM 1045 using Hisat2 2.2.1 (refs. 34–36). Data analysis, coverage plots and scatterplots were generated using the R packages ggplot2 3.3.6 and DEseq2 v1.36.0 (refs. 37,38). Mapped reads were visualized and inspected using IGV 2.13.2 (ref. 39).

**Conjugation of *P. oleovorans***

Genetic constructs for genome editions in the *P. oleovorans* strains were delivered via conjugation following an adapted protocol from ref. 40 and using the helper strain *E. coli* WM3064. As this strain is DAP-auxotroph, 0.3 mM DAP was added to the LB media and agar plates, allowing the...
maintenance of the *E. coli* strain during the first step of conjugation. The absence of DAP allowed the elimination of *E. coli* and visualization of only *P. oleovorans* strain during the first step of conjugation. Plates were incubated for 5–7 h at 37 °C. After incubation, cells were collected by adding 2 ml of plain LB medium to the plate and scraping the agar with an inoculation loop. Cells were washed two times with 1 ml of plain LB medium to remove traces of DAP. Finally, cells were resuspended in 1 ml of LB medium, and serial dilutions of 10−2 and 10−3 were plated onto agar plates supplemented with 50 µg ml−1 kanamycin or 30 µg ml−1 gentamicin. The *P. oleovorans* strain before conjugation was considered as a negative control. Plates were incubated at 37 °C for 36 h.

**Generation of *P. oleovorans* gene knock-in and knock-out strains**

Insertions and deletions of specific elements in the *P. oleovorans* genome were carried out following the adapted protocols for endonucleases-mediated recombination. The suicide vector pEMG was used for delivery of genome-specific regions flanking the place of the desired insertion or deletion of parts. The first construct was designed by fusing mNeonGreen fluorescent protein (mNeonGreen FP) to the C-terminal of the native Cas6 protein in the Type IV-A CRISPR-Cas of *P. oleovorans*. In this construct, mNeonGreen gene flanked by -500 bp of homologous region upstream and downstream of the insertion site of interest (that is, between csf5 and csf1) was cloned into the pEMG suicide vector between BamHI and EcoRI restriction sites. For doing multiple insertions or deletions in the same strain, the helper plasmid pSEVA621S was used for digestion of the suicide vector pEMG.

A second construct was designed to further knockout elements of the native system including the CRISPR array. In this case, two fragments of -500 bp of homologous regions upstream and downstream of the CRISPR array were cloned into the pEMG vector between BamHI and EcoRI restriction sites. A list of plasmids used in this work is found in Supplementary Table 4.

**Transformation efficiency assays**

*E. coli* BL21-AI cells were transformed with plasmids enabling production of recombinant Type IV-A crRNPs. A pETDuet-1 plasmid contained all five Type IV-A cas genes. Individual mutants (ΔDinG, DinG K136A, ΔCsfl, Csfl C30A, Csfl C33A, Csfl C66A or Csfl C69A) were created via Quickchange mutagenesis. A second plasmid, pRSFDuet-1 carried a ΔCsf1, Csf1 C30A, Csf1 C33A, Csf1 C66A or Csf1 C69A) were created via Quikchange mutagenesis. A pETDuet-1 plasmid contained all five Type IV-A cas genes. Individual mutants (ΔDinG, DinG K136A, ΔCsfl, Csfl C30A, Csfl C33A, Csfl C66A or Csfl C69A) were created via Quickchange mutagenesis. A second plasmid, pRSFDuet-1 carried a ΔCsf1, Csf1 C30A, Csf1 C33A, Csf1 C66A or Csf1 C69A) were created via Quikchange mutagenesis. A pETDuet-1 plasmid contained all five Type IV-A cas genes. Individual mutants (ΔDinG, DinG K136A, ΔCsfl, Csfl C30A, Csfl C33A, Csfl C66A or Csfl C69A) were created via Quikchange mutagenesis. A second plasmid, pRSFDuet-1 carried a ΔCsf1, Csf1 C30A, Csf1 C33A, Csfl C66A or Csf1 C69A) were created via Quikchange mutagenesis. A pETDuet-1 plasmid contained all five Type IV-A cas genes. Individual mutants (ΔDinG, DinG K136A, ΔCsfl, Csfl C30A, Csfl C33A, Csfl C66A or Csf1 C69A) were created via Quikchange mutagenesis.

A pETDuet-1 plasmid containing 1 mM IPTG, spectinomycin (100 µg ml−1) and ampicillin (100 µg ml−1). Cells and phages were co-incubated at 30 °C for 10 h and plaques were counted.

**Gene silencing assays**

(1) Genomic lacZ targets in *E. coli*: *E. coli* BL21-AI cells producing all Type IV-A Cas proteins expressed in pETDuet-1 were transformed with a pCDFDuet-1 vector containing a minimal CRISPR array with a spacer targeting the coding or non-coding strand of *lacZ* (Supplementary Table 2) or a pCDFDuet-1 vector carrying a random sequence. As control, a pCas3cRh vector encoding a complete Type I-C CRISPR-Cas system and a crRNA targeting lacZ were used as previously detailed. After transformation, cells were transferred onto LB agar plates containing 0.005% X-gal, 0.2% arabinose and 1 mM IPTG. After overnight incubation at 37 °C, images of plates were captured and analysed with OpenCFU 3.8-BETA. For recognition of white colonies, the colour filter was set to a hue angle of 0–80.

(2) Genomic targets in *P. oleovorans*: pHERD30T plasmids with a Type IV-A repeat-spacer-repeat construct under the control of an arabinose promoter were conjugated into wild-type *P. oleovorans* cells as described above. The crRNA spacer sequences are provided in Supplementary Table 2. Cells producing a crRNA targeting trpE (anthranilate synthase component I) were grown overnight in M9 minimal medium (0.6 mM tryptophan was added to the medium. For the plated assay, cells were grown overnight in LB, washed with M9 minimal medium or type IV-A crRNPs and the transformation efficiency was recorded as described above.

**FACS measurements**

*P. oleovorans* cells harboured a pHERD30T vector expressing sfGFP and a pUCP18 vector with a minimal CRISPR array. Different constructs contained spacers targeting the coding strand of *gfp* with a 5′-AAG-3′ PAM or a 5′-CCG-3′ PAM, the non-coding strand of *gfp* with a 5′-AAG-3′ PAM, or a non-targeting crRNA (Supplementary Table 2). Individual colonies were cultivated in LB medium overnight at 37 °C. The cultures were washed twice and diluted 100 times with phosphate-buffered saline solution. Fluorescence intensity measurements were conducted using a BD Fortessa flow cytometer and GFP was excited by the 488 nm laser line. For each sample, 10,000 events were recorded and the unaged average fluorescence intensity of each measurement was recorded. Data were analysed with BD FACS DIVA 8.0.1.

**Bacteriophage plaque assay**

*E. coli* BL21-AI cells producing all Type IV-A Cas proteins were transformed with a pCDFDuet-1 plasmid carrying a minimal CRISPR array, with a spacer targeting the coding or non-coding strand of geneE of the virulent λvir variant of phage lambda. A pCDFDuet-1 vector with a minimal CRISPR array carrying a random spacer sequence was transformed in control experiments. Individual colonies were inoculated in LB medium for overnight incubation and 100 µl of the overnight culture was pre-incubated with 10 µl of lambda phage (titre 1.2 × 10^12 plaque-forming units per ml). After incubation for 10 min, cultures were mixed with 3 ml of selective 0.7% soft LB agar supplemented with 2 mM MgCl2. The mixture was transferred onto a plain LB agar plate containing 1 mM IPTG, spectinomycin (100 µg ml−1) and ampicillin (100 µg ml−1). Cells and phages were co-incubated at 30 °C for 10 h and plaques were counted.

**RT-qPCR of wild-type and ΔCRISPR array *P. oleovorans* strains**

Primer efficiency test (PETest) was performed for the primers designed for pilN and the housekeeping gene *recA* from *P. oleovorans* using a 3-log dilution series of complementary DNA from WT *P. oleovorans*. RT-qPCR was performed on a CFX Connect real-time PCR detection system (Bio-Rad) using IQ SYBR Green Supermix (Bio-Rad). Amplification conditions were as follows: denaturation at 95 °C for 3 min...
followed by 40 cycles of 95 °C for 30 s and 60 °C for 30 s followed by an additional final extension of 55 °C for 5 s and 95 °C/0.5 °C. Three technical replicates were used per dilution, and five serial concentrations were considered for constructing the linear equation and finding the regression coefficients ($R^2$). Adequate primers with a P-E test of 90–110% and $R^2 > 0.7$ were selected (Supplementary Table S5). The reaction mix consisted of IX IQ SYBR Green Supermix (Bio-Rad), 1 µL 1:50 cDNA, 0.5 µM of each primer (0.5 µl of primer mix) and adjustment with nuclease-free water to a final volume of 10 µl. Three technical replicates were used for each biological replicate (that is, four different colonies per treatment) following the RT–qPCR protocol described previously. Gene expression analysis was performed by normalizing the internal control as the cycle threshold (Ct) values of the housekeeping gene recA and calculating the relative transcript level of the pitN gene using the 2$^\Delta\Delta\text{CT}$ method.

Single-molecule microscopy of mNeonGreen-Cas6 tagged Type IV-A crRNPs

To investigate the intracellular dynamics of the Type IV-A crRNPs complex in its native system and in a system lacking the CRISPR array, mNeonGreen FP was fused with the Type IV-A Cas6 protein, which is a stable part of the Type IV-A crRNPs complex. A wild-type P. oleovorans expressing mNeonGreen in pHERD30T plasmid was used as a control (Supplementary Table 4). P. oleovorans cultures were grown until OD$_{600}$ > 0.3 and prepared for single-molecule microscopy following sample preparation, in an Olympus IX71 microscope with a ×100 objective (UAPON 100-OTIRF; numerical aperture, 1.49; oil immersion). In general, this microscopy technique relies on strong excitation of the fluorophores, in this case mNeonGreen, followed by rapid bleaching, which allows tracking of a few unbleached molecules. Image acquisition was performed continuously during laser excitation with the electron-multiplying CCD camera iXon Ultra (Andor Technology). For each movie, 2,500 frames were taken at an acquisition time of 20 ms. Further processing was done with Outfit 1.2 (ref. 38) to set the cell meshes. Track generation was performed with a minimum track length of five steps U-track 2.2.1 (ref. 38). Bleaching curves were analysed in Imaged 2.0 to verify single-molecule observations. Analytical evaluation was carried out using the SMTTracker 2.0 software.

To determine the presence or absence of signals corresponding to single molecules and the number of Cas6–mNeonGreen molecules per cell, the photon count of single bleaching steps was quantified (that is, single-fluorescent-protein bleaching) towards the end of the acquisitions, and the total fluorescence intensity at the beginning of the acquisition was divided by the fluorescence intensity of single fluorophores, relative to cell size. SMTTracker 2.0 was used to automatically determine background signals in individual cells and to subtract these from the specific point spread functions from single molecules. Estimation of the Cas6 protein copy number was based on the single-molecule-tracking pipeline.

Statistics and reproducibility

RNA-seq, RT–qPCR and Type IV-A CRISPR-Cas activity assays (EoT assays, plasmid curing and phage assay) were performed in triplicate ($n = 3$ biologically independent samples; $n = 4$ for RT–qPCR); all attempts at replication were successful. No sample size calculation was performed and no data were excluded during the analyses. The experiments of this study compare different bacterial cells with defined plasmid sequences and genotypes. A selection bias should not affect the results. Statistical analyses and $P$ values for all experiments were obtained using an unpaired two-sided $t$-test (and $F$-test for the RT–qPCR experiment).

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 manuscript or the Extended Data files. Illumina sequence data generated in this study have been deposited in the NCBI Sequence Read Archive database under project ID PRJEB48544. Raw data from single-molecule microscopy analyses are provided at https://doi.org/10.6084/m9.figshare.20359071. Source data are provided with this paper.

References

1. Barrangou, R. et al. CRISPR provides acquired resistance against viruses in prokaryotes. Science 315, 1709–1712 (2007).
2. Marraffini, L. A. & Sontheimer, E. J. Self versus non-self discrimination during CRISPR RNA-directed immunity. Nature 463, 568–571 (2010).
3. Bolotin, A., Quinquis, B., Sorokin, A. & Ehrlich, S. D. Clustered regularly interspaced short palindromic repeats (CRISPRs) have spacers of extrachromosomal origin. Microbiology 151, 2551–2561 (2005).
4. Mojica, F. J. M., Diez-Villasenor, C., Garcia-Martinez, J. & Almendros, C. Short motif sequences determine the targets of the prokaryotic CRISPR defence system. Microbiology 155, 733–740 (2009).
5. Nunez, J. K. et al. Cas1-Cas2 complex formation mediates spacer acquisition during CRISPR-Cas adaptive immunity. Nat. Struct. Mol. Biol. 21, 528–534 (2014).
6. Brouns, S. J. et al. Small CRISPR RNAs guide antiviral defense in prokaryotes. Science 321, 960–964 (2008).
7. Hale, C. R. et al. RNA-guided RNA cleavage by a CRISPR RNA-Cas protein complex. Cell 139, 945–956 (2009).
8. Marraffini, L. A. CRISPR-Cas immunity in prokaryotes. Nature 526, 55–61 (2015).
9. Mojica, F. J., Diez-Villasenor, C., Garcia-Martinez, J. & Soria, E. Intervening sequences of regularly spaced prokaryotic repeats derive from foreign genetic elements. J. Mol. Evol. 60, 174–182 (2005).
10. Pourcel, C., Salvignol, G. & Vergnaud, G. CRISPR elements in Yersinia pestis: a stable part of the Type IV-A crRNP complex in its native system and in a system lacking the CRISPR array, mNeonGreen FP was fused with the Type IV-A Cas6 protein, which is a stable part of the Type IV-A crRNPs complex.
11. Koonin, E. V., Makarova, K. S. & Zhang, F. Diversity, classification and evolution of CRISPR-Cas systems. Curr. Opin. Microbiol. 37, 67–78 (2017).
12. Makarova, K. S. et al. An updated evolutionary classification of CRISPR-Cas systems. Nat. Rev. Microbiol. 13, 722–736 (2015).
13. Pinilla-Redondo, R. et al. Type IV CRISPR-Cas systems are highly diverse and involved in competition between plasmids. Nucleic Acids Res. 48, 2000–2012 (2020).
14. Kamruzzaman, M. & Iredell, J. R. CRISPR-Cas system in antibiotic resistance plasmids in Klebsiella pneumoniae. Front. Microbiol. 10, 2934 (2019).
15. Makarova, K. S. et al. Evolutionary classification of CRISPR-Cas systems: a burst of class 2 and derived variants. Nat. Rev. Microbiol. 18, 67–83 (2020).
16. Özcan, A. et al. Type IV CRISPR RNA processing and effector complex formation in Aromatoleum aromaticum. Nat. Microbiol. 4, 89–96 (2019).
17. Staals, R. H. J. et al. RNA targeting by the type III-A CRISPR-Cas Csm complex of Thermus thermophilus. Mol. Cell 56, 518–530 (2014).
18. Leenay, R. T. et al. Identifying and visualizing functional PAM diversity across CRISPR-Cas systems. Mol. Cell 62, 137–147 (2016).
19. Crowley, V. M. et al. A type IV-A CRISPR-Cas system in Pseudomonas aeruginosa mediates RNA-guided plasmid interference in vivo. CRISPR J. 2, 434–440 (2019).
42. Geissmann, Q. OpenCFU, a new free and open-source software
41. Hmelo, L. R. et al. Precision-engineering the
37. Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold
35. Kim, D., Langmead, B. & Salzberg, S. L. HISAT: a fast spliced aligner
33. Gilbert, L. A. et al. CRISPR-mediated modular RNA-guided
32. Qi, L. S. et al. Repurposing CRISPR as an RNA-guided platform for
31. Mulepati, S. & Bailey, S. Structural and biochemical analysis
29. Oviedo-Bocanegra, L. M., Hinrichs, R., Rotter, D. A. O., Dersch,
28. Csorgo, B. et al. A compact Cascade-Cas3 system for targeted
27. Makarova, K. S. et al. Evolution and classification of the
25. Li, Z., Zhang, H., Xiao, R. & Chang, L. Cryo-EM structure of a
24. Cass, S. D. et al. The role of Cas8 in type I CRISPR interference.
23. Taylor, H. N. et al. Positioning diverse type IV structures and
22. Westra, E. R. et al. Type I-E CRISPR-cas systems discriminate
21. Jia, N. et al. Type III-A CRISPR-Cas Csm complexes: assembly,
20. Altschul, S. F., Gish, W., Miller, W., Myers, E. W. & Lipman, D. J.
19. J. Mol. Biol. 215, 403–410 (1990).
18. Nature Cell Biol. 73, 264–277.e5 (2019).
17. Front. Microbiol. 12, 671522 (2021).
16. Biostat. Rep. 35, e00197 (2015).
15. Cell Res. 30, 179–181 (2020).
14. Nucleic Acids Res. 49, e112 (2021).
13. J. Biol. Chem. 286, 31896–31903 (2011).
12. Cell 152, 1173–1183 (2013).
11. Cell 154, 442–451 (2013).
10. Anal. Biochem. 548, 38–43 (2018).
9. Genome Biol. 15, 550 (2014).
8. M. Cutadapt removes adapter sequences from
7. Nat. Methods 12, 357–360 (2015).
6. Nature Methods 17, 10–12 (2011).
5. EMNinet J. 17, 10–12 (2011).
4. Genomics Bioworks 3, 264–266 (2015).
3. Integrative genomics viewer. Nat. Biotechnol. 29, 24–26 (2011).
2. Microbiol. Biotechnol. 2019, 1–15 (2019).
1. CRISPR and CRISPR-Cas activity assays. J.V.G.-F, L.M.I. and X.G. conceived,
38. J. Biol. Chem. 286, 31896–31903 (2011).
37. PLoS ONE 8, e54072 (2013).
Extended Data Fig. 1 | RNA-seq coverage plots for crRNAs of *P. oleovorans*. Illumina RNA-seq analysis revealed 8 nt long 5′-terminal repeat tags for crRNAs of three CRISPR arrays (Type IV: 5′-GUGAGCGG-3′, Type I-E: 5′-AUGAACCG-3′, Type I-F: 5′-CUCAGAAA-3′).
Extended Data Fig. 2 | Multiple sequence alignments of conserved Csf1 and DinG sections. The ClustalX default color scheme is applied, positions of point mutations investigated in this study are labeled with an asterisk. **a.** Multiple sequence alignment of Type IV-A Csf1, cysteine residues at position 43, 46, 84 and 87 are conserved. **b.** Multiple sequence alignment of Type IV-A associated DinG proteins; a variant walker-A motif with consensus sequence TGXGK is identified.
Extended Data Fig. 3 | Blue white screening of E. coli colonies with a recombinant Type IV-A CRISPR-Cas system targeting lacZ. a. Targeting of the genomic lacZ by the recombinant P. oleovorans type IV-A CRISPR-Cas system with point mutation in DinG (K136A) generates only blue colonies. b. Targeting of the genomic lacZ by the recombinant P. oleovorans Type IV-A CRISPR-Cas system generates a mixture of blue and white colonies (left). Individual white colonies (W1-W3) and blue colonies (B1 – B3) were picked for a repeated round of blue white screening. Reversibility of the phenotype was observed.
Extended Data Fig. 4 | Recombinant Type IV-A CRISPR targeting of different lacZ regions in E. coli. a. Overview of targeted protospacers in the genome of E. coli BL21-AI which include the promoter region (1), untranslated region (2), a region of lacZ gene (3) and a downstream region of lacZ (4). A crRNA without lacZ target served as a negative control (C−). Employed spacer sequences are provided in Extended data Table 2. b. Quantification of the observed percentage of blue colonies in blue-white screening of E. coli BL21-AI cells producing Type IV-A crRNPs with indicated target regions. Experiments were performed in triplicates (n = 3 biologically independent colonies). Data are presented as mean values +/−SD. P-values were calculated using unpaired t-test (*p = 0.0132; **p = 0.0032; ***p < 0.0001).
Extended Data Fig. 5 | Single Molecule Tracking of Cas6-mNeonGreen in *P. oleovorans*. Representative *P. oleovorans* cell from the WT and ΔCRISPR strains stained with DAPI for nucleoid visualization and overlying all projection tracks from the SMM analysis. Scale bar 2 μm. a-b. Standardized cell model containing the projection tracks of a *P. oleovorans* strain expressing free-diffused mNeonGreen. Experiment was repeated twice (n = 2) using a total of 100 cells per experiment. Scale bar 2 μm. c. Distribution density function of the number of detected fluorophores in all cells. d. Distribution density function of 57329 integrated spot intensities. In the best estimation, there are two populations of average integrated intensity $I_1 \sim 29341$ and $I_2 \sim 27901$ u.a., with a proportion of 57%/43%. Therefore, the estimation of the number of fluorophores after accounting for crRNP complex formation and simulation corrections is 26 per cell.
Extended Data Fig. 6 | Gene silencing assay for trpE in *P. oleovorans* plated on minimal medium agar. The control strain is carrying an empty pHERD30T plasmid instead of pHERD30T coding for the crRNA targeting trpE. Experiments were performed in triplicates (n = 3 biologically independent samples) and four dilutions (10⁻² – 10⁻⁵) were plated, respectively. Statistical analysis was performed using an unpaired t-test. Data are presented as mean values +/- SD with a p = 0.0069 (**).
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.

☑️ a) The exact sample size (n) for each experimental group/condition, given as a discrete number and unit of measurement

☐ A statement that 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 possible.

☒ 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 r, Pearson’s r), indicating how they were calculated

Software and code

Policy information about availability of computer code

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

Raw data from RNA-seq analyses in Fig 1A and 3E are available at EGA (https://www.ebi.ac.uk/ena/browser/home) under the code PRJEB48544. Sourcedata from
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**
Experiments were performed in triplicate (n=3), based on similar studies in the field which generated reproducible results. No statistical methods were used to determine the sample size.

**Data exclusions**
No data were excluded.

**Replication**
Experiments (RNA-seq, Type IV-A CRISPR-Cas activity assays) were performed in triplicate, except qRT-PCR which was performed with 4 biologically independent colonies. All attempts at replication were successful.

**Randomization**
The experiments of this study compare different bacterial cells with defined plasmid sequences and genotypes. A selection bias does not affect the results in these studies. Therefore, randomization is not needed in this experimental design.

**Blinding**
Investigators were not blinded as blinding is not applicable in these studies.

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 | Methods |
|----------------------------------|---------|
| n/a | Involved in the study |
| [ ] Antibodies | [ ] ChIP-seq |
| [ ] Eukaryotic cell lines | [x] Flow cytometry |
| [x] Palaeontology and archaeology | [x] MRI-based neuroimaging |
| [x] Animals and other organisms | |
| [x] Human research participants | |
| [x] Clinical data | |
| [x] Dual use research of concern | |

Flow Cytometry

**Plots**

- [ ] The axis labels state the marker and fluorochrome used (e.g. CD4-FITC).
- [ ] The axis scales are clearly visible. Include numbers along axes only for bottom left plot of group (a 'group' is an analysis of identical markers).
- [ ] All plots are contour plots with outliers or pseudocolor plots.
- [x] A numerical value for number of cells or percentage (with statistics) is provided.

**Methodology**

**Sample preparation**
Pseudomonas oleovorans cells producing indicated crRNA variants with gfp protospacer targets were washed twice and diluted 1:100 with fresh phosphate-buffered saline (PBS).

**Instrument**
BD LSRFortessa SORP flow-cytometer

**Software**
BD FACSDiva 8.0.1
| Cell population abundance | For each measurement, 10,000 events were recorded |
|---------------------------|--------------------------------------------------|
| Gating strategy          | A gating strategy was not applied, all events were used to calculate the fluorescence intensity |

- [ ] Tick this box to confirm that a figure exemplifying the gating strategy is provided in the Supplementary information.