nature chemical biology

ARTICLES

https://doi.org/10.1038/s41589-022-01137-w

Structural basis of AcrIF24 as an anti-CRISPR protein and transcriptional suppressor

Indranil Arun Mukherjee[1], Clinton Gabel[1], Nicholas Noinaj[1], Joseph Bondy-Denomy[2,3,4] and Leifu Chang[1,5]✉

Anti-CRISPR (Acr) proteins are encoded by phages to inactivate CRISPR–Cas systems of bacteria and archaea and are used to enhance the CRISPR toolbox for genome editing. Here we report the structure and mechanism of AcrIF24, an Acr protein that inhibits the type I-F CRISPR–Cas system from *Pseudomonas aeruginosa*. AcrIF24 is a homodimer that associates with two copies of the surveillance complex (Csy) and prevents the hybridization between CRISPR RNA and target DNA. Furthermore, AcrIF24 functions as an anti-CRISPR-associated (Aca) protein to repress the transcription of the *acrIF23-acrIF24* operon. Alone or in complex with Csy, AcrIF24 is capable of binding to the *acrIF23-acrIF24* promoter DNA with nanomolar affinity. The structure of a Csy–AcrIF24–promoter DNA complex at 2.7 Å reveals the mechanism for transcriptional suppression. Our results reveal that AcrIF24 functions as an Acr-Aca fusion protein, and they extend understanding of the diverse mechanisms used by Acr proteins.

The CRISPR and Cas systems are widespread adaptive immune systems in prokaryotes used to combat invading mobile genetic elements (MGEs), including viruses and plasmids. These systems use RNA-guided effector endonucleases to specifically recognize and cleave foreign DNAs and have been developed as tools for genome editing applications. The class 1 CRISPR–Cas systems (types I, III and IV) employ multi-protein effector complexes to cleave foreign nucleic acids, whereas the class 2 systems (types II, V and VI) use a single multi-domain Cas effector[1,2].

However, MGEs evolved anti-CRISPR (Acr) proteins to inhibit the CRISPR–Cas systems. Since the discovery of the first Acr proteins (AcrIF1–5) in 2013 (ref. [3]), a total of 24 Acr protein families have been reported to inactivate type I-F CRISPR–Cas systems in *Pseudomonas aeruginosa*[3–6]. The type I-F systems contain a surveillance complex (the Csy complex) that is composed of four Cas proteins (Cas5f, Cas6f, Cas7f and Cas8f in a stoichiometry of 1:1:6:1) and a 60-nucleotide (nt) CRISPR RNA (crRNA) that integrates all these protein subunits[7–11]. Target DNA recognition and R-loop formation induce conformational changes in Csy, enabling Cas8f to recruit the Cas2/3 nuclease for degradation of invading MGEs[10]. Many AcrIF proteins inhibit the type I-F system by blocking target DNA recognition by Csy, such as AcrIF1 (refs. [7,8]), AcrIF2 (refs. [7–9]), AcrIF6 (ref. [12]), AcrIF7 (ref. [13]), AcrIF8 (ref. [12]), AcrIF9 (refs. [11,12,14]), AcrIF10 (ref. [8]) and AcrIF14 (refs. [13,15]). AcrIF3 adopts a similar structure to the C-terminal helical bundle of Cas8f (Cas8f^{HB}), competitively binding to the Cas2/3 nuclease and preventing its recruitment to Csy[10,16]. AcrIF11 inactivates Csy by specifically ADP-ribosylating a key residue (N250 of Cas8f), preventing the recognition of the protospacer adjacent motif (PAM) of target DNA[17]. AcrIF4 binds to the Cas8f^{HB} and precludes the conformational changes required for activation of Csy[13]; it is also shown to inhibit DNA binding in vivo[18].

AcrIF24 was recently discovered from a *P. aeruginosa* prophage and is the largest AcrIF protein (228 amino acids) to date[6]. Another *acr* gene (*acrIF23*) is found upstream of *acrIF24*, making an *acrIF23-acrIF24* operon that is activated by a single promoter. Usually, *acr* genes are found upstream of a conserved anti-CRISPR-associated (*aca*) gene, which encodes an Aca protein that functions as a repressor of the *acr-aca* operon[19,20]. Without Aca activity, the strong transcription driven by an acr-associated promoter is lethal to a phage[19]. However, the *acrIF23-acrIF24* operon is not associated with an *aca* gene (Extended Data Fig. 1a). Interestingly, the C-terminus of AcrIF24 contains an HTH domain that is conserved in Aca proteins, indicative of a role also as an Aca protein[6].

In this study, we elucidated the mechanisms for the dual function of AcrIF24, combining cryogenic electron microscopy (cryo-EM) and biochemical analysis. Our results reveal that AcrIF24 is a dimer that binds directly to Csy and prevents the hybridization between target DNA and crRNA. Furthermore, AcrIF24 binds tightly to the *acrIF23-acrIF24* promoter DNA and is capable of acting as a transcriptional repressor.

Results

AcrIF24 directly binds to Csy, whereas AcrIF23 does not. We first tested the function of AcrIF24 and AcrIF23 using an in vitro DNA cleavage assay[17]. AcrIF24 can fully inhibit substrate DNA cleavage by Cas2/3 at a 1:1 molar ratio between AcrIF24 and Csy (Extended Data Fig. 1b,c). In comparison, AcrIF23 shows full inhibition at a much lower molar ratio (Extended Data Fig. 1b,c). Electrophoresis mobility shift assay (EMSA) results suggest that AcrIF24 blocks both the formation of the Csy–DNA complex and recruitment of Cas2/3 nuclease, whereas AcrIF23 does not affect either DNA or Cas2/3 recruitment (Extended Data Fig. 1d–f). Size-exclusion chromatography also showed that AcrIF24 directly binds to Csy (Extended Data Fig. 2a).

Overall structure of Csy–AcrIF24. To understand the inhibition mechanism of AcrIF24, we determined the structure of the

[1]Department of Biological Sciences, Purdue University, West Lafayette, IN, USA. [2]Department of Microbiology and Immunology, University of California, San Francisco, San Francisco, CA, USA. [3]Quantitative Biosciences Institute, University of California, San Francisco, San Francisco, CA, USA. [4]Innovative Genomics Institute, Berkeley, CA, USA. [5]Purdue University Center for Cancer Research, Purdue University, West Lafayette, IN, USA. ✉e-mail: lchang18@purdue.edu

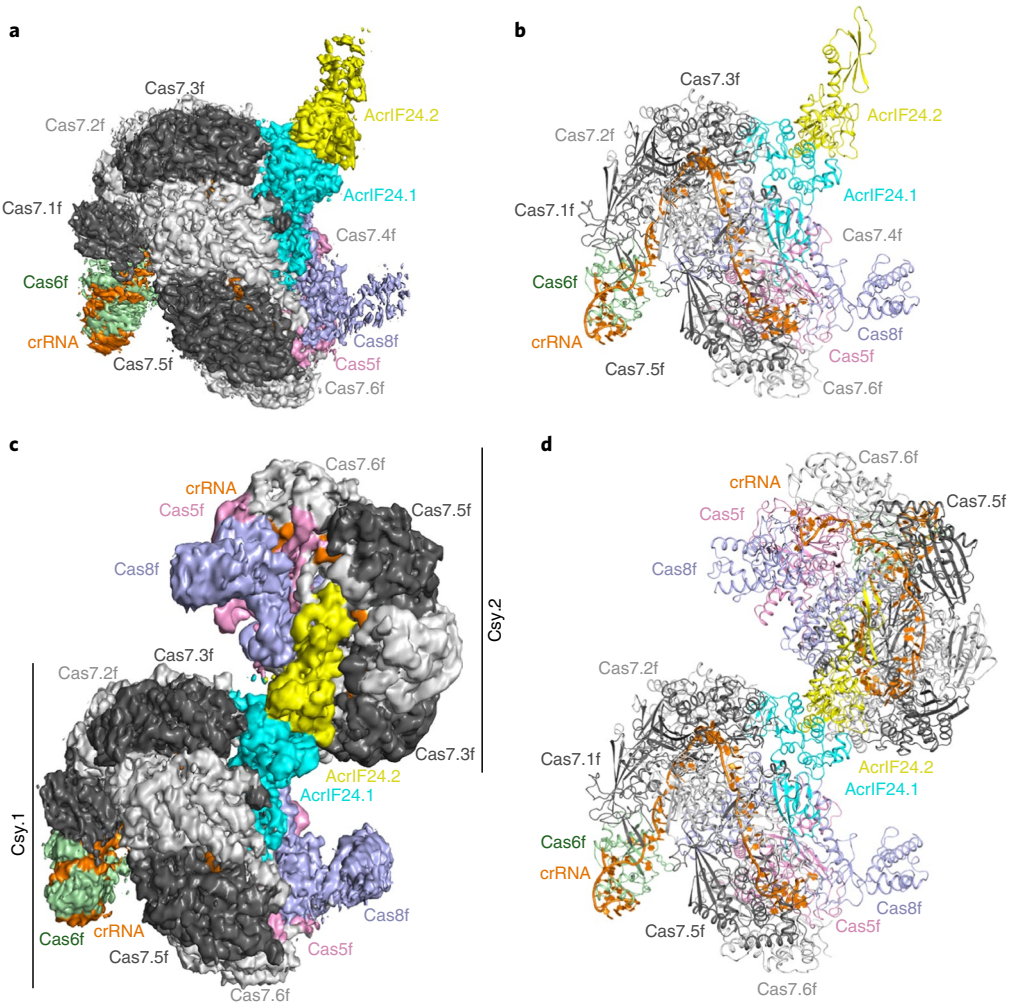


Fig. 1 | Overall structures of Csy-AcrIF24 complexes. a, Cryo-EM map of Csy-AcrIF24 in surface representation with each subunit color-coded. Two subunits in the AcrIF24 dimer are shown in cyan and yellow, respectively. **b**, Atomic model of Csy-AcrIF24 in cartoon representation. **c**, Cryo-EM map of Csy-AcrIF24 dimer in surface representation. **d**, Atomic model of Csy-AcrIF24 dimer in cartoon representation.

Csy–AcrIF24 complex using single-particle cryo-EM (Fig. 1, Extended Data Fig. 3 and Supplementary Table 1). We observed two populations of the Csy–AcrIF24 complex: AcrIF24 dimer bound to one copy of Csy (Csy–AcrIF24; ~91% of total particles) and two copies of Csy (Csy–AcrIF24 dimer; ~9% of total particles), determined at 3.2 Å and 3.5 Å resolutions, respectively (Extended Data Fig. 3). The cryo-EM maps were of sufficient quality for building the atomic model of AcrIF24 de novo and rebuilding the components in Csy (Protein Data Bank (PDB): 7JZW)[13]. Csy does not show major conformational changes upon binding to AcrIF24 (PDB: 6B45)[8]. In brief, Csy is composed of a Cas6f 'head', a Cas8f–Cas5f 'tail' and a 'backbone' of six copies of Cas7f, integrated by crRNA (Fig. 1a,b). An AcrIF24 dimer binds to the middle region of the Cas7f 'backbone', with direct contacts to multiple Cas7 subunits. The AcrIF24 dimer is capable of dimerizing Csy in a 'back-to-back manner', placing the Cas6f 'heads' at both ends of the assembly (Fig. 1c,d).

Structure of AcrIF24. The AcrIF24 monomer contains three domains including an N-terminal domain (NTD, amino acids 1–71), a middle domain (MID, amino acids 72–138) and a C-terminal HTH domain (HTH, amino acids 139–228) (Fig. 2a and Extended Data Fig. 4a,b). The NTD is composed of an anti-parallel β-sheet capped by a helix in a β1β2β3β4α1 topology (Fig. 2a). A DALI search reveals that the NTD is structurally similar to an auxiliary protein of methane monooxygenase hydroxylase (MMOD)[21] (PDB: 6D7K; Z-score 6.5) (Extended Data Fig. 4c). The MID domain contains a beta-sheet (β6β5β7β8) followed by α2, with a long loop between β6 and β7 (amino acids 91–109). However, no structural homology of the MID domain was found in a DALI search. HTH is composed of five short helices (α3–α7) and contacts the NTD through hydrophobic interactions (Extended Data Fig. 4d). A DALI search reveals that this domain shares a similar structure with anti-CRISPR-associated protein 1 (Aca1) (Z-score 8.0; PDB: 7VJN) (Extended Data Fig. 4e)[22]. In vitro DNA cleavage and EMSA assays with domain deletion mutants suggest that the MID and HTH domains are necessary, and that the NTD or HTH are not sufficient for inhibition of DNA cleavage and association between AcrIF24 and Csy (Fig. 2b,c and Extended Data Figs. 2b and 5a,b).

The cryo-EM structures show that AcrIF24 exists as a dimer in the Csy–AcrIF24 complex. To study the stoichiometry of only AcrIF24 in solution, we used in-line size-exclusion chromatography small-angle X-ray scattering (SEC-SAXS). AcrIF24 (theoretical mass 25 kDa) eluted as a single peak from SEC with the SAXS analysis showing that AcrIF24 is a dimer with an experimentally estimated molecular weight of 56 kDa (Extended Data Fig. 6a,b). Furthermore, R_g and D_{max} values are consistent with a dimer for AcrIF24 (Extended Data Fig. 6c). A cryo-EM reconstruction of

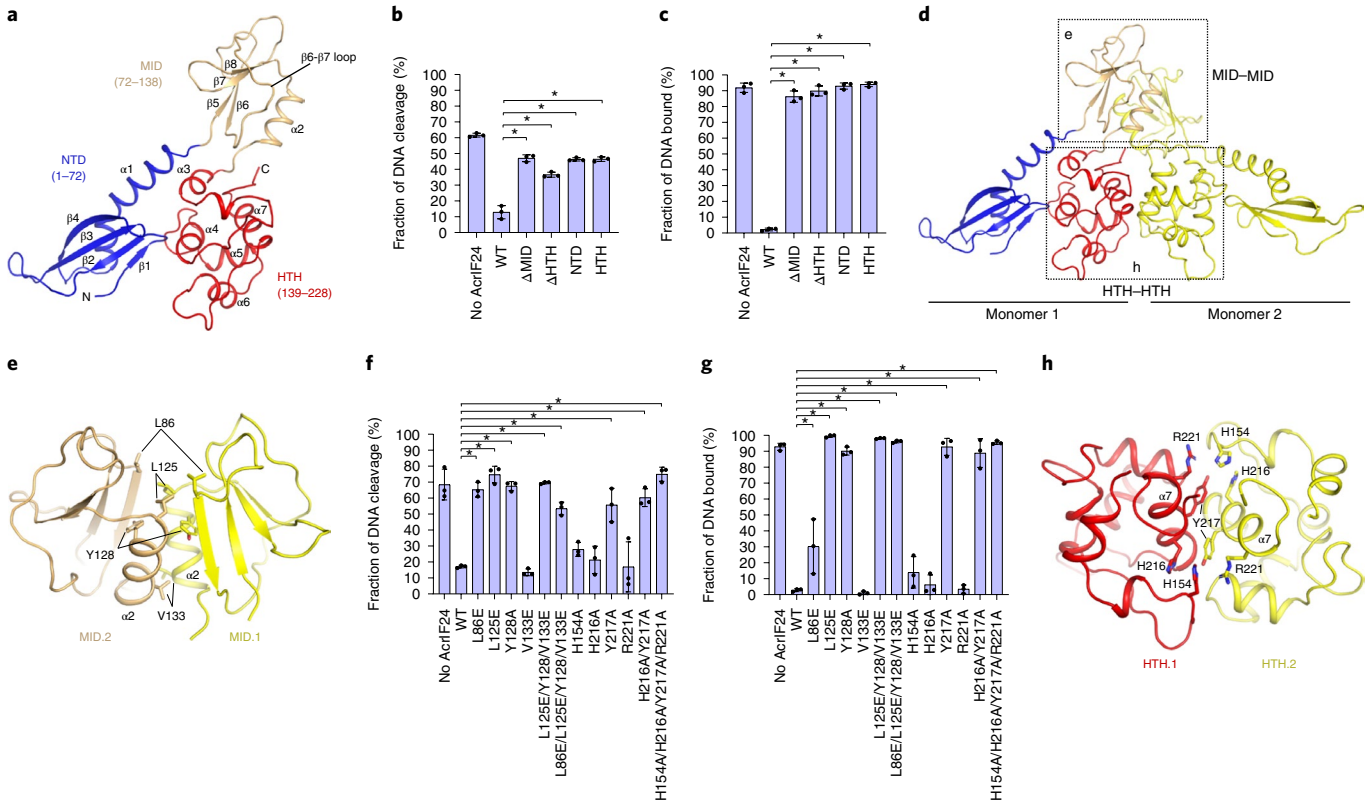


Fig. 2 | Structure of AcrIF24. a, Atomic structure of a monomeric subunit of AcrIF24 in cartoon representation with each domain color-coded. **b**, Quantification of in vitro DNA cleavage assay in the presence of wild-type and domain deletion mutants of AcrIF24. **c**, Quantification of EMSA results showing binding between Csy and target DNA in the presence of wild-type and domain deletion mutants of AcrIF24. **d**, AcrIF24 dimer in cartoon representation. Boxed regions are the dimeric interfaces detailed in **e** and **h**. **e**, MID–MID interaction showing the interacting residues. **f**, Quantification of in vitro DNA cleavage assay using wild-type AcrIF24 and AcrIF24 with mutations of residues involved in the dimerization interfaces. **g**, Quantification of EMSA assay showing the effects of AcrIF24 variants in inhibiting the formation of the Csy–DNA complex. **h**, HTH–HTH interaction with residues in the interface highlighted. The in vitro DNA cleavage assay and EMSA experiments were repeated three times with similar results, and representative raw data are shown in Extended Data Fig. 5. Error bars represent s.d.; $n = 3$. Two-sided t-test was performed ($^*P < 0.05$).

AcrIF24 also supports that AcrIF24 in solution is a dimer (Extended Data Fig. 3e). We also performed SEC-SAXS analysis for AcrIF23 and found that AcrIF23 exists in a monomeric state in solution (Extended Data Fig. 6d–f).

The dimerization interface of AcrIF24 lies at the MID–MID and the HTH–HTH interfaces of the two monomers with a buried surface area of 1,132 Å^2 (Fig. 2d). The MID–MID interface (buried surface 485 Å^2) is mediated primarily by hydrophobic interactions between helix α2 from two protomers. Specifically, sidechains of L86, L125 and Y128 form a hydrophobic patch at the N-terminal end of the helices (Fig. 2e and Extended Data Fig. 4f). Substitution of any of these residues to alanine or glutamic acid significantly alleviates AcrIF24-mediated inhibition of substrate cleavage in the in vitro cleavage assay (Fig. 2f and Extended Data Figs. 2c and 5c). V133 at the C-terminal end of α2 likely augments the interaction, but substitution to glutamic acid showed no detectable effect in the in vitro cleavage assay (Fig. 2f). EMSA results suggest that wild-type AcrIF24 prevents the formation of the Csy–DNA complex, but mutant AcrIF24 (for example, L86E, L125E and Y128A) did so less efficiently (Fig. 2g and Extended Data Fig. 5d). Taken together, the hydrophobic patch at the N-terminal end of α2 is critical for the function of AcrIF24.

The HTH–HTH dimerization interface (buried surface 624 Å^2) is primarily stabilized by interactions between helix α7 of both AcrIF24s (Fig. 2h and Extended Data Fig. 4g). Y217, located in the middle of the helix, is critical for the Acr function of AcrIF24. Alanine substitution of Y217 almost completely abolished the AcrIF24-mediated inhibition in in vitro cleavage assays (Fig. 2f and Extended Data Figs. 2c and 5c). A quadruple mutant containing alanine substitutions of Y217 and bulky residues flanking Y217 (H154A/H216A/Y217A/R221A) showed no detectable inhibition mediated by AcrIF24 (Fig. 2f and Extended Data Figs. 2c and 5c). The SEC-SAXS and MALS data suggest that this quadruple mutant fails to form a dimer (Extended Data Fig. 6g,h). In the presence of this quadruple mutant, DNA binding to Csy is restored in the EMSA assay (Fig. 2g and Extended Data Fig. 5d). Together, these results illustrate that AcrIF24 functions as a dimer and prevents target DNA binding to Csy.

Interactions between AcrIF24 and Csy. Both the NTD and MID domains directly contact the Cas7 'backbone' of Csy, and the interactions would block the hybridization between the target DNA and crRNA (Fig. 3a,b). The NTD of AcrIF24 primarily engages with the thumb domain of Cas7.6f and a loop of Cas7.5f (residues 236–276) through main chain interactions (Fig. 3c and Extended Data Fig. 4h). The β2-β3 loop (16-RGLHYGGLP-24) forms three main chain hydrogen bonds with the thumb domain of Cas7.6f (Fig. 3c). Replacement of the loop with a GSGS linker significantly reduced the AcrIF24 inhibition in the in vitro cleavage assay (Fig. 3d and Extended Data Figs. 2c and 5e). In addition, β4 forms a pair of hydrogen bonds with a loop (residues 236–276) of the palm domain of Cas7.5f (Fig. 3c).

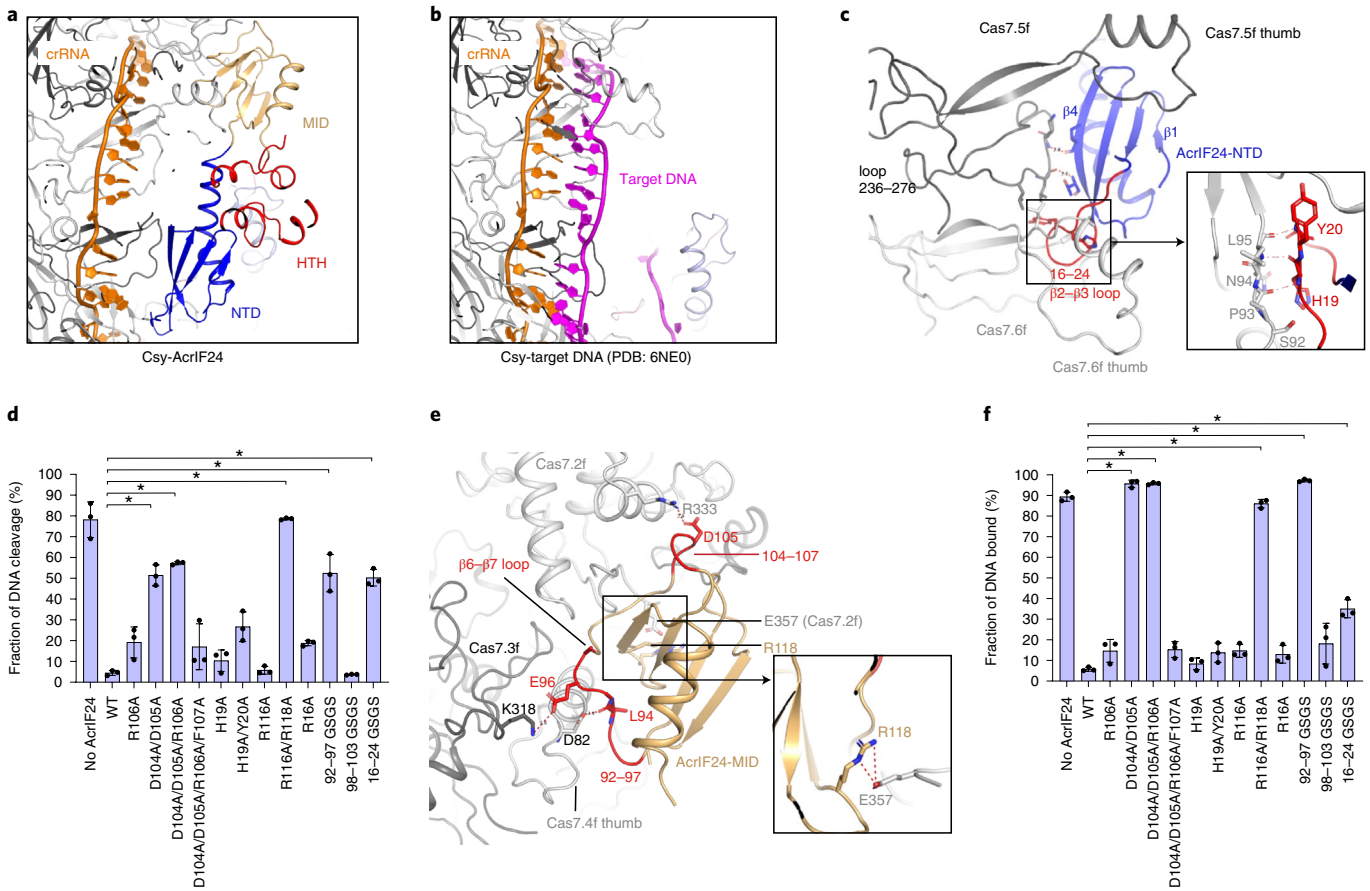


Fig. 3 | Interaction between AcrIF24 and Csy. a, Overview of binding between AcrIF24 and Csy. AcrIF24 binds to the Cas7 'backbone' of Csy. **b**, Structure of Csy–DNA (PDB: 6NE0) shown in the same view of Csy as in **a**. **c**, Detailed interaction between AcrIF24-NTD and Csy. The β2-β3 loop interacting with Csy is colored in red. Specific interactions between the 16–24 region of AcrIF24 and the Cas7.6f thumb domain are detailed in the box on the right. **d**, Quantification of in vitro DNA cleavage assay in the presence of wild-type AcrIF24 and AcrIF24 with mutations of residues involved in Csy interaction shown in **c** and **e**. **e**, Detailed interaction between AcrIF24-MID and Csy. The long β6-β7 loop region interacting with Csy is colored in red. Interactions between R118 of AcrIF24 and the Cas7.2f thumb domain are detailed in the box on the right. **f**, Quantification of EMSA assay showing the effects of wild-type and mutant AcrIF24 variants in inhibiting the formation of Csy-DNA. The in vitro DNA cleavage assay and EMSA experiments were repeated three times with similar results, and representative raw data are shown in Extended Data Fig. 5. Error bars represent s.d.; $n = 3$. Two-sided *t*-test was performed ($^*P < 0.05$).

The MID domain of AcrIF24 interacts with the palm domains of Cas7.2f and Cas7.3f and the thumb domain of Cas7.4f mainly through the long β6-β7 loop (Fig. 3e and Extended Data Fig. 4i). Specifically, the N-terminus of the loop (92-SRLPEA-97) interacts with Cas7.3f and Cas7.4f, with the C=O group of L94 forming a hydrogen bond with the sidechain of D82 and E96 establishing a salt bridge with K318, respectively. Replacement of the 92–97 loop with a GSGS linker reduced AcrIF24 function in the in vitro cleavage assay (Fig. 3d and Extended Data Figs. 2c and 5e). The C-terminus of the loop (104–107) engages Cas7.2f, with a salt bridge formed between D105 of the β5-β6 loop and R333 of Cas7.2f. Alanine substitution of the 104–107 region caused impaired inhibition capacity of AcrIF24 (Fig. 3d and Extended Data Figs. 2c and 5e). In addition to the long β6-β7 loop, the β7-β8 loop also contacts Cas7.2f, with R118 of the loop forming salt bridges with E357 of Cas7.2f; alanine substitution also caused impaired inhibition capabilities of AcrIF24 (Fig. 3d,e and Extended Data Figs. 2c and 5e). EMSA data using wild-type and mutant variants of AcrIF24 suggest that AcrIF24 inhibits Csy by preventing the formation of the Csy–DNA complex (Fig. 3f and Extended Data Fig. 5f).

Csy–AcrIF24 binds DNA independent of the spacer sequence. Our EMSA data show that, in the presence of wild-type AcrIF24, although the formation of the Csy–DNA complex is prevented, a new band at a higher molecular weight is observed (Extended Data Fig. 1d), indicative of the formation of a Csy–AcrIF24–DNA complex. Consistent with this finding, we were able to purify a Csy–AcrIF24–dsDNA complex by incubating the Csy–AcrIF24 complex with a double-stranded DNA (dsDNA) substrate followed by SEC (Extended Data Fig. 2a). The bound dsDNA was only visible in the dimer form and was determined at a lower resolution (8–10 Å) due to low abundance of this species, flexibility or both in this region (Extended Data Fig. 7 and Extended Data Fig. 8a). Nevertheless, the major and minor grooves of the DNA are visible, allowing us to model the bound DNA as a poly(G) sequence.

The structure shows that the two ends of the dsDNA are bound to both the PAM recognition sites of Csy complexes, and the middle region of the dsDNA (~20 base pairs (bp)) is bound to the AcrIF24 dimer, with the connections between the middle region and the two ends barely visible (Extended Data Fig. 8a). The middle region of the dsDNA is bound to the HTH domains of the AcrIF24 dimer, which form a highly positively charged surface (Extended Data

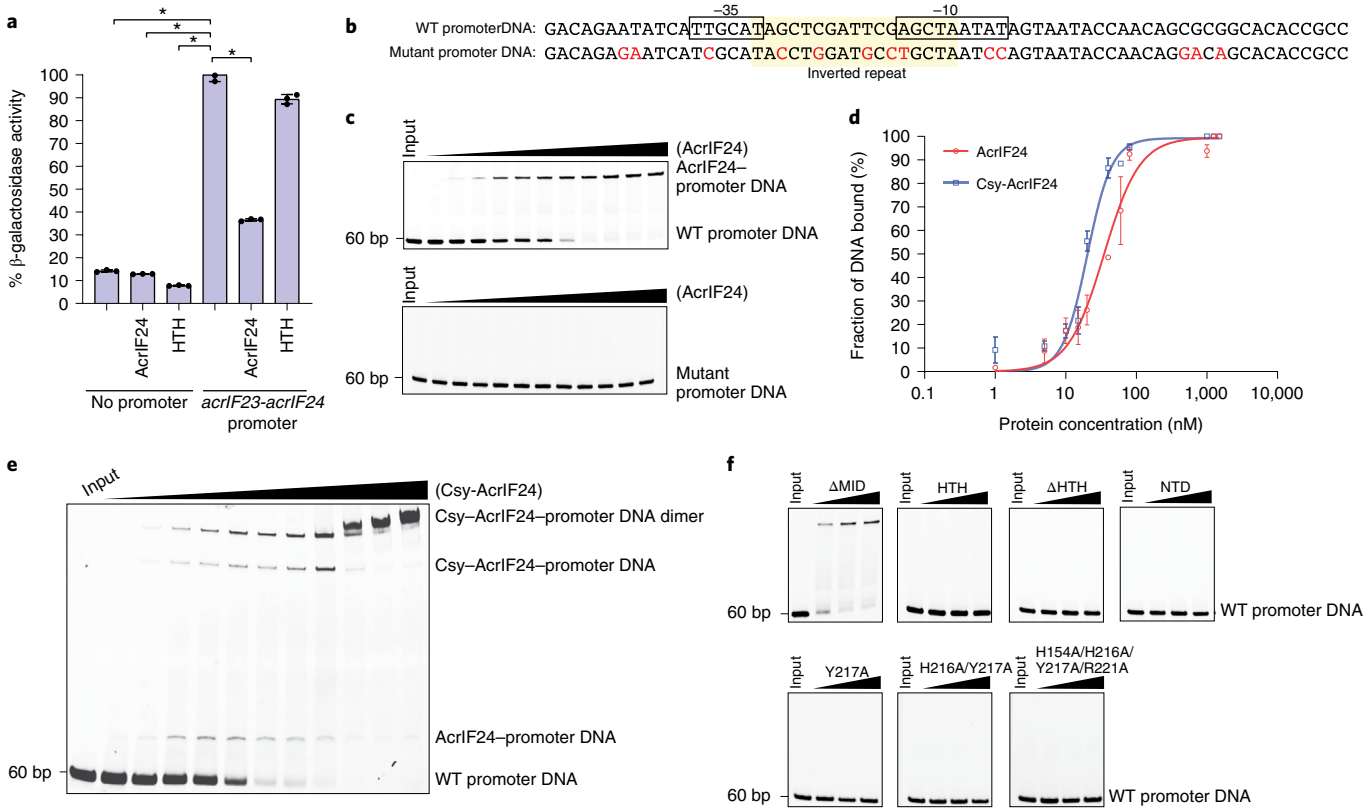


Fig. 4 | AcrIF24 and Csy–AcrIF24 tightly bind to Acr promoter DNA. a, β-galactosidase activity assay showing transcriptional repression by AcrIF24. The activity of β-galactosidase in the presence of an *acrIF23-acrIF24* promoter is calibrated to 100%. The activities of β-galactosidase are shown as mean ± s.d.; $n = 3$. Two-sided *t*-test was performed (*$P < 0.05$). **b**, Sequence of the 66-bp promoter DNA and a mutant version of it where IR sequences are interrupted. **c**, EMSA results showing that AcrIF24 binds to the promoter DNA but not to the mutant version. **d**, Promoter DNA binding curves of AcrIF24 and Csy–AcrIF24. The fraction of DNA bound is determined from EMSA results in **c** and **e**. Calculated K_D values of AcrIF24 and Csy–AcrIF24 to promoter DNA are ~34 nM and ~20 nM, respectively. Data are presented as mean values ± s.d.; $n = 3$. **e**, EMSA results showing that Csy–AcrIF24 binds to the promoter DNA. **f**, EMSA results showing the effect of AcrIF24 mutations on promoter DNA binding.

Fig. 8b). The two ends of the dsDNA are loaded into the PAM recognition sites, similarly to the previously studied Csy–dsDNA complex (PDB: 6NE0)[23]. Based on our modeling, the total length of dsDNA spanning the two PAM recognition sites would be around 60–62 bp.

Because no crRNA–target DNA hybridization is observed in the Csy–AcrIF24–DNA complex, we then checked whether Csy–AcrIF24 is capable of binding non-specific DNA. Our EMSA results suggest that Csy–AcrIF24 binds to both specific and non-specific DNA (Extended Data Fig. 8c). Mutations in the PAM recognition site in Csy (K247E/N250D in Cas8f (ref. [15])) show no detectable binding to either specific or non-specific DNA, suggesting that DNA binding to Csy–AcrIF24 relies on this recognition site (Extended Data Fig. 8d). To check whether the specific and non-specific DNA share the same binding site on the Csy–AcrIF24 complex, we performed competition EMSA experiments. Our results show that non-specific DNA can replace specific DNA in the Csy–AcrIF24–DNA complex and vice versa (Extended Data Fig. 8e,f), suggesting that they share the same binding site. Taken together, Csy–AcrIF24 is capable of recruiting DNA independent of the PAM and spacer sequences, possibly through interaction with the charged surfaces and the PAM recognition sites.

Both AcrIF24 and Csy–AcrIF24 bind tightly to promoter DNA. We next explored the function of AcrIF24 as an Aca protein. We constructed a plasmid with the *acrIF23-acrIF24* promoter upstream of a promoter-less lacZ reporter gene, a strategy previously used to test the function of Aca1 (ref. [19]). Expression of AcrIF24 resulted in a 64% reduction in β-galactosidase activity (Fig. 4a), indicating that AcrIF24 is a transcriptional suppressor. Based on sequence information upstream of the *acrIF23-acrIF24* operon, we synthesized a 66-bp dsDNA that covers the predicted −10 and −35 regions of the promoter (predicted by BPROM[24]) and an inverted repeat (IR) (Fig. 4b). EMSA results suggest that AcrIF24 specifically binds to this promoter DNA (K_D: ~34 nM) but not to a mutant with interrupted IR sequences (Fig. 4c,d). Csy–AcrIF24 shows a slightly higher affinity to promoter DNA (K_D: ~20 nM) than AcrIF24 alone (Fig. 4d,e), possibly due to added interactions contributed by the PAM recognition sites on Csy.

Mutations of key residues in the dimerization interface (for example, H154A/H216A/Y217A/R221A) abolished promoter DNA binding, indicating that dimerization of AcrIF24 is required for binding to promoter DNA (Fig. 4f). Interestingly, the HTH domain of AcrIF24 alone is not capable of binding promoter DNA (Fig. 4f), nor does the HTH domain alone show transcriptional repression in the β-galactosidase activity assay (Fig. 4a). It is likely that other domains, such as NTD, are important for maintaining the native structure and dimeric form of AcrIF24 to fulfill its Aca function.

Structure of Csy–AcrIF24–promoter DNA. To reveal the structural detail of the interaction between AcrIF24 and the promoter DNA, we determined the structure of the Csy–AcrIF24–promoter DNA complex by cryo-EM. Both the Csy–AcrIF24 monomer and dimer states are bound to promoter DNA, and both states were determined

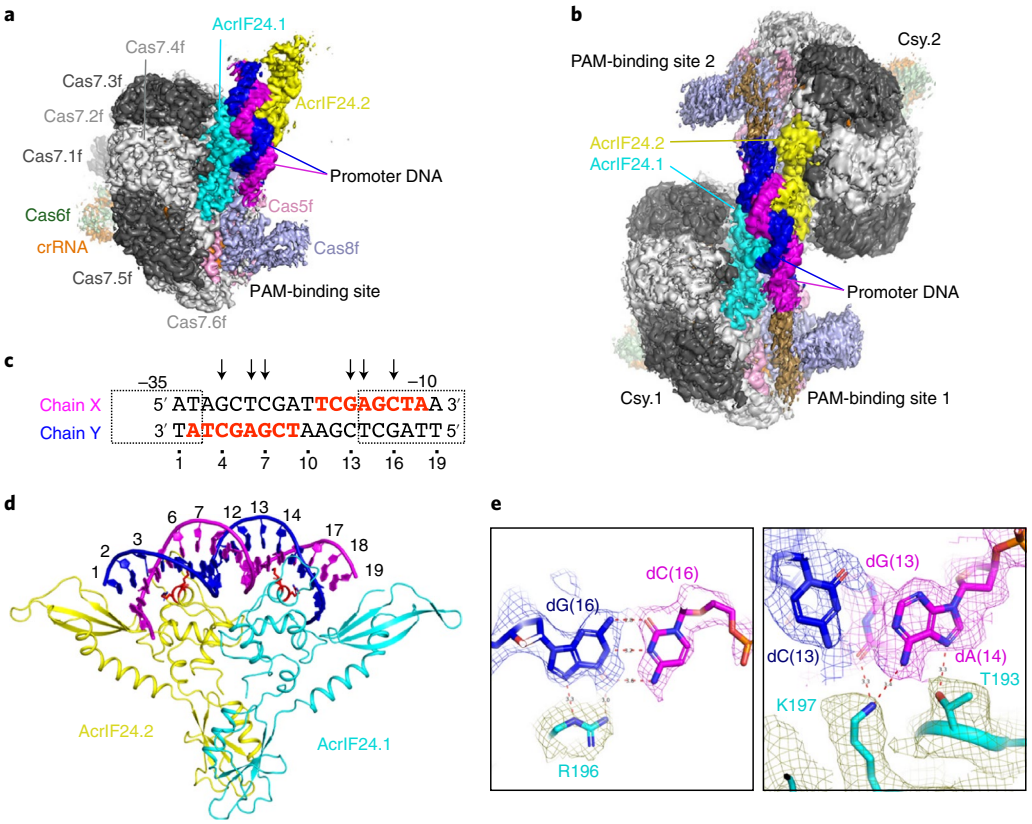


Fig. 5 | Cryo-EM structure of Csy–AcrIF24–promoter DNA complex. a,**b**, Cryo-EM map of Csy–AcrIF24–promoter DNA in monomer (**a**) and dimer (**b**) forms in surface representation with each subunit color-coded as in Fig. 1a. Two strands of DNA are color-coded in blue and magenta, respectively. Unmodeled density at the two ends of promoter DNA is colored brown. **c**, Sequence of the promoter DNA that interacts with AcrIF24 dimer. IRs are indicated by red color. The base pairs involved in specific interaction with AcrIF24 are indicated by arrows. **d**, Structure of AcrIF24–promoter DNA extracted from Csy–AcrIF24–promoter DNA dimer. **e**, Specific interaction of promoter DNA with AcrIF24. Three residues (T193, R196 and K197) from each AcrIF24 protomer are involved in the recognition of promoter DNA. Interactions (hydrogen bonds) are indicated by red dashed lines.

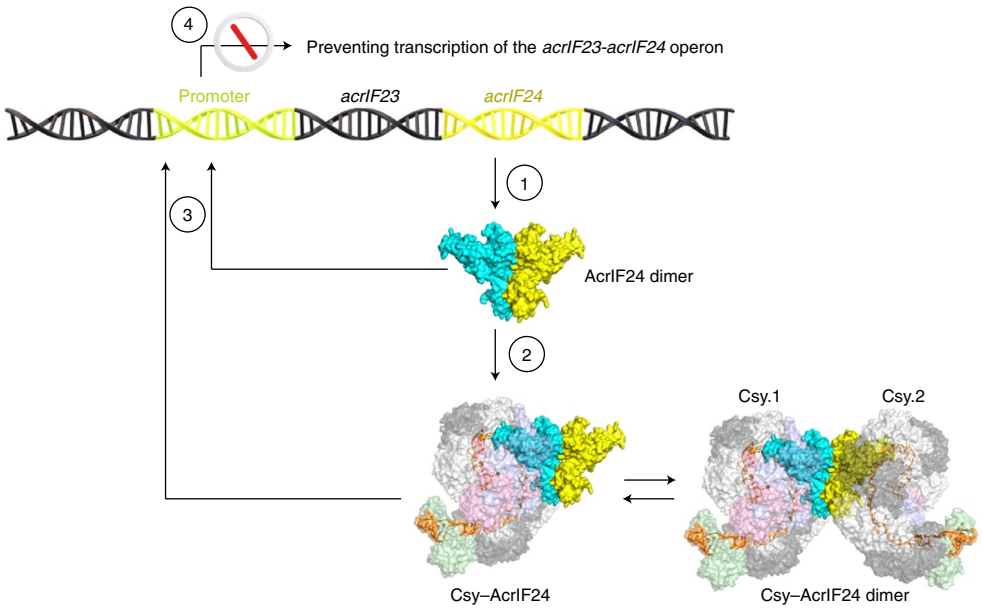


Fig. 6 | Model of the dual function of AcrIF24. A cartoon model showing the Acr and Aca function of AcrIF24. AcrIF24 exists as a dimer (1), which binds to and dimerizes Csy and prevents hybridization between target DNA and crRNA (2). AcrIF24, alone or in complex with Csy, binds to the promoter DNA (3) and suppresses the transcription of the *acrIF23-acrIF24* operon (4).

to around 2.7 Å, which allowed for unambiguous assignment of the bases in the promoter DNA (Fig. 5a,b, Extended Data Fig. 9 and Supplementary Table 1). A 19-bp promoter DNA spanning the predicted −10 and −35 regions was modeled confidently to the EM density map (Fig. 5c,d and Extended Data Fig. 10a). The IR in this region is involved in direct association with both AcrIF24 protomers. Specifically, a G-C base pair at position 16 is recognized by R194 of AcrIF24.1 (similarly, a G-C base pair at position 4 is recognized by R194 of AcrIF24.2) (Fig. 5e). G-C and A-T base pairs at positions 13 and 14 are recognized by K197 and T193 of AcrIF24.1 (similarly, A-T and G-C base pairs at positions 6 and 7 are recognized by K197 and T193 of AcrIF24.2) (Fig. 5e). Binding of the promoter DNA to Csy–AcrIF24 results in bending in the DNA by 43.7° from position 2 to position 18 (calculated by the Curve+ program[25]), similar to the ~42° bending angle observed from the Aca1–promoter DNA structure[22] (Extended Data Fig. 10b). Extra density potentially contributed by DNA flanking the −10 and −35 regions is observed at both PAM recognition sites of Csy; however, the densities are weak, indicative of weak and non-specific interactions (Fig. 5a,b). These results highlight how Csy–AcrIF24 can tightly bind to the *acrIF23-acrIF24* promoter.

Discussion

In this study, we elucidated the structural basis for the dual function of AcrIF24, a fused Acr-Aca protein in type I-F CRISPR–Cas systems or, more broadly, in class 1 CRISPR–Cas systems, and we reported a high-resolution structure of the Csy–AcrIF24–promoter complex. As an Acr, AcrIF24 directly binds to Csy and blocks the hybridization between crRNA and target DNA, thereby inactivating the type I-F CRISPR–Cas system (Fig. 6). Target recognition by Csy involves rapid association and dissociation with dsDNA, searching for a GG PAM sequence[26]. Once the PAM sequence is identified, DNA is partially melted, which allows the hybridization between melted DNA and the crRNA. When the crRNA-DNA hybridization is successful and a stable R-loop structure is formed, the Csy-dsDNA complex is capable of recruiting Cas2/3 nuclease for target cleavage. If not, Csy's search for PAM and spacer sequences continues. AcrIF24 binds to the Cas7 'backbone' and prevents the hybridization between target DNA and crRNA, therefore preventing the formation of the R-loop structure. This mechanism is reminiscent of AcrIF1 (refs. [7,8]), AcrIF9 (refs. [11,12]) and AcrIF14 (refs. [13,15]).

As an Aca, AcrIF24 is capable of binding to the promoter DNA, thereby suppressing the transcription from the *acrIF23-acrIF24* promoter (Fig. 6). The Aca role of AcrIF24 is consistent with a recent study[27]. Here we present a high-resolution structure of Csy–AcrIF24–promoter complex to reveal the structural basis for the Aca role of AcrIF24. In the *acrIF23-acrIF24* promoter, only one IR sequence between the −35 and −10 regions is identifiable. This IR sequence is in an equivalent position to the IR2 of the *acrIF1-aca1* promoter, in which another IR1 is found upstream of the −35 region. The Aca role of Acr proteins was also reported in two studies investigating type II Acrs, including AcrIIA1 (ref. [28]) and AcrIIA13–15 (ref. [29]). The NTD of AcrIIA1 is a critical transcriptional repressor of the strong acr promoter, whereas the C-terminal domain inhibits DNA interference by Cas9 (ref. [28]). The N-termini of AcrIIA13–15 contain HTH motif and are capable of binding the IR regions of the acr promoter[29]. The fact that dual-function inhibitors could span both class 1 and class 2 CRISPR–Cas systems possibly suggests that this could be a widely successful strategy for Acrs.

Although three different domains in AcrIF24 are responsible for separate Acr and Aca roles, AcrIF24 functions as a whole with the dimerization interface being essential for both roles. For example, although both NTD and MID directly bind to the Cas7f backbone of Csy to block the hybridization between crRNA and target DNA, HTH is required for optimal inhibition of Csy (Fig. 2b). On the other hand, although only the HTH domain is involved in DNA binding, as revealed by the structures reported here, the HTH domain alone is not capable of competent binding to promoter DNA (Fig. 4f). One possible explanation is that all domains are required for maintaining the native structure and dimeric form of AcrIF24 for the dual function of Acr and Aca. In summary, our work reveals the structural basis for the dual function of AcrIF24, a fused Acr-Aca protein that inhibits the type I-F CRISPR–Cas system.

Online content

Any methods, additional references, Nature Research reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at https://doi.org/10.1038/s41589-022-01137-w.

Received: 4 January 2022; Accepted: 8 August 2022;
Published online: 26 September 2022

References

1. Makarova, K. S. et al. An updated evolutionary classification of CRISPR–Cas systems. *Nat. Rev. Microbiol.* **13**, 722 (2015).
2. Koonin, E. V., Makarova, K. S. & Zhang, F. Diversity, classification and evolution of CRISPR–Cas systems. *Curr. Opin. Microbiol.* **37**, 67–78 (2017).
3. Bondy-Denomy, J., Pawluk, A., Maxwell, K. L. & Davidson, A. R. Bacteriophage genes that inactivate the CRISPR/Cas bacterial immune system. *Nature* **493**, 429–432 (2013).
4. Pawluk, A. et al. Inactivation of CRISPR-Cas systems by anti-CRISPR proteins in diverse bacterial species. *Nat. Microbiol* **1**, 16085 (2016).
5. Marino, N. D. et al. Discovery of widespread type I and type V CRISPR–Cas inhibitors. *Science* **362**, 240–242 (2018).
6. Pinilla-Redondo, R. et al. Discovery of multiple anti-CRISPRs highlights anti-defense gene clustering in mobile genetic elements. *Nat. Commun.* **11**, 5652 (2020).
7. Chowdhury, S. et al. Structure reveals mechanisms of viral suppressors that intercept a CRISPR RNA-guided surveillance complex. *Cell* **169**, 47–57 (2017).
8. Guo, T. W. et al. Cryo-EM structures reveal mechanism and inhibition of DNA targeting by a CRISPR–Cas surveillance complex. *Cell* **171**, 414–426 (2017).
9. Peng, R. et al. Alternate binding modes of anti-CRISPR viral suppressors AcrF1/2 to Csy surveillance complex revealed by cryo-EM structures. *Cell Res.* **27**, 853–864 (2017).
10. Rollins, M. F. et al. Structure reveals a mechanism of CRISPR-RNA-guided nuclease recruitment and anti-CRISPR viral mimicry. *Mol. Cell* **74**, 132–142 (2019).
11. Hirschi, M. et al. AcrIF9 tethers non-sequence specific dsDNA to the CRISPR RNA-guided surveillance complex. *Nat. Commun.* **11**, 2730 (2020).
12. Zhang, K. M. et al. Inhibition mechanisms of AcrF9, AcrF8, and AcrF6 against type I-F CRISPR–Cas complex revealed by cryo-EM. *Proc. Natl Acad. Sci. USA* **117**, 7176–7182 (2020).
13. Gabel, C., Li, Z., Zhang, H. & Chang, L. Structural basis for inhibition of the type I-F CRISPR–Cas surveillance complex by AcrIF4, AcrIF7 and AcrIF14. *Nucleic Acids Res.* **49**, 584–594 (2021).
14. Lu, W. T., Trost, C. N., Muller-Esparza, H., Randau, L. & Davidson, A. R. Anti-CRISPR AcrIF9 functions by inducing the CRISPR–Cas complex to bind DNA non-specifically. *Nucleic Acids Res.* **49**, 3381–3393 (2021).
15. Liu, X. et al. Insights into the dual functions of AcrIF14 during the inhibition of type I-F CRISPR–Cas surveillance complex. *Nucleic Acids Res.* **49**, 10178–10191 (2021).
16. Wang, X. F. et al. Structural basis of Cas3 inhibition by the bacteriophage protein AcrF3. *Nat. Struct. Mol. Biol.* **23**, 868–870 (2016).
17. Niu, Y. et al. A type I-F anti-CRISPR protein inhibits the CRISPR–Cas surveillance complex by ADP-ribosylation. *Mol. Cell* **80**, 512–524 (2020).
18. Bondy-Denomy, J. et al. Multiple mechanisms for CRISPR–Cas inhibition by anti-CRISPR proteins. *Nature* **526**, 136–139 (2015).
19. Stanley, S. Y. et al. Anti-CRISPR-associated proteins are crucial repressors of anti-CRISPR transcription. *Cell* **178**, 1452–1464 (2019).
20. Birkholz, N., Fagerlund, R. D., Smith, L. M., Jackson, S. A. & Fineran, P. C. The autoregulator Aca2 mediates anti-CRISPR repression. *Nucleic Acids Res.* **47**, 9658–9665 (2019).
21. Kim, H. et al. MMOD-induced structural changes of hydroxylase in soluble methane monooxygenase. *Sci. Adv.* **5**, eaax0059 (2019).

22. Liu, Y. et al. Structural basis for anti-CRISPR repression mediated by bacterial operon proteins Aca1 and Aca2. *J. Biol. Chem.* **297**, 101357 (2021).
23. Rollins, M. F. et al. Structure reveals a mechanism of CRISPR-RNA-guided nuclease recruitment and anti-CRISPR viral mimicry. *Mol. Cell* **74**, e135 (2019).
24. Li, R. W. *Metagenomics and Its Applications in Agriculture, Biomedicine, and Environmental Studies* (Nova Science Publishers, 2011).
25. Lavery, R., Moakher, M., Maddocks, J. H., Petkeviciute, D. & Zakrzewska, K. Conformational analysis of nucleic acids revisited: Curves+. *Nucleic Acids Res.* **37**, 5917–5929 (2009).
26. Rollins, M. F., Schuman, J. T., Paulus, K., Bukhari, H. S. & Wiedenheft, B. Mechanism of foreign DNA recognition by a CRISPR RNA-guided surveillance complex from *Pseudomonas aeruginosa*. *Nucleic Acids Res.* **43**, 2216–2222 (2015).
27. Yang, L. G. et al. Insights into the inhibition of type I-F CRISPR-Cas system by a multifunctional anti-CRISPR protein AcrIF24. *Nat. Commun.* **13**, 1931 (2022).
28. Osuna, B. A. et al. Critical anti-CRISPR locus repression by a bi-functional Cas9 inhibitor. *Cell Host Microbe* **28**, 23–30 e25 (2020).
29. Watters, K. E. et al. Potent CRISPR–Cas9 inhibitors from *Staphylococcus* genomes. *Proc. Natl Acad. Sci. USA* **117**, 6531–6539 (2020).

Methods

Plasmid constructs and protein purification. The DNA sequences encoding AcrIF23 and AcrIF24 were ordered as gBlocks from Integrated DNA Technologies (IDT). The sequnces were cloned individually into pET His6 Sumo TEV LIC cloning vectors (1S) (Addgene, 29659) using the Gibson Assembly Master Mix (NEB, E2611S). After sequencing verification, these plasmids were transformed into *Escherichia coli* BL21(DE3) cells for expression in Terrific Broth (TB) medium. Protein expression was induced when the optical density at 600 nm (OD_{600}) reached 0.6 by adding 0.5 mM isopropyl β-D-1-thiogalactopyranoside (IPTG), followed by incubation at 16 °C overnight. After harvesting, cell pellets were resuspended in lysis buffer (50 mM HEPES (pH 7.5), 500 mM NaCl, 5% glycerol and 5 mM β-mercaptoethanol) supplemented with 0.2 mM phenylmethylsulfonyl fluoride (PMSF). Cells were disrupted by sonication. After centrifugation, the supernatant was incubated with Ni-NTA beads for 30 minutes and loaded to a gravity flow column, followed by an extensive wash using lysis buffer. The target AcrIF24 (or AcrIF23) protein was eluted in a buffer containing 50 mM HEPES (pH 7.5), 500 mM NaCl, 5% glycerol, 1 mM tris(2-carboxyethyl) phosphine hydrochloride (TCEP) and 350 mM imidazole. After digestion by TEV protease overnight to remove the 6×His and SUMO tags, AcrIF24 (or AcrIF23) protein was purified with ion-exchange chromatography using HiTrap Q HP. AcrIF24 (or AcrIF23) protein was then concentrated and further purified over a Superdex 200 or a Superdex 75 column (Cytiva) in a buffer containing 20 mM HEPES (pH 7.5), 150 mM NaCl, 5% glycerol and 1 mM TCEP.

For purification of the PA14 Csy complex, we used a similar protocol as described for AcrIF24. For purification of Cas2/3, the pMBPCas2/3 plasmid (Addgene, 89238) was transformed into BL21(DE3) cells. Protein expression was induced as described above for AcrIF24. After harvesting, cell pellets were resuspended in lysis buffer (50 mM HEPES (pH 7.5), 500 mM NaCl, 5% glycerol and 1 mM TCEP) supplemented with 0.2 mM PMSF and disrupted by sonication. After centrifugation, the supernatant was passed through a StrepTrap HP column for binding the tagged Cas2/3, followed by washing with lysis buffer. The protein was eluted with lysis buffer supplemented with 2.5 mM desthiobiotin. The collected elution fractions were concentrated and subjected to a Superdex 200 column (Cytiva) for further purification.

Mutagenesis. All AcrIF24 and Csy mutations were generated by PCR-based QuickChange site-directed mutagenesis. All mutations were confirmed by DNA sequencing. The primers used for making these mutations are listed in Supplementary Table 2. The mutant proteins were expressed and purified using the same protocol as used for the wild-type proteins.

Substrate DNA preparation. To prepare dsDNA substrate for in vitro cleavage assays and EMSAs, a 54-nt non-target DNA strand labeled with 5′-fluorescein (FAM) (synthesized by GENEWIZ) and a 54-nt target strand without a label (synthesized by IDT) were mixed at a molar ratio of 1:1.25, denatured at 95 °C for 5 minutes and then annealed at room temperature. The DNA sequences are listed in Supplementary Table 2.

Complex assembly. To assemble the Csy–AcrIF24 complex, purified Csy was incubated with AcrIF24 at a molar ratio of 1:10 for 1 hour on ice. The mixer was then subjected to a Superdex 200 column (Cytiva) equilibrated in buffer containing 20 mM HEPES (pH 7.5), 100 mM KCl, 5% glycerol and 1 mM TCEP. Elution fractions contatining the assembled complex were examined by SDS-PAGE analysis. To assemble the Csy–AcrIF24–dsDNA complex, Csy was incubated with a ten-fold molar excess of AcrIF24 and a 1.5 molar excess of target dsDNA for 1 hour on ice. The mixer was then subjected to a Superdex 200 column (Cytiva) equilibrated in buffer containing 20 mM HEPES (pH 7.5), 100 mM KCl, 5% glycerol, 1 mM TCEP and 5 mM $MgCl_2$. Elution fractions were examined by SDS-PAGE and 15% polyacrylamide gels containing 7 M urea.

In vitro DNA cleavage assay. Reactions were typically performed with final concentrations of Csy at 1.2 μM, Cas2/3 at 0.15 μM and substrate DNA at 0.04 μM. The concentrations of AcrIF24 and its variants are specified in the figures or legends.

We first incubated AcrIF24 with Csy complex at 4 °C in the buffer containing 20 mM HEPES (pH 7.5), 100 mM KCl, 5% glycerol and 1 mM TCEP for 30 minutes. Substrate DNA was added and incubated for another 15 minutes. Then, Cas2/3 was introduced along with 5 mM $MgCl_2$, 75 μM $NiSO_4$, 5 mM $CaCl_2$ and 1 mM ATP. After 30 minutes at 37 °C, the reaction was quenched with 1% SDS and 8 mM EDTA. The products were separated by electrophoresis over 15% TBE polyacrylamide gels containing 7 M urea and visualized by fluorescence imaging using a GE Amersham Imager 600 system.

For quantification, the intensities of uncleaved substrate DNA were quantified using ImageJ software[30]. DNA input was calibrated to 100%, and uncleaved substrate DNA in each lane was measured. Then, the percentage of cleaved DNA was calculated and illustrated as a bar graph using GraphPad Prism.

Cas2/3 recruitment assay. The Cas2/3 recruitment assay was performed by incubating 1.2 μM Csy complex with 1.2 μM AcrIF23 or AcrIF24 for 30 minutes at 4 °C in a buffer containing 20 mM HEPES (pH 7.5), 150 mM KCl, 1 mM $MgCl_2$ and 5% glycerol. Substrate DNA (0.04 μM) was then introduced into the reaction mixture and further incubated for 15 minutes at 37 °C. Cas2/3 was then added to the reaction mixture and was further incubated for 30 minutes at 37 °C. Products of the reaction were separated using 10% TBE polyacrylamide gels and visualized by fluorescence imaging.

EMSA. To test the effects of AcrIF24 (or its variants) on the recognition of substrate DNA by Csy (Extended Data Fig. 5b,d,f), 1.2 μM Csy complex was incubated with AcrIF24 (or its variants) in a 1:1 molar ratio at 4 °C for 30 minutes. Then, 40 nM DNA substrate was added and incubated at 4 °C for another 30 minutes. Products were separated using 6% TBE polyacrylamide gels and visualized by fluorescence imaging. All binding reactions were conducted in buffer A containing 20 mM HEPES (pH 7.5), 150 mM KCl, 1 mM $MgCl_2$ and 5% glycerol.

To test DNA binding to Csy–AcrIF24 (Extended Data Fig. 8c,d), 20 nM 5′-FAM-labeled DNA was incubated with increasing concentration of Csy–AcrIF24 (1, 20, 50, 100, 150, 200, 250, 300, 600, 1,200 and 1,500 nM) in buffer A at 4 °C for 30 minutes. Products were separated using 6% TBE polyacrylamide gels and visualized by fluorescence imaging.

In competition EMSA experiments (Extended Data Fig. 8e,f), Csy and AcrIF24 were incubated in a 1:1 molar ratio at 4 °C for 30 minutes, followed by the addition of 0.25 μM FAM-labeled specific dsDNA (dsDNAsp), and the reaction mixture was incubated at 37 °C for 30 minutes in buffer A. Then, the non-specific dsDNA (dsDNAns with no FAM label) was added at increasing concentrations (0.125, 0.25, 0.5 and 1 μM) and incubated at 4 °C for another 30 minutes. Products were separated using 6% TBE polyacrylamide gels and visualized by fluorescence imaging.

To calculate the affinities between AcrIF24 (or Csy–AcrIF24) and the promoter DNA (Fig. 4c,e), 20 nM 5′-FAM-labeled wild-type or mutant promoter DNA was incubated with increasing concentrations (1, 5, 10, 15, 20, 40, 60, 80, 1,000, 1,250 and 1,500 nM) of AcrIF24 (or Csy–AcrIF24) in buffer A for 30 minutes at 4 °C. Products were separated using 10% (for AcrIF24) and 6% (for Csy–AcrIF24) TBE polyacrylamide gels and visualized by fluorescence imaging.

To test the effect of AcrIF24 mutations on promoter DNA binding (Fig. 4f), 20 nM 5′-FAM-labeled promoter DNA was incubated with an increasing concentration of AcrIF24 mutants (1.5 μM, 3 μM and 6 μM) in buffer A for 30 minutes at 4 °C. Products were separated using 10% TBE polyacrylamide gels and visualized by fluorescence imaging.

The quantification of the EMSA data was performed by measuring the intensities of free DNA bands using ImageJ[30] with DNA input calibrated at 100%. The percentage of bound DNA was calculated and illustrated as bar graphs. For calculation of binding affinity, the binding curves were fitted using the 'specific binding with Hill slope' function in GraphPad Prism.

β-galactosidase assay. A promoter-less β-galactosidase reporter shuttle vector pQF50 and a vector with constitutive promoter pCM-Str were gifts from the Alan R. Davidson laboratory[18]. The whole plasmids were sequenced for confirmation and construction purposes. AcrIF24 was PCR-amplified from its gBlock using primers listed in Supplementary Table 2 and cloned into pCM-Str using Gibson isothermal assembly. To insert the promoter DNA to pQF50, a pair of primers with 30-nt overhangs matching the promoter DNA were designed. pQF50 plasmid was PCR-amplified using the primers. PCR products were cleaned up with a PCR clean-up kit (Qiagen), and then DpnI was digested to remove the original pQF50 with no promoter insertion. Blunt ends of the PCR product were annealed using Blunt/TA Ligase Master Mix (New England Biolabs). The insertion of promoter and *acrIF24* gene was sequence verified.

The modified pQF50 and pCM-Str plasmids designed for the β-galactosidase assay were transformed into competent DH5α *E. coli* cells. A single colony was selected and cultured overnight. The cells were then diluted by 1:100 and grown for ~3 hours with shaking until the OD_{600} reached ~0.6. Next, 20 μl of culture was mixed with 80 μl of permeabilization solution (0.8 mg ml^{-1} of CTAB, 0.4 mg ml^{-1} of sodium deoxycholate, 100 mM Na_2HPO_4, 20 mM KCl, 2 mM $MgSO_4$ and 5.4 μl ml^{-1} of β-mercaptoethanol) and incubated at 30 °C for 30 minutes. Then, 600 μl of substrate solution (60 mM Na_2HPO_4, 40 mM NaH_2PO_4 and 1 mg ml^{-1} of *o*-nitrophenyl-β-galactosidase) was added, and the reaction was allowed to proceed at 30 °C for 30 minutes. The reaction was stopped with the addition of 700 μl of 1 M Na_2CO_3; absorbance at 420 nm (A420) was measured; and Miller units were calculated. The β-galactosidase activity was then quantified using a method derived from ref. [31].

SEC-SAXS. Purified samples were subjected to SEC-SAXS with or without multi-angle light scattering (MALS) using either a Superdex 200 Increase GL or a Superose 6 column (Cytiva). The scattering data were collected at beamline 18-ID by the Biophysics Collaborative Access Team (BioCAT) of the Advanced Photon Source, Argonne National Laboratory. In brief, for SEC-SAXS, the sample(s) were analyzed using an ÄTKA Pure (Cytiva) FPLC, whereas, for SEC-MALS-SAXS, the sample was analyzed using a 1260 Infinity II HPLC (Agilent Technologies) in line with a MALS + DLS detector (DAWN HELEOS II, Wyatt Technology) and an RI detector (Optilab T-rEX, Wyatt Technology). The SAXS data were collected as the

sample was passed through a co-flow cell using a Pilatus3 X 1M detector (Dectris). The SAXS data were analyzed, and final plots were made using BioXTAS RAW[32] and ATSAS[33]. The data were reduced, and the data range for scattering curves was selected. Upon averaging the data, the q-range and molecular weight information were obtained by Guinier analysis. The pair–distance distribution curves were calculated using GNOM. Theoretical scattering for the X-ray and cryo-EM structures were calculated and compared with experimental scattering curves using CRYSOL, with structural superpositions performed with SUPCOMB[34]. Analysis of the MALS data was performed using ASTRA 7 software (Wyatt Technology).

EM. Aliquots of 3-µl Csy–AcrIF24 or Csy–AcrIF24–DNA complexes at ~2 mg ml^{-1} were applied to glow-discharged UltrAuFoil holey gold grids (R1.2/1.3, 300 mesh). The grids were prepared using a Thermo Fisher Scientific Mark IV Vitrobot and transferred into a Titan Krios microscope operated at 300 kV. Cryo-EM data were collected at a nominal magnification of ×81,000, resulting in a calibrated physical pixel size of 1.05 Å per pixel. The images were recorded on a K3 electron direct detector in super-resolution mode at the end of a GIF Quantum energy filter operated with a slit width of 20 eV. Imaged were collected using Leginon[35] with a defocus range of 1.2–2.5 µm. A dose rate of 20 electrons per pixel per second and an exposure time of 3.12 seconds were used, generating 40 movie frames with a total dose of ~54 electrons per Å^2. Statistics for cryo-EM data are listed in Supplementary Table 1.

The Csy–AcrIF24–promoter DNA complex was studied similarly, except for that the data were collected on a Titan Krios G4 microscope, and the images have a calibrated physical pixel size of 1.054 Å per pixel.

Image processing. As for the Csy–AcrIF24 dataset, the movie frames were imported to cryoSPARC[36]. Movie frames were aligned using patch motion correction with a binning factor of 2. Contrast transfer function (CTF) parameters were estimated using Patch CTF[37]. A few thousand particles were auto-picked without a template to generate two-dimensional (2D) averages for subsequent template-based auto-picking. The auto-picked and extracted particles were screened by 2D classification and selection to exclude false and bad particles that fall into 2D averages with poor features. A subset of the particles (100,000 particles) was used to generate an initial model in cryoSPARC.

The screened particles were imported to RELION-3 (ref. [38]) for three-dimensional (3D) classification, using an initial model generated by cryoSPARC. All three resultant 3D classes contain AcrIF24; however, class 2 lacks the Cas8–Cas5 heterodimer, and class 3 shows poor features (Extended Data Fig. 3f). Particles in class 1 (111,955) were further refined in RELION, resulting in a map at an overall resolution of 3.2 Å based on the FSC = 0.143 criterion. Focused refinements around the inhibitors were further performed to improve the local resolutions. A combined map was calculated by the 'vop maximum' program in Chimera[39] and used for model building. A workflow for data processing is summarized in Extended Data Fig. 3.

A similar workflow was used for the Csy–AcrIF24–DNA and Csy–AcrIF24–promoter DNA datasets, except for that all processing was done in cryoSPARC (Extended Data Figs. 7 and 9).

Model building and refinement. De novo model building of the AcrIF24 structures was performed manually in Coot[40]. Secondary structure predictions by PSIPRED[41] were used to assist manual building. To build models of Csy, each subunit of the structure of the Csy–AcrIF4 complex (PDB: 6NE0) was fitted into the maps as a rigid body in UCSF Chimera[39] and manually adjusted in Coot. Refinement of the structure models against corresponding maps was performed using the 'phenix.real_space_refine' tool in Phenix[42]. The buried interface area was calculated by the PDBePISA web server. The AcrIF23 model used for fitting the SEC-SAXS bead model was generated by AlphaFold2 (ref. [43]). The poly(G) model was built by the 'other modeling tools' in Coot. Specifically, a double-stranded poly(G) structure was built in an ideal B-form. The structure was fitted onto the cryo-EM map as a rigid body using the 'fit in map' program in UCSF Chimera. The model was then fitted in Coot with all molecule self-restrains to maintain hydrogen bonds and stacking interactions between bases.

Visualization. Figures were generated using PyMOL and UCSF Chimera[39].

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

Data availability

Cryo-EM reconstructions of Csy–AcrIF24, Csy–AcrIF24 dimer, Csy–AcrIF24–DNA dimer, Csy–AcrIF24–Promoter DNA and Csy–AcrIF24–Promoter DNA dimer have been deposited in the Electron Microscopy Data Bank under accession numbers EMD-25660, EMD-25661, EMD-25662, EMD-25789 and EMD-25788, respectively. Coordinates for atomic models of Csy–AcrIF24, Csy–AcrIF24 dimer, Csy–AcrIF24–DNA dimer, Csy–AcrIF24–Promoter DNA and Csy–AcrIF24–Promoter DNA dimer have been deposited in the Protein Data Bank under accession numbers 7T3J, 7T3K, 7T3L, 7TAX and 7TAW, respectively. Structures of Csy alone and Csy–AcrIF4 are available in the Protein Data Bank under accession numbers 6B45 and 7JZW, respectively. Source data are provided with this paper.

References

30. Schneider, C. A., Rasband, W. S. & Eliceiri, K. W. NIH Image to ImageJ: 25 years of image analysis. *Nat. Methods* **9**, 671–675 (2012).
31. Zhang, X. & Bremer, H. Control of the *Escherichia coli* rrnB P1 promoter strength by ppGpp. *J. Biol. Chem.* **270**, 11181–11189 (1995).
32. Hopkins, J. B., Gillilan, R. E. & Skou, S. BioXTAS RAW: improvements to a free open-source program for small-angle X-ray scattering data reduction and analysis. *J. Appl. Crystallogr.* **50**, 1545–1553 (2017).
33. Manalastas-Cantos, K. et al. ATSAS 3.0: expanded functionality and new tools for small-angle scattering data analysis. *J. Appl. Crystallogr.* **54**, 343–355 (2021).
34. Kozin, M. B. & Svergun, D. I. Automated matching of high- and low-resolution structural models. *J. Appl. Crystallogr.* **34**, 33–41 (2001).
35. Suloway, C. et al. Automated molecular microscopy: the new Leginon system. *J. Struct. Biol.* **151**, 41–60 (2005).
36. Punjani, A., Rubinstein, J. L., Fleet, D. J. & Brubaker, M. A. cryoSPARC: algorithms for rapid unsupervised cryo-EM structure determination. *Nat. Methods* **14**, 290–296 (2017).
37. Zhang, K. Gctf: real-time CTF determination and correction. *J. Struct. Biol.* **193**, 1–12 (2016).
38. Zivanov, J. et al. New tools for automated high-resolution cryo-EM structure determination in RELION-3. *eLife* **7**, e42166 (2018).
39. Pettersen, E. F. et al. UCSF Chimera—a visualization system for exploratory research and analysis. *J. Comput. Chem.* **25**, 1605–1612 (2004).
40. Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. Features and development of Coot. *Acta Crystallogr D Biol. Crystallogr.* **66**, 486–501 (2010).
41. Jones, D. T. Protein secondary structure prediction based on position-specific scoring matrices. *J. Mol. Biol.* **292**, 195–202 (1999).
42. Afonine, P. V. et al. Real-space refinement in PHENIX for cryo-EM and crystallography. *Acta Crystallogr. D Struct. Biol.* **74**, 531–544 (2018).
43. Jumper, J. et al. Highly accurate protein structure prediction with AlphaFold. *Nature* **596**, 583–589 (2021).

Acknowledgements

We thank T. Klose and J. Sun for help with cryo-EM; S. Wilson for computation; R. V. Stahelin for providing access to the GE Amersham Imager 600 system; A. R. Davidson for sharing plasmids; J. Hopkins for assistance with the SAXS and MALS experiments and analyses; and other beamline staff of the BioCAT beamline at the Advanced Photon Source, Argonne National Laboratory, for their help with data collection. This work was supported by National Institutes of Health (NIH) grant R01GM138675 and a Core Pilot grant from the Indiana Clinical and Translational Sciences Institute to L.C. C.G. is supported by a grant from the NIH (T32GM132024). The BioCAT resources are supported by grant P30 GM138395 from the National Institute of General Medical Sciences (NIGMS) of the NIH. The use of the Pilatus3 1M detector was provided by grant 1S10OD018090 from the NIGMS. This work made use of the Purdue Cryo-EM Facility and resources of the Advanced Photon Source, a US Department of Energy (DOE) Office of Science User Facility operated for the DOE Office of Science by the Argonne National Laboratory under contract number DE-AC02-06CH11357.

Author contributions

L.C. supervised the study. I.A.M. prepared samples, with help from C.G. I.A.M. performed the biochemical assays. I.A.M., C.G. and L.C. collected and processed cryo-EM data. I.A.M. and N.N. performed SAXS experiments and analyzed the data. L.C. and I.A.M. built the structural models. J.B.-D. identified the promoter sequence and elements. I.A.M. and L.C. prepared the manuscript, with input from C.G., N.N. and J.B.-D.

Competing interests

J.B.-D. is a scientific advisory board member of SNIPR Biome and Excision Biotherapeutics and a scientific advisory board member and co-founder of Acrigen Biosciences. The Bondy-Denomy laboratory receives research support from Felix Biotechnology. The remaining authors declare no competing interests.

Additional information

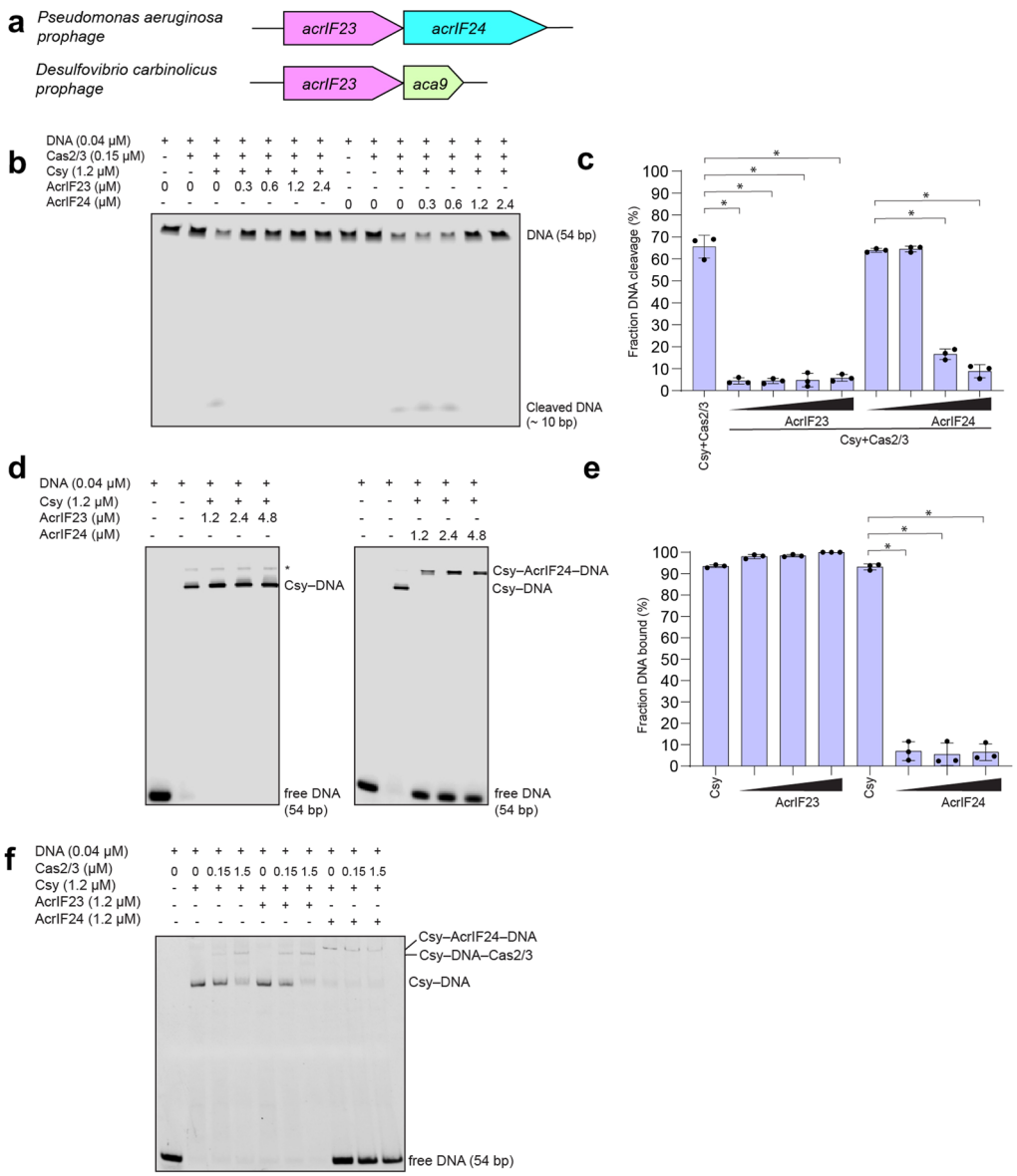
Extended data is available for this paper at https://doi.org/10.1038/s41589-022-01137-w.

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41589-022-01137-w.

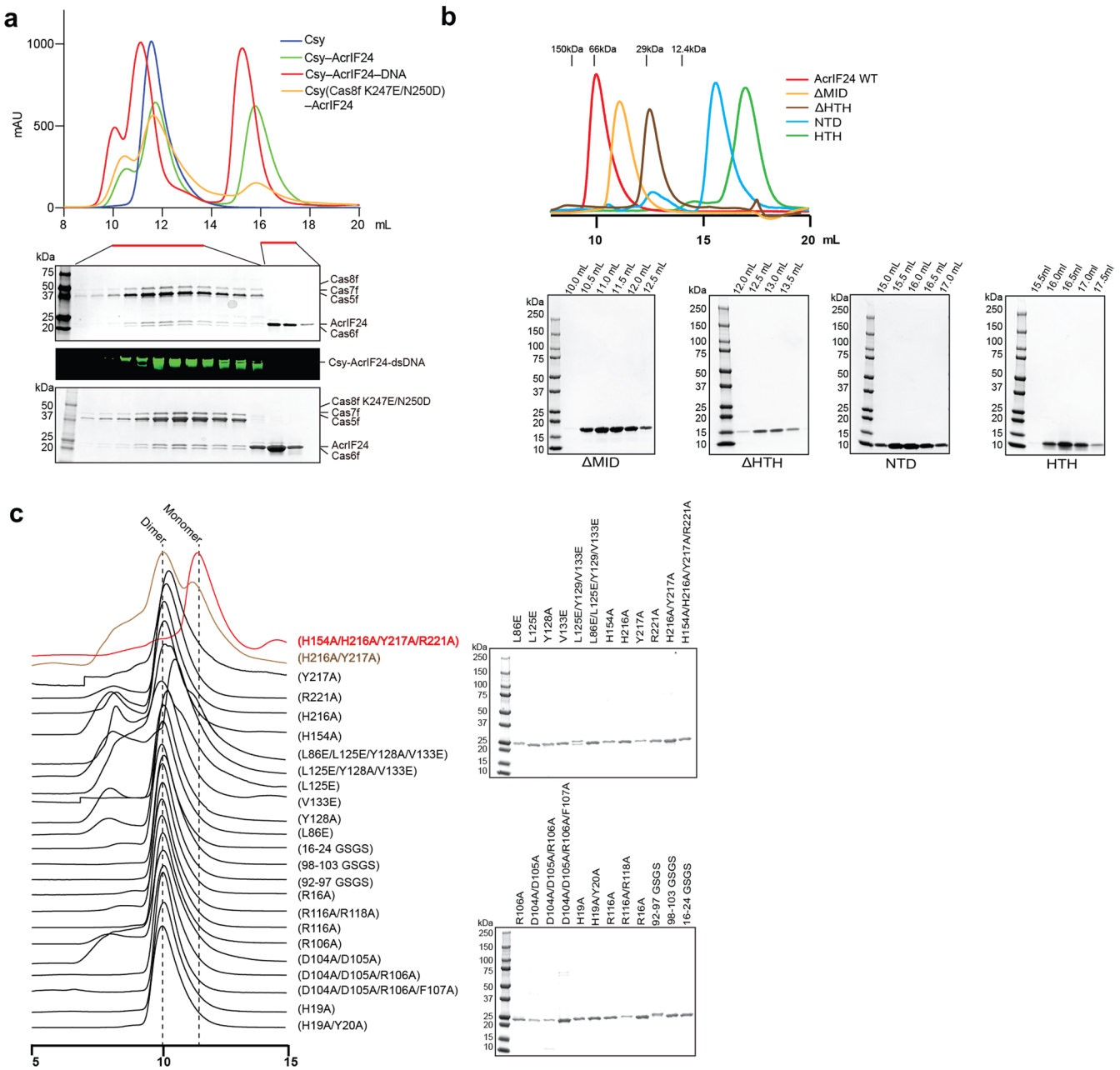
Correspondence and requests for materials should be addressed to Leifu Chang.

Peer review information *Nature Chemical Biology* thanks Tina Liu and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

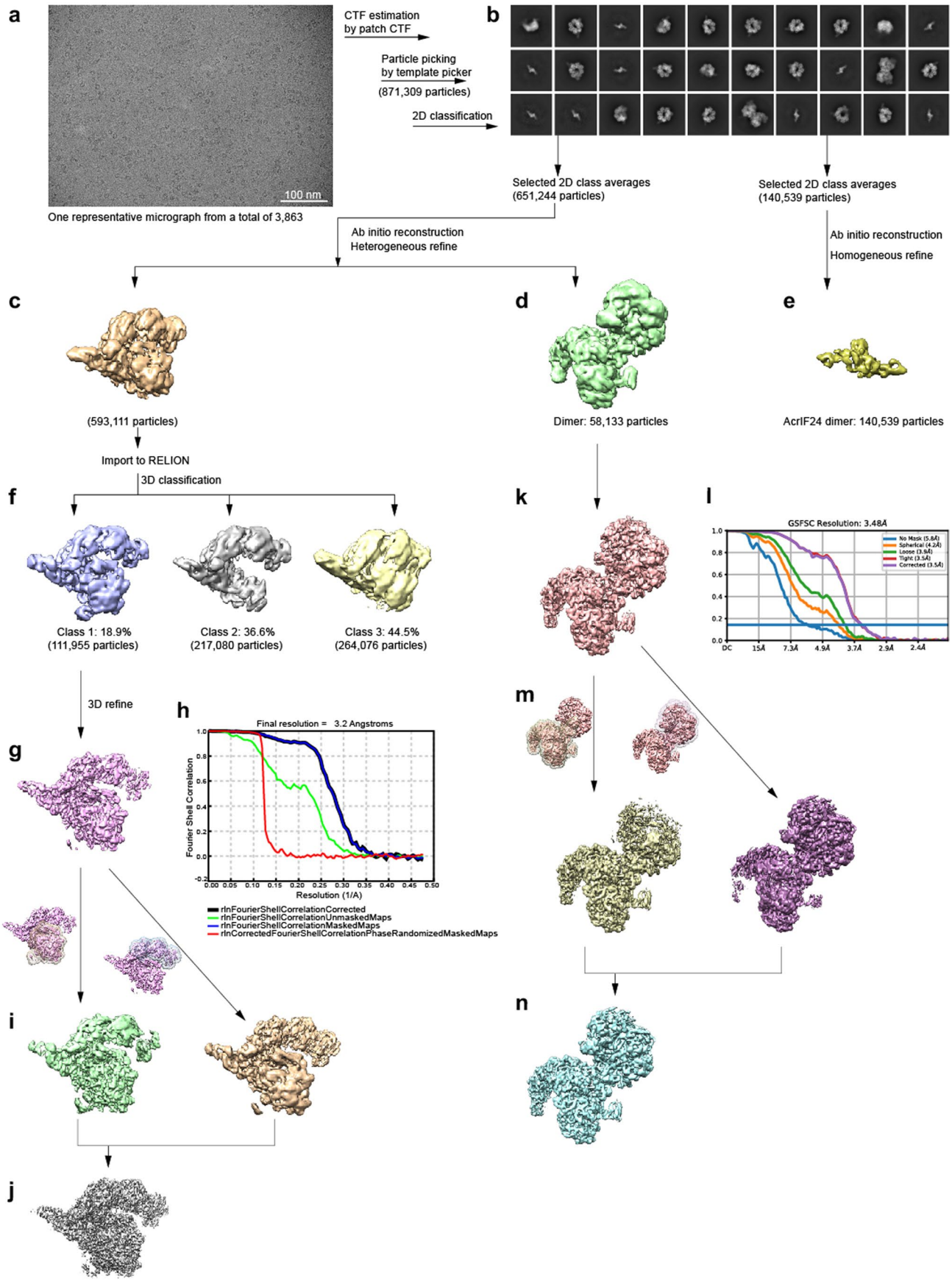
Reprints and permissions information is available at www.nature.com/reprints.



Extended Data Fig. 1 | AcrIF23 and AcrIF24 inhibit *in vitro* DNA cleavage by Csy and Cas2/3. (a) Cartoon diagram of the *acrIF23-acrIF24* operon found in a *P. aeruginosa* prophage. No *aca* gene downstream of the acrIF23–24 locus was found. In comparison, the *aca9* gene was found downstream of *acrIF23* in a *D. carbinolicus* prophage. Adapted from[6]. **(b)** AcrIF23 and AcrIF24 inhibits substrate DNA cleavage by the type I-F CRISPR-Cas system in the *in vitro* DNA cleavage assay. **(c)** Quantitative result of the experiment in **(b)**. Error bars represent SD; $n = 3$. Two-sided *t* test was performed ($^*P < 0.05$). **(d)** EMSA results reveal that AcrIF24 inhibits substrate DNA binding to Csy whereas AcrIF23 does not. Note: Csy–AcrIF24 is capable of binding DNA independent of PAM and spacer sequences, resulting in a high molecular weight shift. * indicates a faint band where two copies of Csy bound to DNA, similar to previous report[14] (doi.org/10.1093/nar/gkab092). **(e)** Quantitative result of the experiment in **(d)**. Error bars represent SD; $n = 3$. Two-sided *t* test was performed ($^*P < 0.05$). **(f)** EMSA results reveal that Cas2/3 recruitment by Csy is abolished by AcrIF24 but not by AcrIF23.

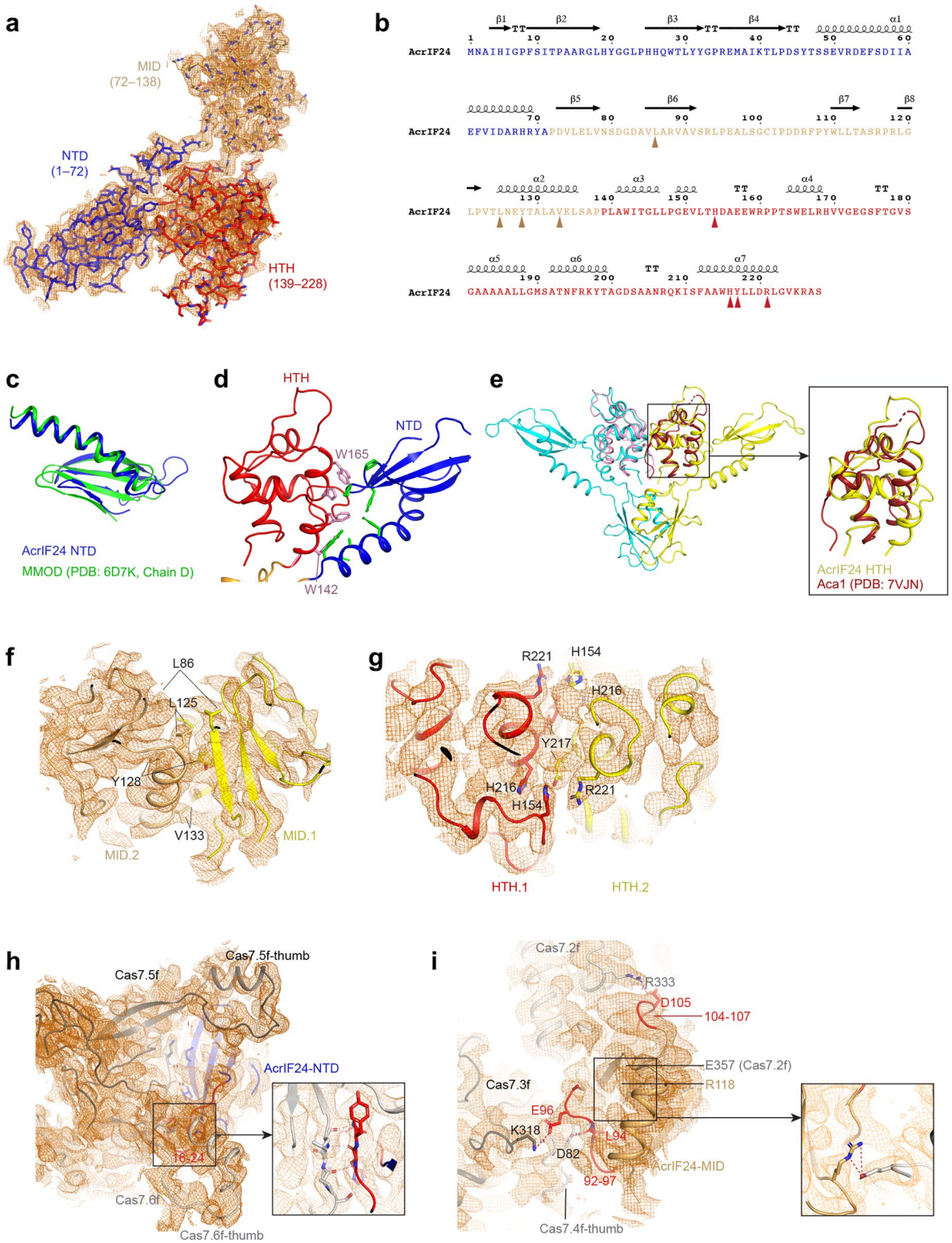


Extended Data Fig. 2 | Biochemical and biophysical analysis of AcrIF23 and AcrIF24. (a) Size exclusion chromatography (SEC) profiles of Csy, Csy–AcrIF24, Csy(Cas8f K247E/N250D)–AcrIF24 and Csy–AcrIF24–DNA complexes. Fractions of Csy–AcrIF24–DNA as indicated were analyzed by SDS-PAGE (to visualize proteins) and 10% Native PAGE gel (to visualize DNA), results of which are shown below the SEC profiles. SDS-PAGE of fractions of Csy(Cas8f K247E/N250D)–AcrIF24 is also shown below the SEC profiles. **(b)** SEC profiles (top) and SDS-PAGE (bottom) of AcrIF24 and its domain deletion mutants. **(c)** SEC profiles (left) and SDS-PAGE (right) of AcrIF24 mutants.

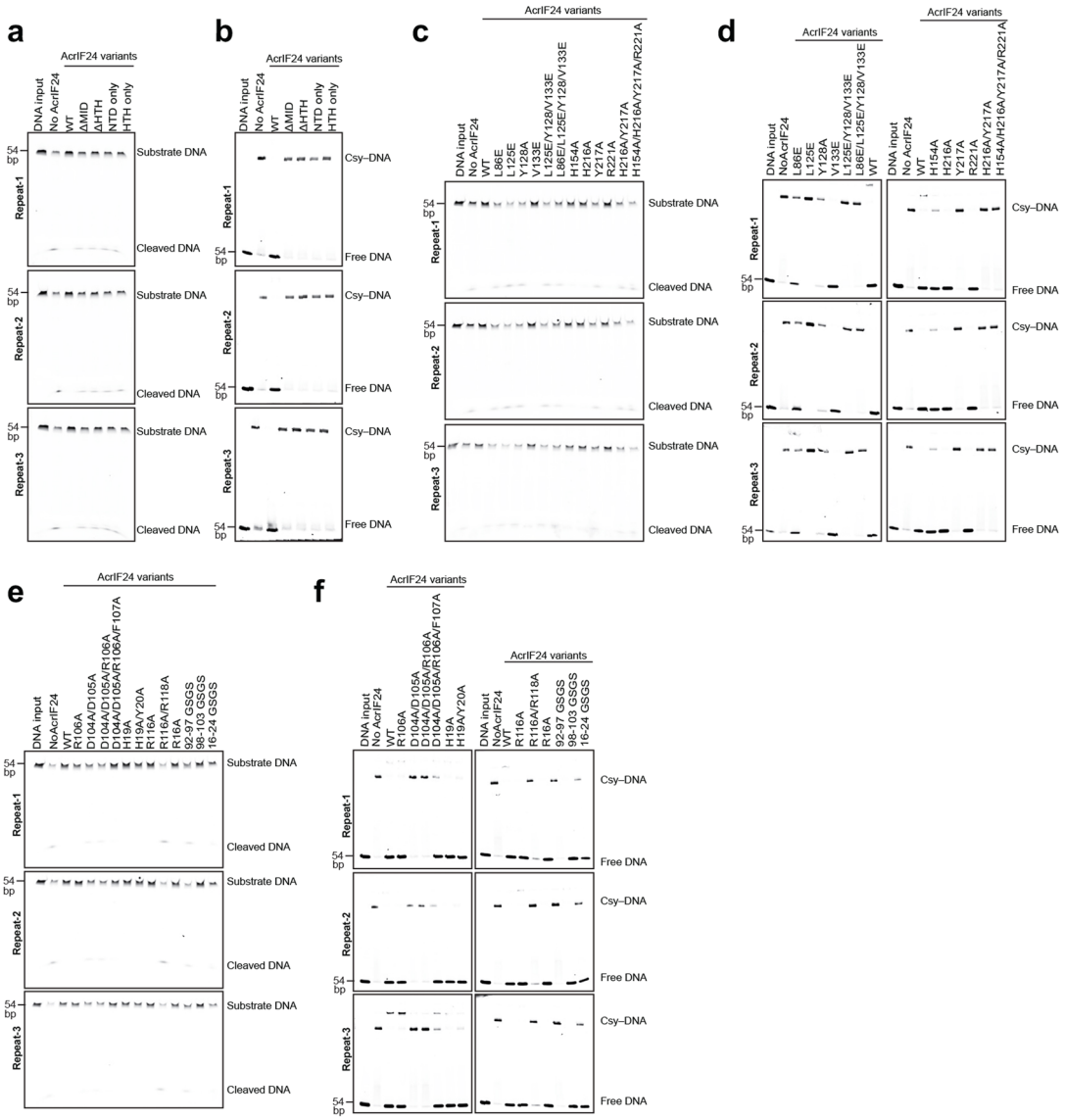


Extended Data Fig. 3 | See next page for caption.

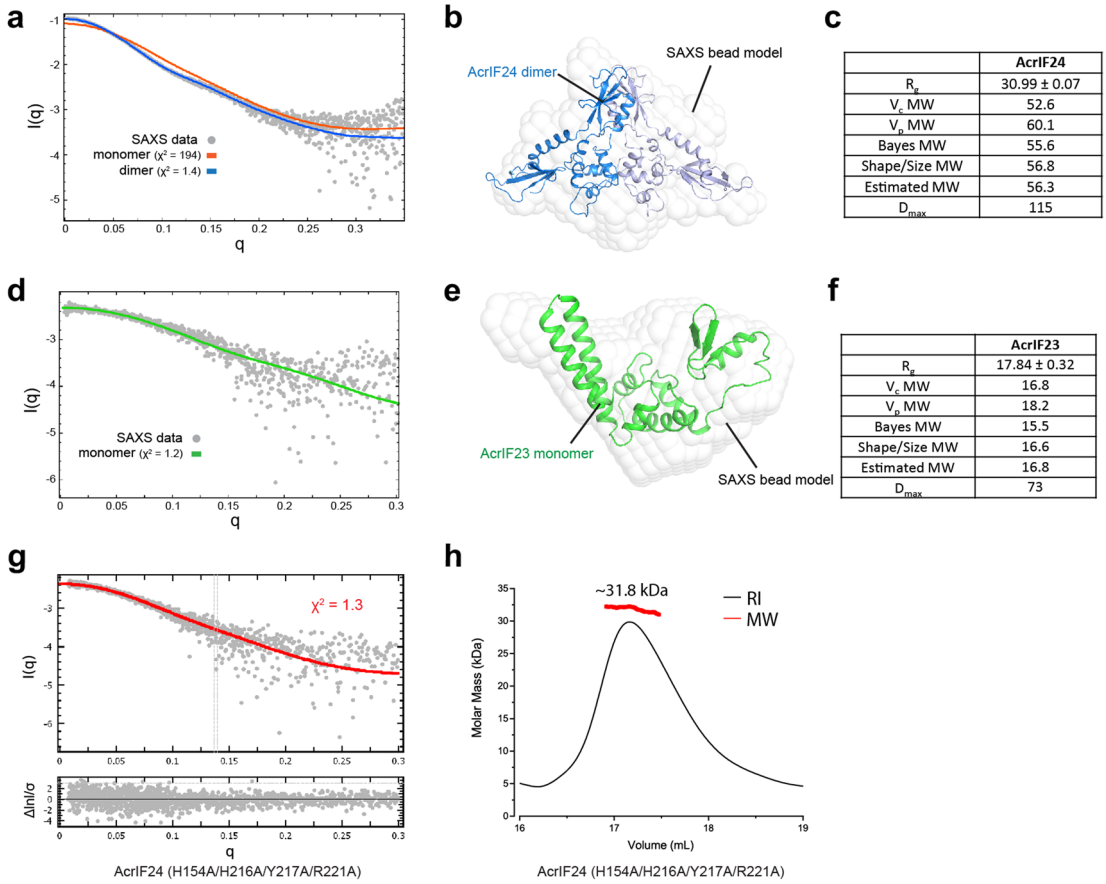
Extended Data Fig. 3 | Cryo-EM data processing workflow for Csy–AcrIF24. (a) A representative raw cryo-EM micrograph of Csy–AcrIF24. **(b)** Representative 2D class averages. **(c–e)** Three major 3D classes were identified from 2D classification and heterogenous refinement: Csy–AcrIF24 **(c)**, Csy–AcrIF24 dimer **(d)**, and AcrIF24 only **(e)**. **(f)** Major 3D classes from the 3D classification of Csy–AcrIF24. Class 2 shows missing density for Cas5f, and Cas8f subunits. **(g)** 3D refinement for particles from the 3D classification of Csy–AcrIF24 as indicated. **(h)** The plot of the global half map FSC of Csy–AcrIF24 indicates an average resolution of 3.2 Å. **(i)** Two focused refinements with a soft mask either in the head or tail of the Csy complex to improve local resolutions. **(j)** The final Csy–AcrIF24 map generated by combining maps from focused refinement. **(k)** 3D refinement for particles from the 3D classification of Csy–AcrIF24 dimer as indicated. **(l)** The plot of the global half map FSC of Csy–AcrIF24 dimer indicates an average resolution of ~ 3.5 Å. **(m)** Two focused refinements with a soft mask to each Csy complex bound to AcrIF24 to improve local resolutions. **(n)** The final Csy–AcrIF24 dimer map generated by combining maps from focused refinement.



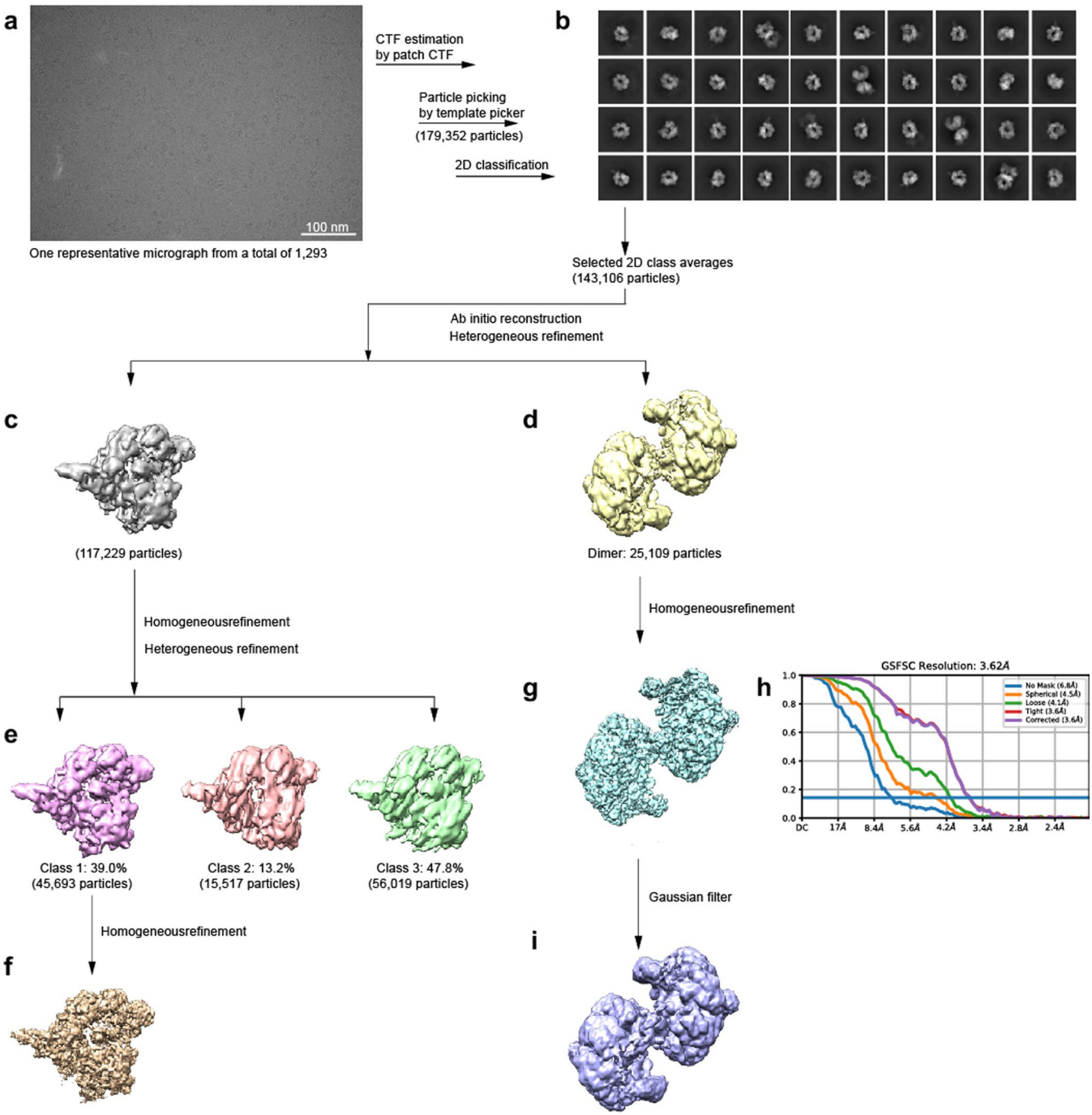
Extended Data Fig. 4 | Structure of Csy–AcrIF24. (a) Structure of AcrIF24 protomer color-coded as in Fig. 2a. EM density map is shown in orange mesh, and the atomic model of AcrIF24 is shown in sticks. **(b)** The protein sequence and secondary structure of AcrIF24. Residues involved in the dimeric interaction and tested by mutagenesis are marked with triangles. **(c)** Structural superimposition between AcrIF24 NTD domain and MMOD (PDB: 6D7K). **(d)** Interaction between HTH and NTD, with hydrophobic residues in the interface highlighted. **(e)**Structural superimposition between AcrIF24 HTH domain and Aca1 (PDB: 7CJK). **(f)** MID-MID dimerization interface of AcrIF24 as in Fig. 2e, with corresponding EM density shown in mesh. **(g)** HTH-HTH dimerization interface of AcrIF24 as in Fig. 2h, with corresponding EM density shown in mesh. **(h)** Interaction between AcrIF24-NTD and Csy as in Fig. 3c, with corresponding EM density shown in mesh. **(i)** Interaction between AcrIF24-MID and Csy as in Fig. 3e, with corresponding EM density shown in mesh.



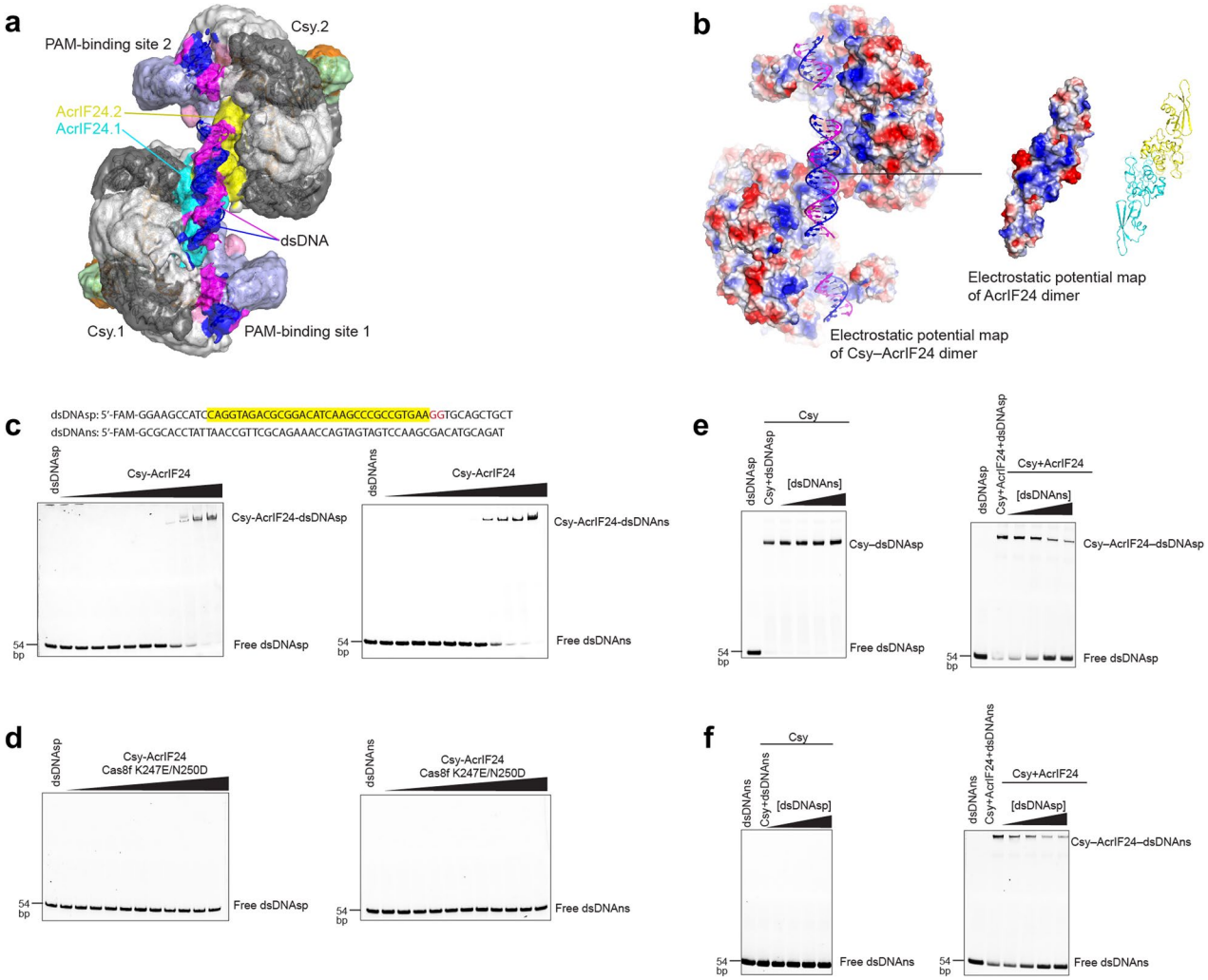
Extended Data Fig. 5 | Effect of AcrIF24 mutants on DNA cleavage and binding by Csy. (a,b) *In vitro* DNA cleavage assays **(a)** and EMSAs **(b)** with Csy and target DNA in the presence of wild-type or domain deletion mutants of AcrIF24. **(c,d)** *In vitro* DNA cleavage assays **(c)** and EMSAs **(d)** with Csy and target DNA in the presence of wild-type AcrIF24 or AcrIF24 with mutations of residues in the dimerization interfaces. **(e,f)** *In vitro* DNA cleavage assays **(e)** and EMSAs **(f)** in the presence of wild-type AcrIF24 or AcrIF24 with mutations of residues involved in Csy interaction.



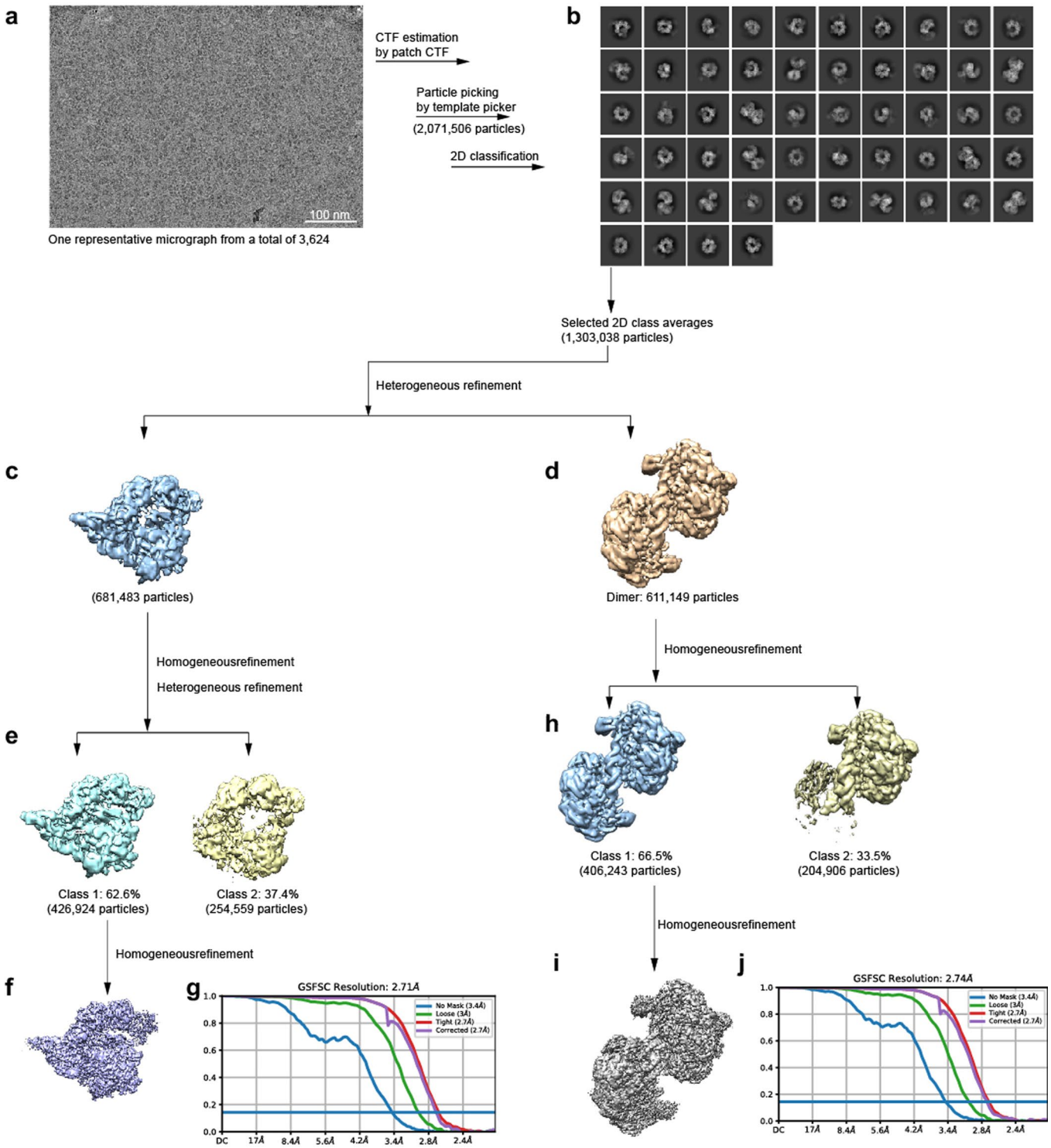
Extended Data Fig. 6 | SAXS analysis of AcrIF23 and AcrIF24. (a) SAXS assay of purified wild type AcrIF24 in solution. The experimental SAXS scattering curve for AcrIF24 is shown as gray points, while calculated curves for the monomeric (red) and dimer (blue) forms of AcrIF24 are shown as solid lines. The fit of each calculated curve to the experimental data is indicated by χ^2 values of 194 and 1.4, respectively. **(b)** A superposition of the AcrIF24 dimer with an *ab initio* bead model from the SAXS analysis (DAMMIF/N). **(c)** The 'estimated molecular weight' of AcrIF24 as the average of Vc, Vp, Bayes, and Shape/Size. **(d)** The experimental SAXS scattering curve for AcrIF23 is shown as gray points, while calculated curve for the AcrIF23 monomer model (AlphaFold) (green) is shown as a solid line. The fit of the calculated curve to the experimental data is indicated by the χ^2 value of 1.2. **(e)** A superposition of the AcrIF23 monomeric model (AlphaFold) with the *ab initio* bead model from the SAXS analysis (DAMMIF/N). **(f)** The 'estimated molecular weight' of AcrIF23 as the average of V_c, V_p, Bayes, and Shape/Size. AcrIF23 has a theoretical molecular weight of 17.5 kDa. **(g)** SEC-SAXS scattering curve (gray; D_{max} of 75.0 Å; R_g of 25.3 Å; MW_{avg} of 36 kDa) for the AcrIF24(H154A/H216A/Y217A/R221A) mutant aligned with the calculated scattering curve for the AcrIF24 monomer (red, χ^2 of 1.3). The residual plot from the fit of the experimental data (gray) and fit of the model (red) is shown in the bottom panel. **(h)** SEC-MALS analysis of the AcrIF24(H154A/H216A/Y217A/R221A) mutant showing a calculated molecular weight consistent with a monomer.



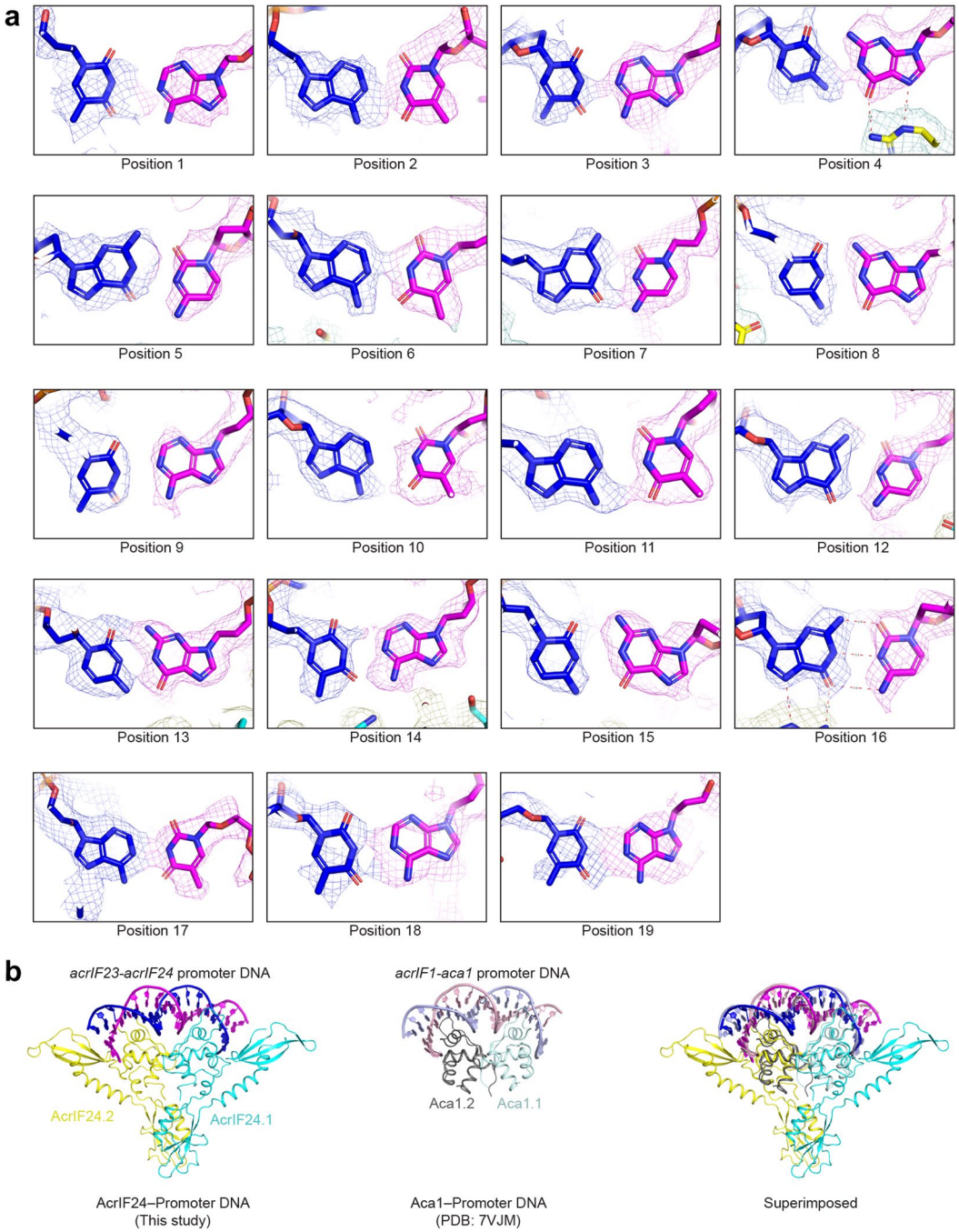
Extended Data Fig. 7 | Cryo-EM data processing workflow for Csy–AcrIF24–DNA. (a) A representative raw cryo-EM micrograph of the Csy–AcrIF24–dsDNA complex. **(b)** Representative 2D class averages. **(c,d)** Two 3D classes separated by heterogenous refinement: Csy–AcrIF24 in monomer **(c)** and dimer **(d)** states. **(e)** Three major classes from the heterogenous refinement of Csy–AcrIF24 monomer. **(f)** Refined map of Csy–AcrIF24 in monomer state. **(g)** Homogenous refinement for particles from the dimer state. **(h)** The plot of the global half map FSC of Csy–AcrIF24–dsDNA indicates an average resolution of 3.62 Å. **(i)** Refined map of Csy–AcrIF24–dsDNA after applying Gaussian filter to show the weak density of DNA.



Extended Data Fig. 8 | Csy–AcrIF24 dimer recruits DNA independent of PAM and spacer sequences. (a) Cryo-EM structure of Csy–AcrIF24 dimer bound to dsDNA. Csy–AcrIF24 subunits are color-coded as Fig. 1a. Two strands of DNA are color-coded in blue and magenta, respectively. **(b)** Electrostatic potential maps of Csy–AcrIF24 dimer (left) and AcrIF24 dimer (right), showing that DNA is bound to the positively changed surface of the assembly. **(c)** EMSA assay showing that Csy–AcrIF24 binds to both specific DNA (dsDNAsp) and non-specific DNA (dsDNAns) without a PAM or a spacer sequence. Sequences of dsDNAsp and dsDNAns are shown at the top, with the spacer highlighted and the GG PAM sequence colored in red. **(d)** EMSA assay showing that Csy (Cas8f K247E/N250D)–AcrIF24 is not capable of binding to both dsDNAsp or dsDNAns. **(e,f)** Competition EMSA assay to test whether the binding sites of dsDNAsp and dsDNAns overlap. In **e**, Csy or Csy–AcrIF24 was incubated with dsDNAsp (5′-FAM in the TS), followed by adding increasing concentrations of unlabeled dsDNAns. In **f**, Csy or Csy–AcrIF24 was incubated with dsDNAns (5′-FAM in the TS), followed by adding increasing concentrations of unlabeled dsDNAsp.



Extended Data Fig. 9 | Cryo-EM data processing workflow for Csy–AcrIF24–promoter DNA. (a) A representative raw cryo-EM micrograph of the Csy–AcrIF24–promoter DNA complex. **(b)** Representative 2D class averages. **(c,d)** Two 3D classes separated by heterogenous refinement: Csy–AcrIF24–promoter DNA in monomer **(c)** and dimer **(d)** states. **(e)** Two major classes from the heterogenous refinement of Csy–AcrIF24–promoter DNA monomer. **(f)** Refined map of Csy–AcrIF24–promoter DNA in monomer state. **(g)** The plot of the global half map FSC of Csy–AcrIF24–promoter DNA in monomer state indicates an average resolution of ~ 2.7 Å. **(h)** Two major classes from the heterogenous refinement of Csy–AcrIF24–promoter DNA dimer. **(i)** Refined map of Csy–AcrIF24–promoter DNA in dimer state. **(j)** The plot of the global half map FSC of Csy–AcrIF24–promoter DNA in dimer state indicates an average resolution of ~ 2.7 Å.



Extended Data Fig. 10 | Structure of AcrIF24–promoter DNA. (a) Modeling of the 19-bp promoter DNA bound to AcrIF24. Shown are individual base pairs of the 19-bp promoter DNA, with corresponding EM densities in mesh. **(b)** Structural comparison between AcrIF24–promoter DNA and Aca1–promoter DNA.

nature research

Corresponding author(s): Leifu Chang

Last updated by author(s): Jun 29, 2022

Reporting Summary

Nature Research 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 Research policies, see our Editorial Policies and the Editorial Policy Checklist.

Statistics

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

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

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

Software and code

Policy information about availability of computer code

Data collection	Leginon 3.5, Thermo Scientific EPU 2,
Data analysis	Relion 3.1, cryoSPARC v3.3.1, PyMOL 2.4.0, UCSF Chimera 1.13, Coot 0.9.5, Phenix 1.19, ESPript 3.0, ImageJ 1.53q, GraphPad Prism 9.2, ASTRA 7, BioXTAS RAW, ATSAS, AlphaFold2

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

Data

Policy information about availability of data

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

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

Data availability

Cryo-EM reconstructions of Csy–AcrIF24, Csy–AcrIF24 dimer, Csy–AcrIF24–DNA dimer, Csy–AcrIF24–Promoter DNA, and Csy–AcrIF24–Promoter DNA dimer have been deposited in the Electron Microscopy Data Bank under the accession numbers EMD-25660, EMD-25661, EMD-25662, EMD-25789, and EMD-25788, respectively. Coordinates for atomic models of Csy–AcrIF24, Csy–AcrIF24 dimer, Csy–AcrIF24–DNA dimer, Csy–AcrIF24–Promoter DNA, and Csy–AcrIF24–Promoter DNA dimer have been deposited in the Protein Data Bank under the accession numbers 7T3J, 7T3K, 7T3L, 7TAX, and 7TAW, respectively. Structures of Csy alone and Csy–AcrIF4 are available in the Protein Data Bank under the accession numbers 6B45 and 7JZW, respectively. Source data are provided with this paper.

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	No sample-size calculation was performed. For biochemical assays, three independent experiments were performed. For cryo-EM, the actual sample size was increased so that desired structure features (i.e. side chains) were obtained in the cryo-EM maps.
Data exclusions	Cryo-EM particle images that fall into 2D averages or 3D classes with poor features were excluded from final reconstruction.
Replication	All biochemical experiments were carried out as triplicates, and were readily and reliably reproduced.
Randomization	Randomization is not relevant for the biochemical assays in this study. For cryo-EM, randomization was internally used by data processing software in initial particle assignment for 2D classifications and generation of half maps for resolution estimation.
Blinding	Blinding was not relevant to this study. Blinding is not typically performed for the biochemical and structural studies reported in this paper.

Reporting for specific materials, systems and methods

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

Materials & experimental systems

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

Methods

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