Self-Targeting Type IV CRISPR interference in Pseudomonas oleovorans

Xiaohan Guo
Philipps Universität Marburg
Mariana Sanchez-Londono
Philipps Universität Marburg
Vicente Gomes-Filho
Philipps Universität Marburg
Rogelio Hernández-Tamayo
Philipps Universität Marburg https://orcid.org/0000-0003-4666-3929
Selina Rust
Philipps Universität Marburg
Leah Immelmann
Philipps Universität Marburg
Pascal Schäfer
Philipps Universität Marburg
Julia Wiegel
Philipps Universität Marburg
Peter Graumann
LOEWE-Zentrum für Synthetische Mikrobiologie https://orcid.org/0000-0002-8033-5171
Lennart Randau (lennart.randau@staff.uni-marburg.de)
Philipps Universität Marburg

Research Article

Keywords:

Posted Date: March 15th, 2022

DOI: https://doi.org/10.21203/rs.3.rs-1138708/v1

Version of Record: A version of this preprint was published at Nature Microbiology on September 29th, 2022. See the published version at https://doi.org/10.1038/s41564-022-01229-2.

Self-Targeting Type IV CRISPR interference in *Pseudomonas oleovorans*

Xiaohan Guo[1†], Mariana Sanchez-Londono[1†], Vicente Gomes-Filho[1], Rogelio Hernandez-Tamayo[2], Selina Rust[1], Leah M. Immelmann[1], Pascal Schäfer[1], Julia Wiegel[1], Peter L. Graumann[2,3] and Lennart Randau*[1,3]

[1] Department of Biology, Philipps Universität Marburg, Marburg, Germany
[2] Department of Chemistry, Philipps-Universität Marburg, Marburg, Germany
[3] SYNMIKRO, LOEWE Center for Synthetic Microbiology, Marburg, Germany

* To whom correspondence should be addressed. Tel: +49 (0)6421 28 23320; Fax: +49 (0)6421-178 599; Email: lennart.randau@staff.uni-marburg.de

† Joint Authors

Many bacteria contain Class I CRISPR-Cas systems and utilize multi-subunit ribonucleoprotein complexes to interfere with mobile genetic elements. The activities of these complexes are described in detail for Type I effector complexes, showing CRISPR-RNA (crRNA) mediated DNA recognition that (i) relies on the presence of a short DNA sequence termed protospacer adjacent motif (PAM) and (ii) results in the recruitment of the target DNA nuclease Cas3. Type IV CRISPR-Cas systems also belong to Class I, but their substrate requirements have not been defined and a DNA nuclease has not been identified. Here we show that the native *Pseudomonas oleovorans* Type IV-A CRISPR-Cas system targets DNA in PAM-dependent manner and elicits interference in the absence of 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. The deletion of the native Type IV-A CRISPR array resulted in upregulation of *pilN* operon transcription. Reporter gene targeting assays with endogenous and heterologous Type IV-A CRISPR-Cas machineries verified effective CRISPR interference in the absence of DNA nuclease activity. Our results demonstrate that nuclease-free Type IV-A CRISPR-Cas systems can function in host gene regulation. The observed activity resembles CRISPR interference (CRISPRi) methodology using dCas9 or Type I effectors without Cas3. Therefore, Type IV-A CRISPR-Cas activity represents a natural CRISPRi system that is found in many *Pseudomonas* and *Klebsiella* species and allows for their manipulation using synthetic crRNAs.

Clustered Regularly Interspaced Short Palindromic Repeats (CRISPR) and CRISPR-associated (Cas) proteins elicit adaptive immunity in many bacterial species[1,2]. 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[3-5]. Transcription and processing of the CRISPR array results in mature CRISPR RNAs (crRNAs)[6,7] that guide CRISPR ribonucleoprotein complexes (crRNPs) towards

complementary nucleic acid target sequences[3,8-10]. Two Classes and six Types of CRISPR-Cas systems have been classified[11] and Type IV CRISPR-Cas remains as the only Type without description of its endogenous activity. Subtype IV-A CRISPR-Cas was first discovered in the genome of *Acidithiobacillus ferrooxidans*[12] and later shown to contain the signature proteins Csf1 and the crRNP backbone subunits Csf2[13]. Type IV-A systems are usually present on large plasmids (>200 kb)[14] and lack adaptation modules and apparent DNA nucleases like Cas3 or Cas10[13,15]. Identification of plasmid-borne protospacers targeting conjugative elements suggested that Type IV systems are involved in inter-plasmid competition[13]. The characterization of recombinant Cas proteins of the Type IV-A system of *Aromatoleum aromaticum*[16] indicated the presence of crRNPs that resemble Type I effectors[16,17]. Heterologous anti-plasmid interference activity was reported for a Type IV-A system of *Pseudomonas aeruginosa*[16,18]. However, it is not known how its crRNPs can combat plasmids without a target DNA nuclease. In order to pinpoint the native activity of a Type IV-A CRISPR-Cas system, we identified and characterized a complete *cas* gene module (*csf1* (*cas8*-like), *csf2* (*cas7*-family), *csf3* (*cas5*-family), *csf5* (*cas6*-family) and *csf4* (*dinG*)) and a neighboring CRISPR array on a megaplasmid of *Pseudomonas oleovorans*.

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. 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* DSM 1045 were sequenced using Illumina RNA-Seq methodology and sequencing reads were found to map to all three CRISPR arrays (Fig 1A). 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 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 twenty spacers, generating twenty crRNAs with different targeting potential. BLAST search analyses[19] identified seven targets 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 1B). In addition, we identified two targets of Type I-E crRNAs and both protospacers also exhibited a 5′-AAG-3′ PAM (Extended data 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 sfGFP and different versions of a second plasmid coding for crRNAs that target

selected *gfp* protospacers (Extended data Table 2). The GFP signal was quantified using Fluorescence Activated Cell Sorting (FACS) 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 1C). 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 (i) target and (ii) crRNA-production plasmids and cultivated for 12 hours without antibiotic selection (Extended data 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 crRNP can interfere with plasmid replication (Fig 1D).

To analyze this targeting mechanism in greater detail, we designed a heterologous *E. coli* BL21-AI system for the production of recombinant *P. oleovorans* Type IV-A crRNPs. The genes *csf1-5* and a minimal CRISPR array consisting of a single spacer-repeat-spacer unit were provided on two plasmids (Extended data Table 2). Target plasmids 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 and to prevent DNA cleavage[18,20]. 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 (Extended data Table 2). We showed that the presence of 5′-GTG-3′ and 5′-AAN-3′ PAM sequences results in significant EOT reduction (Fig 1F). Other PAM sequences did not support anti-plasmid activity. 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 analyzed Cas protein mutants in the recombinant system. The 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 1). 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 1E and G, Extended data Table 2). These results agree with the requirement for DinG activity in EOT assays using recombinant *P. aeruginosa* crRNPs[21]. The function of the signature protein Csf1 is elusive, but it was suggested that this protein might fulfil the crRNP large subunit's role of PAM recognition and R-loop stabilization[22,23]. 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 1G).

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 geneE 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 lytic lambda phages and plaque formation was quantified. We observed PAM-dependent immunity against lytic lambda phages (Fig 1E). 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 if 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 as a consequence of the activity of the DNA nuclease Cas3[24]. We used this Type I-C system activity 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 (Extended data Table 2). The presence of a non-targeting crRNA resulted in all blue colonies. For both, Type I-C and Type IV-A CRISPR-Cas systems, the presence of crRNAs with *lacZ* protospacer targets resulted in significant and comparable amounts of white colonies (Fig 2A). Protospacers in both strands generated white colonies, providing further support for a DNA targeting mechanism. Next, we PCR-amplified the *lacZ* gene of selected blue and white colonies. All blue colonies contained intact *lacZ*. Strikingly, the PCR amplification of fourteen white colonies indicated no bands for the Type I-C system and wildtype bands for the Type IV-A system (Fig 2B and G). Sequencing of these bands confirmed that only Type IV-A-mediated CRISPR interference of *lacZ* did not involve DNA degradation.

Self-targeting activity of the native *P. oleovorans* Type IV-A CRISPR-Cas 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 1 a perfect match for the host chromosome gene *pilN*, which codes 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. To visualize the scanning of self-targets in the *P. oleovorans* genome, we fused mNeonGreen fluorescent protein to Cas6, allowing us to quantify changes in the crRNP dynamics at the single-molecule level. Diffusion patterns of Cas6-mNeonGreen revealed two molecule populations mainly distributed in the bacterial nucleoid (Fig 3A and B), where most of the dwell events or confinement of molecules location suggest the existence of a target region (Fig 3C and Extended data Fig. 2B). Among the mobile molecules, 83% exposed higher diffusion rates, suggesting a scanning behavior of the crRNPs attached to the fluorophore as compared with the static population with 17% fraction (Extended data Fig 2A), suggesting scanning of specific target, which might represent the self-target gene *pilN*. SMTracker 2.0 software[25] was used to estimate the number of crRNPs and we observed an average of 26 crRNPs in a single *P. oleovorans* cell (Extended data Fig 2C and D). Therefore, the presence of a significant number of Type IV 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 wildtype *P. oleovorans* cells and a Type IV-A CRISPR array deletion 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 neighboring genes in the ΔCRISPR strain (Fig 3D and F). qRT-PCR verified significantly increased *pilN* transcript levels in this strain (Fig 3E). These results highlight 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[26]. Thus, it is intriguing to speculate that a CRISPR-Cas system that evolved to target foreign DNA might benefit from down-regulating 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 analyze the efficiency of CRISPRi 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). As a second target, a crRNA against gene *hmgA* was expected to stimulate production of the pigment pyomelanin (Fig. 4B). We detected clear phenotypes upon induction of synthetic crRNA production, indicating that the introduced crRNAs can compete with endogenous crRNAs for incorporation into wild type Type IV-A effectors and enable sequence-specific inactivation of targeted pathways.

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 down-regulate gene expression in the absence of DNA nuclease activity. The mechanism of this system resembles CRISPRi methodology which uses CRISPR-Cas proteins with inactivated catalytic sites (e.g., dead Cas9 (dCas9)) or missing nucleases (Cascade without Cas3) to stably bind targets and block the transcription machinery. Thus, Type IV-A CRISPR activity at genomic DNA can be viewed as nature's original version of these CRISPRi tools. In return, one can utilize endogenous Type IV-A CRISPR-Cas systems (e.g. found in many pathogenic *Pseudomonas* and *Klebsiella* strains) to conveniently regulate host gene expression by providing synthetic crRNA sequences.

Figure 1

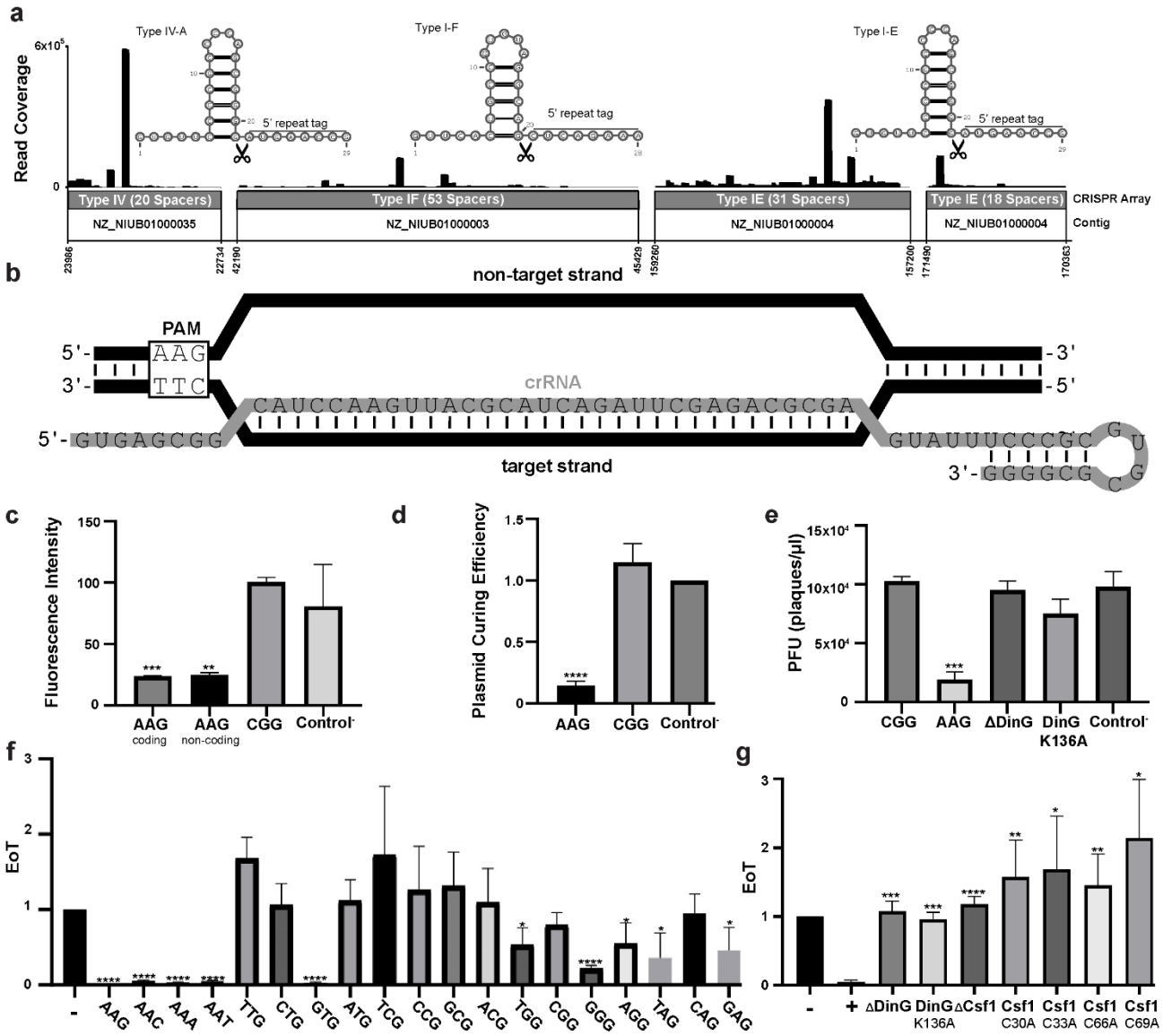

Figure 1. ***In vivo*** **activity of the** ***P. oleovorans*** **Type IV-A CRISPR-Cas system. a)** Coverage plots of small RNA-Seq analysis of three *P. oleovorans* CRISPR-Cas systems highlighting transcription and effective processing of individual crRNAs. High abundance of Type IV-A crRNA8 correlates with a point mutation in repeat 7. Repeat structures and their observed processing sites are indicated. **b)** Schematic representation of Type IV-A R-loop formation of crRNA1 and its genomic *pilN* target. The indicated PAM sequence was observed for all identified protospacer targets of this CRISPR array. **c)** Fluorescence intensity measurement of a *gfp*-containing plasmid targeted by engineered crRNAs in *P. oleovorans*. The target protospacer strand and its PAM are indicated. A crRNA without a protospacer target served as control. **d**) Plasmid curing efficiency was calculated in reference to colony forming units of cells containing non-targeting crRNAs. The protospacer PAM is indicated. **e**) Plaque-forming units (PFU) were identified following lytic lambda phage infection of *E. coli* cells producing recombinant Type IV-A crRNPs. Activities of two crRNAs targeting geneE with 5′-CGG-3′ and 5′-AAG-3′ PAMs are compared and DinG deletion and K136A mutants were assayed

during AAG protospacer targeting. **f**) EOT assays of recombinant Type IV-A crRNPs targeting protospacers on an electroporated plasmid with the indicated PAM sequence in *E. coli*. The AAG-PAM (see b) served as positive control, a crRNA without protospacer match as negative control. **g)** EOT assays were performed with an AAG protospacer and indicated *cas* gene deletions and mutations. **c-g)** All experiments were performed in triplicates. P-Values were calculated using unpaired T-test (* p<0.05, ** p<0.01, *** pp<0.001, **** p<0.0001).

Figure 2

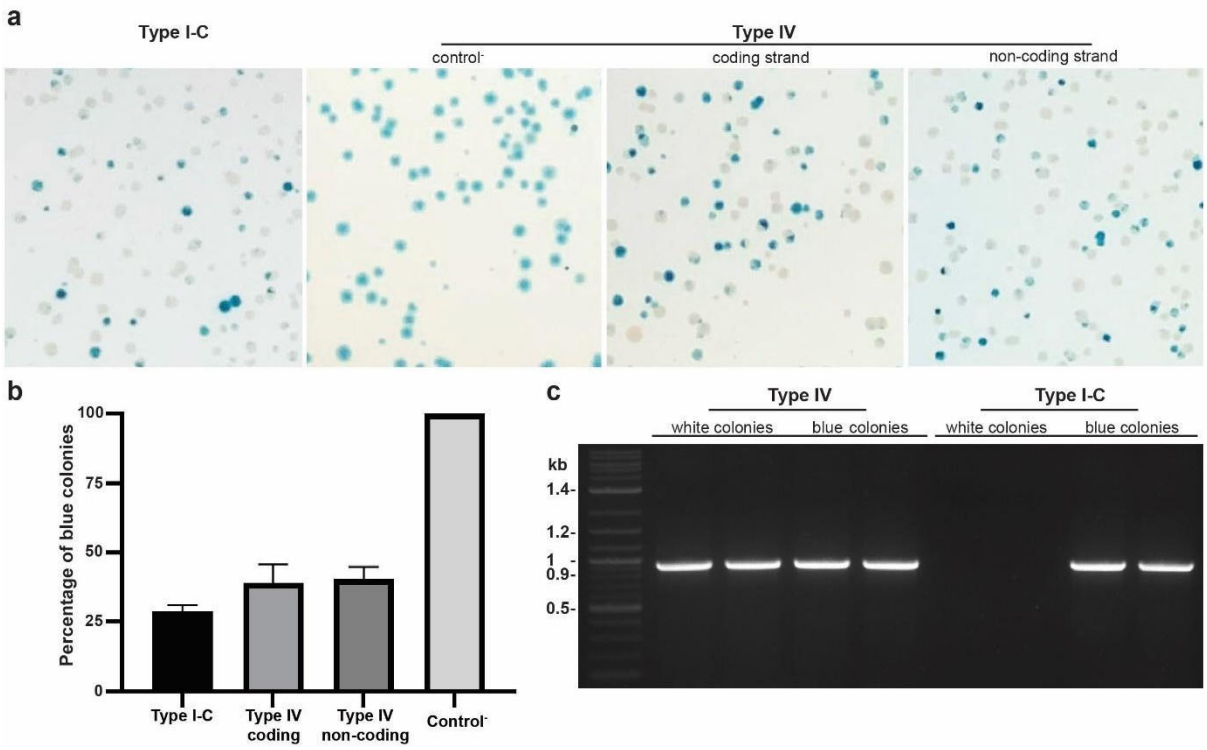

Figure 2. Self-targeting CRISPR interference of recombinant Type IV-A crRNPs in *E. coli* **a**) Representative blue-white screening of *E. coli* BL21-AI cells producing either recombinant Type I-C or Type IV-A crRNPs. Individual crRNAs are targeting *lacZ* and the protospacer strand is indicated. A crRNA without *lacZ* target served as a negative control. **b)** Quantification of the observed percentage of blue colonies for all (n=9852) counted colonies on triplicate plates. **c)** Representative PCR amplification of a 900 bp *lacZ* product. All investigated colonies revealed the absence of *lacZ* following Type I-C CRISPR-Cas genome editing while all white colonies with Type IV-A crRNPs maintained *lacZ*.

Figure 3

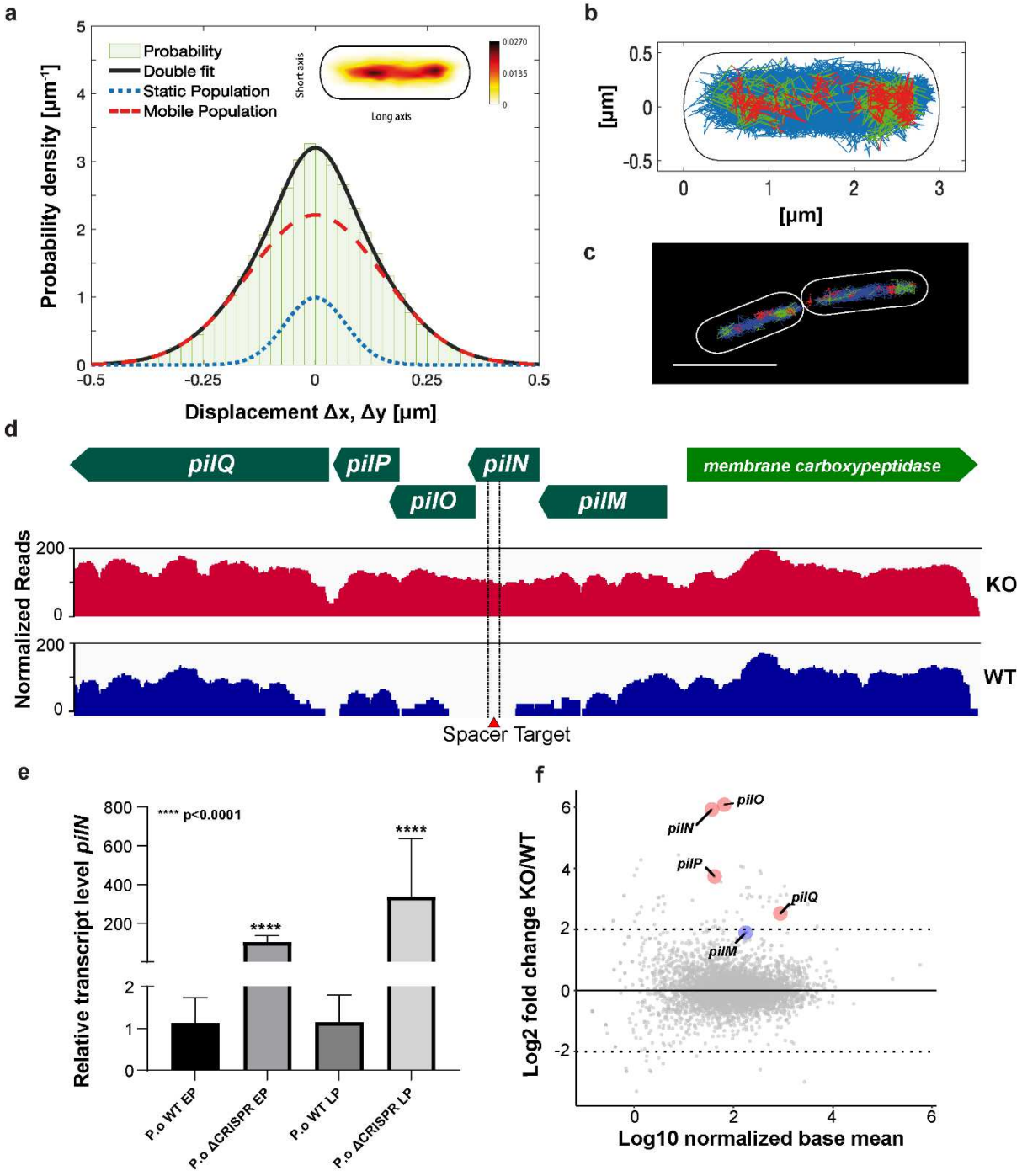

Figure 3. Self-targeting Type IV-A CRISPR interference in *P. oleovorans* cells. a. Single molecule tracking of crRNPs via Cas6-mNeonGreen in wildtype *P. oleovorans*. Gaussian mixture model (GMM) analyses of frame-to-frame displacements in x- and y-directions were performed: the black line represents the sum of the two Gaussian distributions; dotted red and blue lines represent the single Gaussian distributions corresponding to the static and mobile fractions. **b**. Heat map of the spatial distribution of fluorophores mainly located in the nucleoid of the cell. A darker color indicates that more tracks with longer scanning time are distributed in that area **c**. Projection of 89300 tracks analyzed from 56 cells 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 behavior) and are shown in green. Freely diffusive molecules lacking considerable parts of confinement are shown in blue. **d**. Illumina RNA-Seq coverage plots of the *pilN* operon region highlight specific loss of *pilN* reads for the wildtype strain in comparison to the ΔCRISPR (KO) strain. **e**. 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 in triplicates. **f**. Scatter plot of the genes upregulated in the delta CRISR 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 Extended data Table 3.

Figure 4

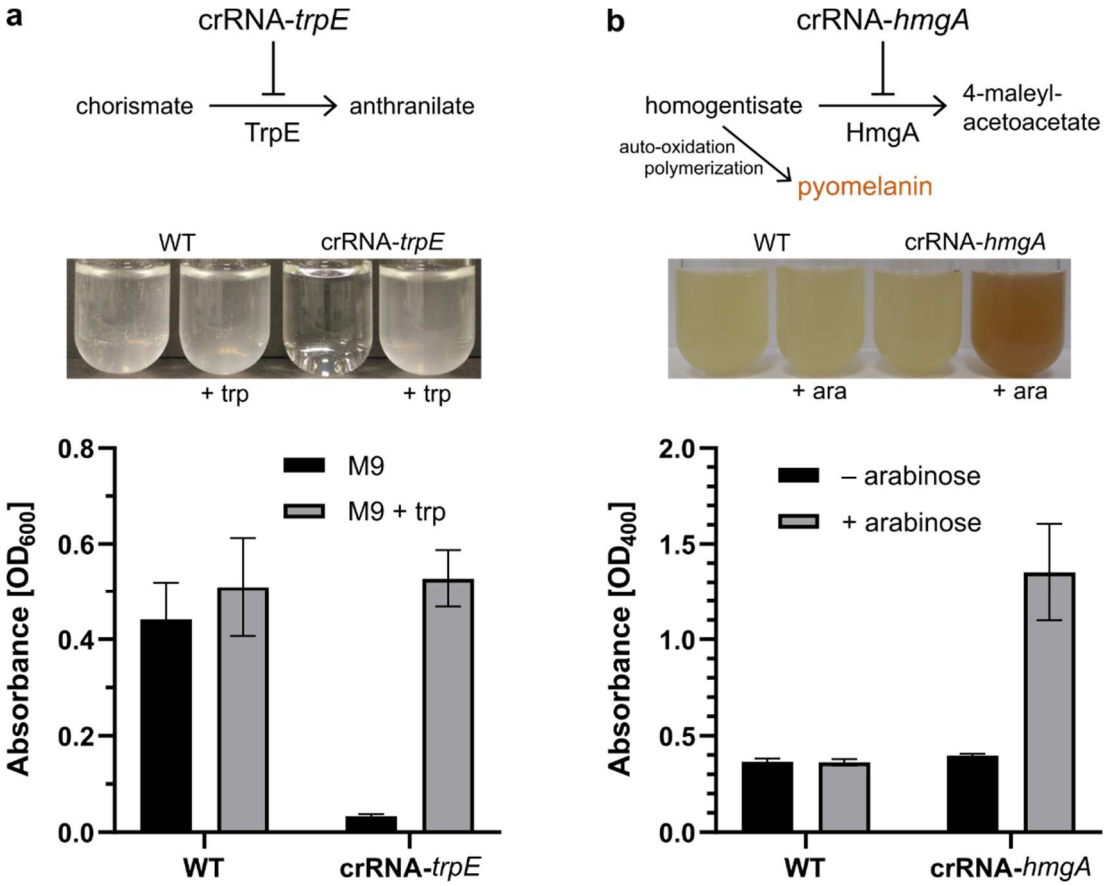

Figure 4. CRISPRi activity in *P. oleovorans* using engineered crRNAs. Plasmids with a Type IV-A repeat-spacer-repeat construct were conjugated into wild type *P. oleovorans* cells. The crRNA spacers (Extended Table 2) are complementary to the non-coding strand of the indicated genes in the host genomes and CRISPRi phenotypes were analyzed after the induction of crRNA expression with arabinose (ara). **a.** Targeting of *trpE* induces tryptophan auxotrophy, which can be complemented by the addition of tryptophan (trp). Cell growth for indicated conditions was quantified for OD_{600} values after 24 hours (bottom) **b.** Targeting of

hmgA induces formation of pyomelanin. Pigment accumulation was quantified for OD_{400} measurements of the supernatant after 48 hours (bottom). All experiments were performed in triplicates.

Materials and 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 IPTG and 0.2 % arabinose. *E. coli* WM3064 cultures were grown at 37°C in LB media supplemented with diaminopimelic acid (DAP).

Illumina RNA Sequencing

P. oleovorans DSM 1045 cells were grown until OD_{600}=0.4. To analyze 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 to 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 delta CRISPR array strains, total RNA was extracted with a mirVana™ isolation kit (Ambion) from cells exponentially growing at two different phases, early phase (EP) until OD_{600}=0.3 and late phase (LP) OD_{600}=1. The extracted RNA was treated with DNase I (NEB), rRNA depleted using a NEBNext® rRNA Depletion Kit (Bacteria) and submitted to 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™ System in single-end mode generating150 nt reads. Data quality was analyzed using FastQC, reads were trimmed with Cutadapt, and aligned to the genome of *P. oleovorans* DSM 1045 using Hisat2[27-29]. Data analysis, coverage plots, and scatters plots were generated using the R packages ggplot2 and DEseq2[30,31]. NCBI Sequence Read Archive database

Conjugation of *P. oleovorans*

Genetic constructs for genome editions in the *P. oleovorans* strains were delivered via conjugation following an adapted protocol from (Wirth et al., 2020)[32] 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* transformed cells. For the first event of conjugation 1 ml and 0.5 ml of a fresh overnight culture of *P. oleovorans* and *E. coli* WM3064, respectively, were harvested separately by centrifugation at 8000 rpm for 3 min at room temperature. Cell pellets were washed twice with LB supplemented with 0.3 mM DAP (LB-DAP) and finally resuspended in a total volume of 100 µl of LB-DAP. The whole suspension was pipetted as a single drop onto a LB-DAP agar plate. Plates were incubated for five to seven hours at 37ºC. After incubation, cells were harvested 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 1ml of plain LB medium to remove traces of DAP. Finally, cells were resuspended in 1 ml of LB medium and serial dilutions of 10^{-1} and 10^{-2} were plated onto agar plates supplemented with 50 μg/ml kanamycin or 30 μg/ml gentamycin. A negative control was considered as the *P. oleovorans* strain before conjugation. Plates were incubated at 37ºC for 36 hours.

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[33,34]. 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 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 ~500bp of homologous region upstream and downstream of the insertion site of interest (i.e., 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 pSEVA6213S 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 are found in Extended data 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 contained all five Type IV-A *cas* genes. Individual mutants (ΔDinG, DinG K136A, ΔCsf1, Csf1 C30A, Csf1 C33A, Csf1 C66A or Csf1 C69A) were created via Quikchange mutagenesis. A second plasmid, pRSFDuet-1 carried a minimal CRISPR array with crRNA1 from *P. oleovorans*. *E. coli* cells producing Type IV-A crRNP variants were then transformed with a target pCDFDuet-1 vector carrying a perfectly complementary protospacer1 against crRNA1 with a 5′-AAG-3′ PAM. A pCDFDuet-1 vector with a protospacer1 and a 5′-CGG-3′ PAM or a non-base-pairing protospacer served as controls. Transformation efficiency was calculated with the formula: Transformation efficiency = CFU (sample) / CFU (non-matching spacer control). To identify functional PAM elements, protospacer1 sequences were synthesized with different 3 nt PAM combinations and cloned into vector pCDFDuet-1. These vectors were transformed into *E. coli* BL21-AI cells containing recombinant wildtype Type IV-A crRNPs and the transformation efficiency was recorded as described above.

FACS measurements

P. oleovorans cells harbored 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′-CGG-3′ PAM, the non-coding strand of *gfp* with a 5′-AAG-3′ PAM, or a non-targeting crRNA (Extended data Table S2). Individual

colonies were cultivated in LB medium overnight at 37°C. The cultures were washed twice and diluted 100x 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 ungated average fluorescence intensity of each measurement was recorded.

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 lambda phage geneE. A pCDFDuet-1 vector with a minimal CRISPR array carrying a random 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 (titer 1.2 x 10^7 PFU/ml). After incubation for 10 minutes, cultures were mixed with 3 ml of selective 0.7 % soft LB agar supplemented with 2 mM $MgCl_2$. The mixture was transferred onto a plain LB agar plate containing 1 mM IPTG, spectinomycin (100 µg/ml) and ampicillin (100 µg/ml). Cells and phages were co-incubated at 30°C for 10 hours and plaques were counted.

CRISPRi assays

(i) Genomic *lacZ*-targets in *E. coli* - *E. coli* BL21-AI cells producing all Type IV-A Cas proteins were transformed with a pCDFDuet-1 vector containing a minimal CRISPR array with a spacer targeting the coding or non-coding strand of *lacZ* (extended data 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 detailed before [24]. After transformation, cells were transferred onto LB agar plates containing 0.005 % X-gal, 0.2 % arabinose and 1 mM of IPTG. After overnight incubation at 37°C, images of plates were captured and analyzed with OpenCFU[35]. For recognition of white colonies, its color filter was set to a hue angle of 0 to 80.

(ii) Genomic targets in *P. oleovorans* – pHERD30T plasmids with a Type IV-A repeat-spacer-repeat construct under control of an araBAD promoter were conjugated into wild type *P. oleovorans* cells as described above. The crRNA spacer sequences are provided in Extended Table 2. Cells producing a crRNA targeting *trpE* (anthranilate synthase component I) were grown for 24 hours in M9 minimal medium (Na_2HPO_4 x $2H_2O$, 5.6 g/l; KH_2PO_4, 3 g/l; NaCl, 0.5 g/l; NH_4Cl, 1 g/l; 2 mM $MgSO_4$; 0.1 mM $CaCl_2$; 0.4 % (w/v) glucose) supplemented with 0.2 % arabinose for crRNA induction and absorbance at OD_{600} was monitored to analyze tryptophan auxotrophy. For complementation, 0.6 mM tryptophan was added to the medium. Cells producing a crRNA targeting *hmgA* (coding for homogentisate 1,2-dioxygenase) were grown in LB medium for 48 hours and absorbance at OD_{400} of the supernatant was monitored to quantify pyomelanin accumulation in the medium.

Quantitative real-time PCR (RT-qPCR)

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 cDNA from *P. oleovorans* WT. RT-qPCR was performed on a CFX Connect™ Real-Time PCR Detection

System (BioRad, Hercules, CA) using iQ™ SYBR® Green Supermix (Bio-Rad, Hercules, CA, USA). Amplification conditions were as follows: denaturation at 95ºC for 3 minutes, followed by 40 cycles of 95 °C for 30 seconds, and 60 °C for 30 seconds, followed by additional final extension of 55ºC for 5 seconds and 95ºC/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 were selected with a PETest between 90-110% and R^2 >0.7 (Extended data Table 5). Reaction mix consisted of 1X of iQ™ SYBR® Green Supermix (Bio-Rad, Hercules, CA, USA), 1 µL of 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 (i.e., four different colonies per treatment) following the RT-qPCR protocol described previously. Gene expression analysis was performed by normalizing the internal control as the Ct values of the housekeeping gene *recA*, and calculating the relative transcript level of *pilN* gene using the $2^{-\Delta\Delta CT}$ method[36].

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

To investigate the intracellular dynamics of the Type IV-A crRNP complex in its native system, mNeonGreen FP was fused with the Type IV-A Cas6 protein, which is a stable part of the Type IV-A crRNP complex[16]. *P. oleovorans* cultures were grown until OD_{600} ~0.3 and prepared for single molecule microscopy following the samples preparation by Hernández-Tamayo et al., 2021 in an Olympus IX71 microscope[37], with a ×100 objective (UAPON 100×OTIRF; numerical aperture [NA], 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 (EMCCD) camera iXon Ultra (Andor Technology, Belfast, UK). For each movie, 2,500 frames were taken with an acquisition time of 20 msec. Further processing was done with Oufti[38], to set the cell meshes. Track generation was performed with a minimum track length of five steps U-track[39]. Bleaching curves were analyzed in ImageJ to verify single-molecule observations. Analytical evaluation was carried out using the SMTracker 2.0 software[25]. To determine the presence or absence of signals corresponding to single molecules and to determine the number of Cas6-mNeonGreen molecules per cell, photon count of single bleaching steps was quantified (i.e., 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. SMTracker 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[25]. Estimation of the Cas6 protein copy number was based on the Single Molecule Tracking pipeline[25].

Data availability

All data are available in the manuscript or the Extended data file. Illumina sequence data generated in this study have been deposited in the NCBI Sequence Read Archive database under project ID PRJEB48544. Raw data from RNA-Seq analysis, FACS, and single molecule microscopy are available upon request from the corresponding author.

Acknowledgements

The authors thank Silvia González Sierra for assistance with flow cytometry, Bálint Csörgő & Joseph Bondy-Denomy for providing plasmid pJW31, Alan Davidson for providing plasmid pHERD30T, Victor de Lorenzo for providing plasmid pEMG and Nathalie Klein for assistance. This work was supported by the DFG-SPP2141 (to X.G. and L.R.) and the Max Planck Society (to M.S-L.).

Contributions

X.G., S.R. and J.W. performed Type IV-A CRISPR-Cas activity assays. V.G. analyzed the RNA-Seq data. M.S-L. performed qRT-PCR analyses. M.S.-L., R.H. and P.G. conceived, performed and analyzed fluorescence microscopy studies. P.S. and L.M.I. performed CRISPRi assays in *E. coli* and *P. oleovorans*, respectively. L.R., X.G. and M.S.-L. conceived the experiments. L.R. wrote the manuscript with support from all other authors.

References

1. Barrangou, R. *et al.* CRISPR provides acquired resistance against viruses in prokaryotes. *Science* **315**, 1709-12 (2007).
2. Barrangou, R. & Marraffini, L.A. CRISPR-Cas systems: Prokaryotes upgrade to adaptive immunity. *Mol Cell* **54**, 234-44 (2014).
3. Bolotin, A., Quinquis, B., Sorokin, A. & Ehrlich, S.D. Clustered regularly interspaced short palindrome repeats (CRISPRs) have spacers of extrachromosomal origin. *Microbiology (Reading)* **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 (Reading)* **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**, 528-34 (2014).
6. Pougach, K. *et al.* Transcription, processing and function of CRISPR cassettes in Escherichia coli. *Mol Microbiol* **77**, 1367-79 (2010).
7. Pul, U. *et al.* Identification and characterization of E. coli CRISPR-cas promoters and their silencing by H-NS. *Mol Microbiol* **75**, 1495-512 (2010).
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**, 174-82 (2005).
10. Pourcel, C., Salvignol, G. & Vergnaud, G. CRISPR elements in Yersinia pestis acquire new repeats by preferential uptake of bacteriophage DNA, and provide additional tools for evolutionary studies. *Microbiology (Reading)* **151**, 653-663 (2005).
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**, 722-36 (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. Ozcan, A. *et al.* Type IV CRISPR RNA processing and effector complex formation in Aromatoleum aromaticum. *Nat Microbiol* **4**, 89-96 (2019).
17. Staals, Raymond H.J. *et al.* RNA Targeting by the Type III-A CRISPR-Cas Csm Complex of Thermus thermophilus. *Molecular Cell* **56**, 518-530 (2014).
18. Leenay, R.T. *et al.* Identifying and Visualizing Functional PAM Diversity across CRISPR-Cas Systems. *Mol Cell* **62**, 137-47 (2016).
19. Altschul, S.F., Gish, W., Miller, W., Myers, E.W. & Lipman, D.J. Basic local alignment search tool. *J Mol Biol* **215**, 403-10 (1990).
20. Jia, N. *et al.* Type III-A CRISPR-Cas Csm Complexes: Assembly, Periodic RNA Cleavage, DNase Activity Regulation, and Autoimmunity. *Mol Cell* **73**, 264-277 e5 (2019).
21. 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).
22. Cass, S.D. *et al.* The role of Cas8 in type I CRISPR interference. *Biosci Rep* **35**(2015).
23. Li, Z., Zhang, H., Xiao, R. & Chang, L. Cryo-EM structure of a type I-F CRISPR RNA guided surveillance complex bound to transposition protein TniQ. *Cell Res* **30**, 179-181 (2020).
24. Csorgo, B. *et al.* A compact Cascade-Cas3 system for targeted genome engineering. *Nat Methods* **17**, 1183-1190 (2020).
25. Oviedo-Bocanegra, L.M., Hinrichs, R., Rotter, Daniel Andreas O., Dersch, S. & Graumann, P.L. Single molecule/particle tracking analysis program SMTracker 2.0 reveals different dynamics of proteins within the RNA degradosome complex in Bacillus subtilis. *Nucleic Acids Research* (2021).
26. van Schaik Erin, J. *et al.* DNA Binding: a Novel Function of Pseudomonas aeruginosa Type IV Pili. *Journal of Bacteriology* **187**, 1455-1464 (2005).
27. Thrash, A., Arick, M., 2nd & Peterson, D.G. Quack: A quality assurance tool for high throughput sequence data. *Anal Biochem* **548**, 38-43 (2018).
28. Kim, D., Langmead, B. & Salzberg, S.L. HISAT: a fast spliced aligner with low memory requirements. *Nat Methods* **12**, 357-60 (2015).
29. Martin, M. Cutadapt removes adapter sequences from high-throughput sequencing reads. *2011* **17**, 3 (2011).
30. Love, M.I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-Seq data with DESeq2. *Genome Biol* **15**, 550 (2014).
31. Wickham, H. *ggplot2: Elegant Graphics for Data Analysis*, (Springer-Verlag New York, 2016).
32. Wirth, N.T., Kozaeva, E. & Nikel, P.I. Accelerated genome engineering of Pseudomonas putida by I-SceI-mediated recombination and CRISPR-Cas9 counterselection. *Microb Biotechnol* **13**, 233-249 (2020).
33. Wirth, N.T., Kozaeva, E. & Nikel, P.I. Accelerated genome engineering of Pseudomonas putida by I-SceI—mediated recombination and CRISPR-Cas9 counterselection. *Microbial Biotechnology* **13**, 233-249 (2020).
34. Hmelo, L.R. *et al.* Precision-engineering the Pseudomonas aeruginosa genome with two-step allelic exchange. *Nature Protocols* **10**, 1820-1841 (2015).
35. Geissmann, Q. OpenCFU, a new free and open-source software to count cell colonies and other circular objects. *PLoS One* **8**, e54072 (2013).
36. Livak, K.J. & Schmittgen, T.D. Analysis of Relative Gene Expression Data Using Real-Time Quantitative PCR and the 2−ΔΔCT Method. *Methods* **25**, 402-408 (2001).
37. Hernández-Tamayo, R., Schmitz, H. & Graumann, P.L. Single-Molecule Dynamics at a Bacterial Replication Fork after Nutritional Downshift or Chemically Induced Block in Replication. *mSphere* **6**, e00948-20 (2021).

38. Paintdakhi, A. *et al.* Oufti: an integrated software package for high-accuracy, high-throughput quantitative microscopy analysis. *Molecular Microbiology* **99**, 767-777 (2016).
39. Jaqaman, K. *et al.* Robust single-particle tracking in live-cell time-lapse sequences. *Nature Methods* **5**, 695-702 (2008).
40 Martinez-Garcia, E. & de Lorenzo, V. Engineering multiple genomic deletions in Gram-negative bacteria: analysis of the multi-resistant antibiotic profile of Pseudomonas putida KT2440. Environ Microbiol **13**, 2702-16 (2011).
41 Brink, K.R. et al. High-throughput discovery of peptide activators of a bacterial sensor kinase. bioRxiv, 2021.06.01.446581 (2021).
42 Wirth, N.T., Kozaeva, E. & Nikel, P.I. Accelerated genome engineering of Pseudomonas putida by I-SceI—mediated recombination and CRISPR-Cas9 counterselection. Microbial Biotechnology **13**, 233-249 (2020).
43 Thompson, J.D., Gibson, T.J., Plewniak, F., Jeanmougin, F. & Higgins, D.G. The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucleic Acids Res **25**, 4876-82 (1997).
44 Qiu, D., Damron, F.H., Mima, T., Schweizer, H.P. & Yu, H.D. PBAD-based shuttle vectors for functional analysis of toxic and highly regulated genes in Pseudomonas and Burkholderia spp. and other bacteria. Appl Environ Microbiol **74**, 7422-6 (2008).
45 Schweizer, H.P. Escherichia-Pseudomonas shuttle vectors derived from pUC18/19. Gene **97**, 109-21 (1991).

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- ExtendeddatatypeIVCRISPRRandauNatMicrobiol.pdf