

# Widespread gene regulator Psu inhibits transcription termination factor p by forced hyper-oligomerization

Daniela Gjorgjevikj<sup>1</sup>, Naveen Kumar<sup>2</sup>, Bing Wang<sup>3</sup>, Tarek Hilal<sup>1,4</sup>, Nelly Said<sup>1</sup>, Bernhard Loll<sup>1</sup>, Irina Artsimovitch<sup>3</sup>, Ranjan Sen<sup>2,5</sup>, Markus C. Wahl<sup>1,6,\*</sup>

<sup>1</sup> Freie Universität Berlin, Institute of Chemistry and Biochemistry, Laboratory of Structural Biochemistry, Takustr. 6, D-14195 Berlin, Germany

<sup>2</sup> Laboratory of Transcription, Centre for DNA Fingerprinting and Diagnostics, Inner Ring Road, Uppal, Hyderabad-500039, India.

<sup>3</sup> Department of Microbiology and Center for RNA Biology, The Ohio State University, Columbus, OH, USA

<sup>4</sup> Freie Universität Berlin, Institute of Chemistry and Biochemistry, Research Center of Electron Microscopy and Core Facility BioSupraMol, Fabeckstr. 36a, 14195 Berlin, Germany

<sup>5</sup> School of Biological Sciences, Indian Association for the Cultivation of Science, 2A & B, Raja S. C. Mullick Road, Jadavpur, Kolkata 700032, India

<sup>6</sup> Helmholtz-Zentrum Berlin für Materialien und Energie, Macromolecular Crystallography, Albert-Einstein-Str. 15, D-12489 Berlin, Germany

\* Correspondence to: markus.wahl@fu-berlin.de

## Abstract

Many bacteriophages modulate the host transcription machinery for efficient expression of their own genomes. Phage P4 polarity suppression protein, Psu, is a building block of the viral capsid and inhibits the hexameric transcription termination factor,  $\rho$ , by presently unknown mechanisms. We elucidated cryogenic electron microscopy structures of  $\rho$ -Psu complexes, showing that Psu dimers laterally clamp two inactive, open  $\rho$  rings and promote their expansion to higher-oligomeric states. Systematic ATPase, nucleotide binding and nucleic acid binding studies revealed that Psu hinders  $\rho$  ring closure and traps nucleotides in their binding pockets on  $\rho$ . Structure-guided mutagenesis in combination with growth, pull-down and termination assays further delineated the functional  $\rho$ -Psu interfaces. Bioinformatic analyses suggested that, in addition to guarding its own genome against  $\rho$ , Psu enables expression of diverse phage-defense systems commonly found in P4-like mobile genetic elements across bacteria. Thus, Psu is a widespread gene regulator that inhibits  $\rho$  *via* forced hyper-oligomerization.

# Main

More than 90 % of bacterial species encode  $\rho$ , a hexameric, ring-shaped RNA-dependent NTPase.<sup>1</sup>  $\rho$  is long known as a mediator of transcription termination<sup>2</sup> that defines the ends of 20-30 % of transcription units in *Escherichia coli*<sup>3</sup>.  $\rho$  also mediates attenuation in 5'-untranslated regions (UTRs)<sup>4</sup>, limits antisense transcription<sup>5,6</sup>, and safeguards bacterial genomes by restricting R-loops<sup>7</sup>. A major function of  $\rho$  is the silencing of horizontally acquired genes and invading foreign DNA<sup>8</sup>. Thus, phages that utilize the bacterial transcription machinery must counteract  $\rho$  activity to efficiently express their genomes. While lambdoid phages implement protein- or RNA-based anti-termination mechanisms to render the host RNA polymerase (RNAP) termination-resistant<sup>9-15</sup>, the enterobacterial pirate phage, P4, has acquired a unique anti-termination system that acts directly upon  $\rho$ . Its polarity suppression (Psu) protein is a building block of its capsid<sup>16</sup>, but can also inhibit  $\rho$ <sup>17,18</sup>, as first noted by its suppression of  $\rho$ -dependent polarity<sup>19</sup>. The precise mode of Psu-dependent  $\rho$  inhibition has so far remained elusive.

A  $\rho$  protomer comprises N-terminal and C-terminal domains (NTD and CTD) connected by a flexible hinge region.<sup>20,21</sup> The NTD harbors a primary RNA-binding site (PBS).<sup>20,21</sup> The CTDs together form a secondary RNA-binding site (SBS) in the central pore of the hexamer and six nucleotide-binding pockets between the subunits.<sup>20,21</sup> The  $\rho$  hexamer adopts an NTPase-inactive, washer-like, open-ring conformation until NTPs (preferentially ATP) are engaged and RNA is bound at the SBS, leading to ring closure and RNA entrapment inside the ring.<sup>22</sup> The SBS and the sixth nucleotide-binding site are fully formed only in the closed  $\rho$  ring. After RNA engagement at the PBS and the SBS,  $\rho$  can act as an NTP hydrolysis-driven 5'-to-3' RNA-translocase and helicase.<sup>20,21</sup>

Extensive research over decades suggested two mechanisms of  $\rho$  termination. In a tethered-tracking model,  $\rho$  in the open conformation engages a C-rich, secondary-structure-deficient  $\rho$ -utilization (*rut*) site on the nascent RNA (a  $\rho$ -dependent terminator) *via* its PBSes.<sup>20,21</sup>

Subsequently, downstream RNA can enter the SBS, followed by ring closure and NTPase-driven 5'-to-3' translocation on the transcript, during which the PBses are thought to remain bound to *rut*.<sup>20,21</sup> When RNAP reaches a pause signal,  $\rho$  catches up and, using its strong motor activity, disassembles the elongation complex (EC).<sup>23-25</sup> In an alternative, allosteric model, the NTDs of an open  $\rho$  ring engage ECs *via* direct contacts to RNAP and bound general transcription factors (TFs) NusA and NusG<sup>26-28</sup> and  $\rho$  traffics with the EC<sup>29</sup>. Once the EC pauses,  $\rho$  can stepwise inactivate RNAP without resorting to its motor activity.<sup>26-28</sup> Recent single-molecule spectroscopic studies have suggested that, at least *in vitro*, both mechanisms may be at work to a different extent depending on the transcription unit or termination scenario.<sup>30,31</sup>

Here, we present cryogenic electron microscopy (cryoEM)/single-particle analysis (SPA)-based structures of  $\rho$ -Psu complexes, defining the precise  $\rho$ -Psu interfaces and revealing a unique inhibitory mechanism. Multiple Psu dimers staple two open  $\rho$  complexes together and promote the formation of higher-order  $\rho$  oligomers. Psu subunits bind between two neighboring  $\rho$  CTDs across the nucleotide-binding sites. Our systematic ATPase, nucleotide-binding and nucleic-acid-binding studies are fully consistent with Psu blocking the  $\rho$  nucleotide-binding sites and stabilizing  $\rho$  in an open conformation, while leaving  $\rho$  PBses unobstructed. Furthermore, we delineate key residues on Psu and  $\rho$  required for  $\rho$  inhibition *in vitro* and *in vivo*. We also find Psu proteins encoded in the neighborhood of known and putative phage defense genes in many bacterial genomes. Based on our results, we propose that Psu inhibits  $\rho$ -dependent transcription termination *via* an unconventional molecular mechanism of forced hyper-oligomerization, which may well be employed by diverse bacteria to regulate the expression of phage defense systems.

## Results

### Psu dimers bridge and promote the expansion of two open $\rho$ hexamers

Nucleotide-free, wild-type (wt)  $\rho$  did not form a stable complex with Psu in analytical size-exclusion chromatography (SEC; Extended Data Fig. 1a). However, in the presence of ADP-BeF<sub>3</sub> or ATPyS,  $\rho^{\text{wt}}$  and Psu partially co-eluted earlier from the SEC column than either protein alone, suggesting that dynamic  $\rho^{\text{wt}}$ -Psu complexes are stabilized by the nucleotides (Extended Data Fig. 1a). To elucidate the structural basis of  $\rho^{\text{wt}}$ -ATPyS-Psu complex formation, we subjected corresponding mixtures to cryoEM/SPA (Extended Data Fig. 2 and 3; Supplementary Table 1). Multi-particle 3D refinement with ~750,000 particle images yielded several cryoEM reconstructions (Extended Data Fig. 3). The assemblies each contain two open  $\rho^{\text{wt}}$  complexes, but with varying numbers of subunits, and different numbers of Psu dimers that bridge the  $\rho^{\text{wt}}$  complexes. As exemplars, we further refined and analyzed one of the smallest and the largest assembly imaged.

In both assemblies, two open  $\rho^{\text{wt}}$  rings, referred to as  $\rho^{\text{wt}}$ (A) and  $\rho^{\text{wt}}$ (B), face each other with their CTDs, while the  $\rho^{\text{wt}}$  NTDs are facing outwards (Fig. 1a-d). Multiple Psu dimers laterally bridge the two  $\rho^{\text{wt}}$  rings. The lower-oligomeric complex (complex I; global resolution 3.65 Å) contains two open  $\rho^{\text{wt}}$  hexamers and six Psu dimers (Fig. 1a,b). Designating individual subunits of  $\rho^{\text{wt}}$ (A)/ $\rho^{\text{wt}}$ (B) by increasing Arabic numerals from the inner-most to the outer-most subunits, Psu dimers bridge  $\rho^{\text{wt}}$ (A) subunit 6 with  $\rho^{\text{wt}}$ (B) subunit 1,  $\rho^{\text{wt}}$ (A) subunit 5 with  $\rho^{\text{wt}}$ (B) subunit 2, etc. (Fig. 1b). The highest-oligomeric complex (complex II; global resolution 4.25 Å) can be envisaged to emerge from complex I by the addition of one centrally positioned  $\rho^{\text{wt}}$  subunit and two peripheral  $\rho^{\text{wt}}$  subunits to each of the  $\rho^{\text{wt}}$ (A) and  $\rho^{\text{wt}}$ (B) hexamers of complex I (Fig. 1c-e). The expansion of the  $\rho^{\text{wt}}$ (A)/ $\rho^{\text{wt}}$ (B) complexes generates two additional pairs of  $\rho^{\text{wt}}$  subunits that can be bridged by two additional Psu dimers (Fig. 1d). The most peripheral  $\rho^{\text{wt}}$  subunits of the  $\rho^{\text{wt}}$ (A)/ $\rho^{\text{wt}}$ (B) rings in complex II are not stabilized by contacts to a Psu dimer, as no appropriately-spaced, unoccupied

$\rho^{\text{wt}}$  subunit remains in the opposite  $\rho^{\text{wt}}$  ring. These peripheral subunits are less well defined in the cryoEM reconstructions, suggesting that they are more flexibly or sub-stoichiometrically connected (Fig. 1c). Due to the additional  $\rho^{\text{wt}}$  subunits incorporated at the center of complex II, the minimum spacing of the inner-most  $\rho^{\text{wt}}(\text{A})/\rho^{\text{wt}}(\text{B})$  subunits is decreased from  $\sim 50$  Å in complex I to  $\sim 10$  Å in complex II. As a consequence, the contacting Psu subunits also reciprocally interact *via* their C-terminal  $\alpha 7$  helices (Fig. 1d). Several other complexes observed upon multi-particle 3D refinement represent intermediate oligomeric states (e.g., marked by dashed boxes in Extended Data Fig. 3), supporting the picture of a dynamic ensemble of  $\rho^{\text{wt}}$ -ATPyS-Psu complexes in solution (Fig. 1e).

In isolated, open  $\rho$  hexamers (PDB ID: 1PV4)<sup>32</sup>, the average inter-subunit rise (8.6 Å) yields an overall  $\rho 1$ - $\rho 6$  rise (42.9 Å), which is too small to allow oligomerization beyond six  $\rho$  protomers. The  $\rho^{\text{wt}}$  hexamers in  $\rho^{\text{wt}}$ -ATPyS-Psu complex I exhibit slightly increased inter-subunit and  $\rho^{\text{wt}} 1$ - $\rho^{\text{wt}} 6$  rise values (9.5 Å and 46.5 Å, respectively). The incorporation of additional  $\rho$  subunits supported by additional, bridging Psu dimers in complex II is facilitated by a further increase in the inter-subunit and  $\rho^{\text{wt}} 1$ - $\rho^{\text{wt}} 6$  rise values (12.0 Å and 57.1 Å, respectively). As a consequence, the helical pitch in complex II is large enough to allow expansion by additional  $\rho^{\text{wt}}$  protomers, which can associate at both ends of the central  $\rho^{\text{wt}}$  hexamers.

## **Increased stability of a $\rho^{\text{P167L}}$ -Psu interaction is based on local refolding of the $\rho^{\text{P167L}}$ variant**

A proline at the 167-equivalent position of *E. coli*  $\rho$  is invariant across the bacterial kingdom (Extended Data Fig. 4a). A  $\rho$  variant harboring a P167L substitution has been observed to bind Psu more tightly than  $\rho^{\text{wt}}$ .<sup>33,34</sup> Consistently, we observed  $\rho^{\text{P167L}}$ -Psu complexes in analytical SEC even in the absence of nucleotides, and the more pronounced shift in the elution volume compared to  $\rho^{\text{wt}}$ -nucleotide-Psu complexes suggested the formation of more stable complexes compared to  $\rho^{\text{wt}}$  (Extended Data Fig. 1b). However, in  $\rho^{\text{wt}}$ -ATPyS-Psu structures, P167 of  $\rho$  does not engage in direct contacts to Psu that could explain these effects. To explore the structural

basis underlying the higher stability of  $\rho^{P167L}$ -Psu complexes, we also determined cryoEM structures of  $\rho^{P167L}$ -ATPyS-Psu complexes.

We obtained two main cryoEM reconstructions from ~764,000 particle images (Extended Data Fig. 5 and 6; Supplementary Table 1). The first reconstruction (P167L-complex I; global resolution 3.09 Å) represents two open hexameric  $\rho^{P167L}$  rings, similar to  $\rho^{wt}$ -ATPyS-Psu complex I, but with the  $\rho^{P167L}$ (B) hexamer rotated by one subunit position relative to the  $\rho^{P167L}$ (A) hexamer (Fig. 2a,b). As a consequence, the two opposed open  $\rho^{P167L}$  rings present only five pairs of  $\rho^{P167L}$  subunits that can be bridged by Psu dimers, which connect  $\rho^{P167L}$ (A) subunit 6 with  $\rho^{P167L}$ (B) subunit 2,  $\rho^{P167L}$ (A) subunit 5 with  $\rho^{P167L}$ (B) subunit 3, etc. (Fig. 2b). Assemblies with altered rotational orientation of Psu-bridged  $\rho^{wt}$  rings, and thus differing numbers of bridging Psu dimers, are also observed in the ensemble of the  $\rho^{wt}$ -ATPyS-Psu complexes (e.g., marked by gray boxes in Extended Data Fig. 3).

The second refined  $\rho^{P167L}$ -ATPyS-Psu assembly (P167L-complex II; global resolution 2.95 Å) resembles the first arrangement but with an additional  $\rho^{P167L}$  subunit appended to the center of the  $\rho^{P167L}$ (A) ring (Fig. 2c,d). Upon joining of this additional subunit, the helical pitch of the  $\rho^{P167L}$ (A) ring does not increase, such that the NTD of the additional  $\rho^{P167L}$  subunit in its canonical conformation would clash with the most peripheral subunit of the  $\rho^{P167L}$ (A) ring (Fig. 2d). As a consequence, the NTD of the additional  $\rho^{P167L}$  subunit is displaced and not defined in the cryoEM reconstruction. The CTD of the additional subunit in its canonical conformation would clash with the CTD of  $\rho^{P167L}$ (B) subunit 1 (Fig. 2d); nevertheless, the additional subunit can be accommodated as both CTDs refold and form domain-swapped  $\beta$ -sheet arrangements (Fig. 2d,e). Upon refolding, the Psu interaction interface is altered, preventing Psu association with the central, refolded  $\rho^{P167L}$  subunits. Consequently, no additional Psu dimers can join as observed for the highest-oligomeric  $\rho^{wt}$ -ATPyS-Psu complex II, preventing higher-level oligomerization of the  $\rho^{P167L}$  rings in this constellation.

In the refolded  $\rho^{P167L}$  protomers, the region preceding L167 (residues 141-163), encompassing the C-terminal part of the hinge region, becomes disordered, as indicated by a lack of interpretable cryoEM density (Fig. 2e). C-terminally of L167, the  $\rho^{wt}$  fold is maintained up to residue L345. In the following region up to residue Q378, helical elements are repositioned and changed in register. Between residues Q378 and N401, the helical elements of the canonical fold adopt an extended  $\beta$ -hairpin structure that constitutes the main swapped element (Fig. 2e). The swapped  $\beta$ -hairpins connect the central, six-stranded  $\beta$ -sheets of the CTDs of the domain-swapped  $\rho^{P167L}$  subunits into a continuous, 16-stranded sheet (Fig. 2e). Residues beyond N401 again lack interpretable cryoEM density.

Mutation-induced domain swapping has been described previously and is frequently associated with destabilization of the fold of the affected protein variants.<sup>35</sup> Indeed, differential scanning fluorimetry revealed that the apparent melting temperature of  $\rho^{P167L}$  is decreased by about 7 °C compared to  $\rho^{wt}$  (Fig. 2f). Thus, the P167L substitution leads to a decrease in the fold stability of  $\rho$  subunits that allows refolding of part of the CTD. As a consequence, a more stable interlocking of  $\rho^{P167L}$  complexes is possible *via* the observed domain swapping, leading to the formation of overall more stable  $\rho^{P167L}$ -Psu complexes.

Similar to  $\rho^{wt}$ -ATPyS-Psu complexes, additional  $\rho^{P167L}$ -ATPyS-Psu assemblies imaged can be considered to emerge from P167L-complexes I or II by the addition or dissociation of  $\rho^{P167L}$  subunits and Psu dimers. E.g., one such additional assembly contains two hexameric  $\rho^{P167L}$  rings interlocked by refolded, domain-swapped  $\rho^{P167L}$  subunits in the center and bridged by four Psu dimers (Fig. 2g, right); it could emerge from P167L-complex II upon dissociation of the peripheral  $\rho^{P167L}$ (A)6 subunit and loss of the associated Psu dimer (marked by a dashed box in Extended Data Fig. 6). Due to the loss of the peripheral  $\rho^{P167L}$ (A)6 subunit, the NTDs of both domain-swapped central  $\rho^{P167L}$  protomers exhibit density for NTDs in the canonical conformation in this assembly. We suggest that  $\rho^{P167L}$  prefers to form Psu-bridged complexes that can be stabilized by domain swapping of the central, refolded  $\rho^{P167L}$  protomers.



## **Psu binds across p nucleotide-binding pockets**

ATPyS is bound similarly between all neighboring p CTDs in the  $\rho^{\text{wt}}$ -ATPyS-Psu or  $\rho^{\text{P167L}}$ -ATPyS-Psu assemblies (Fig. 3a). The nucleotides are bound as previously observed for open p complexes<sup>36</sup> (Fig. 3b). Thus, the adenine base is sandwiched between F355 and M186, K181 contacts the  $\gamma$ -phosphate and T185 together with the  $\beta/\gamma$ -phosphates coordinates the  $\text{Mg}^{2+}$  ion (Fig. 3b). The  $\gamma$ -phosphate is additionally bound by R212 (arginine valve), and by R366 (arginine finger) of the adjacent p CTD (Fig. 3b). Compared to the hydrolytic situation, the catalytic glutamate (E211) and the Walker B aspartate (D265) are displaced from the ATPyS phosphates and there is no evidence for a bound attacking water molecule.

Except at the inner-most p(A/B) subunits of  $\rho^{\text{wt}}$ -ATPyS-Psu complexes, which lack a preceding p subunit, each Psu molecule binds in a similar fashion between two neighboring p CTDs in all  $\rho^{\text{wt/P167L}}$ -ATPyS-Psu assemblies (root-mean-square deviations of 0.6-2.0 Å for ~8,100 atom pairs upon global, pairwise superposition of p-p-Psu sub-complexes; Fig. 3a). Psu interacts predominantly with one p CTD *via* the tip of its long coiled-coil (helices  $\alpha 1$  and  $\alpha 2$ ), the N-terminal part of helix  $\alpha 2$  and the C-terminal helix  $\alpha 7$ . Psu helix  $\alpha 7$  (residues 170-186) is placed in a cleft between the two neighboring p CTDs, covering the corresponding nucleotide-binding pocket (Fig. 3a). The interaction is supported by shape and electrostatic complementarity (Fig. 3c).

Psu R43 at the tip of the coiled-coil engages in cation- $\pi$  interactions with Y197 of the main interacting p subunit (Fig. 3d). Psu D49 at the N-terminus of helix  $\alpha 2$  (residues 45-91) forms an electrostatic interaction with p R146 (Fig. 3d). Psu E56 (helix  $\alpha 2$ ) hydrogen bonds with the p Q374 backbone amide (Fig. 3e). Psu Q173, K180 and Q183 (helix  $\alpha 7$ ) form hydrogen bonds with the p Q378 side chain, the backbone carbonyls of p E369/T372, and the side chain of p N142, respectively (Fig. 3e). Furthermore, there is a weak electrostatic interaction between Psu R179 (helix  $\alpha 7$ ) and E218 of the adjacent p subunit (Fig. 3f). Finally, Psu P171, A172 and M175 (helix  $\alpha 7$ ) engage in contacts with the T154-loop of the neighboring p subunit and cover the Watson-

Crick side of the ATP $\gamma$ S base bound at that subunit, stabilizing it between p M186 and F355 (Fig. 3f). In the largest p<sup>wt</sup>-ATP $\gamma$ S-Psu complex II, additional contacts ensue between the dimerization region of the central Psu dimers and the NTDs of the central p subunits (Fig. 1d). The lack of conservation in the Psu-interacting residues in p across all bacteria (Extended Data Fig. 4a; Supplementary Dataset 1A) must be one of the reasons why the phage P4 host range is restricted to *Enterobacteriaceae*; ~99 % of *psu* genes are found in *Enterobacteriaceae* (Supplementary Dataset 1B). The conservation pattern of p-interacting residues in Psu supports the functional importance of helix  $\alpha$ 7-p contacts revealed by the structure: Psu residues P171, A172 and Q173 are highly conserved, and only K or R is found at position 180 (Extended Data Fig. 4b).

## **p-contacting residues of Psu are required for complex formation and inhibition of p ATPase activity**

Effects of previously investigated Psu and p variants<sup>16,19,33,37,38</sup> are fully reconciled by our structures (Supplementary Table 2). To further test the importance of specific contacts observed in our structures, we exchanged Psu residues R43 and D49 (contacts to p Y197 and R146, respectively; Fig. 3d), to alanine and arginine, respectively, and tested the binding of these variants to p<sup>P167L</sup> that forms complexes with Psu<sup>wt</sup> independent of added nucleotides. Psu<sup>R43A</sup> showed reduced binding to p<sup>P167L</sup> in analytical SEC, as indicated by a reduced shift in elution volume compared to p<sup>P167L</sup>-Psu<sup>wt</sup>, while binding of Psu<sup>D49R</sup> was completely abolished (Fig. 3g).

We next investigated Psu effects on p ATPase activity<sup>38</sup> by measuring ATP hydrolysis by p<sup>wt</sup> and p<sup>P167L</sup>, stimulated by RNA containing the *rut* site of the strong  $\lambda t_{R1}$  terminator. Reaction mixtures contained a fixed amount of p variants and increasing amounts of Psu<sup>wt</sup> under pseudo-first-order conditions. Reactions were started by the addition of ATP, samples were taken at various time points, and nucleotides were separated by thin-layer chromatography (TLC; Source Data) and quantified. The time traces of hydrolyzed ATP could be fitted to a single exponential equation (Fig. 3h). The final percentage of hydrolyzed ATP (reaction amplitude) was comparable

for  $\rho^{\text{wt}}$  and  $\rho^{\text{P167L}}$ , but the ATP turnover rate of  $\rho^{\text{P167L}}$  was approximately doubled compared to  $\rho^{\text{wt}}$  (Fig. 3h). For both  $\rho$  variants, addition of increasing amounts of  $\text{Psu}^{\text{wt}}$  ( $\text{Psu}^{\text{wt}}$  dimer: $\rho$  hexamer 10:1 and 50:1) led to the stepwise decrease of the  $\rho$  ATPase rate (1.5-fold and 2-fold, respectively, for  $\rho^{\text{wt}}$ ; 2-fold and 5-fold, respectively, for  $\rho^{\text{P167L}}$ ) and amplitude (by 20 % and 70 %, respectively, for  $\rho^{\text{wt}}$ ; by 55 % and 70 %, respectively, for  $\rho^{\text{P167L}}$ ; Fig. 3h).  $\text{Psu}^{\text{wt}}$  inhibited RNA-stimulated ATPase activities of  $\rho^{\text{wt}}$  and  $\rho^{\text{P167L}}$  with comparable  $IC_{50}$  values (0.9  $\mu\text{M}$  and 0.8  $\mu\text{M}$ , respectively; Fig. 3i).  $\text{Psu}^{\text{R43A}}$  was partially defective, with an  $IC_{50}$  value 10-fold higher (12.9  $\mu\text{M}$ ) than for  $\text{Psu}^{\text{wt}}$  (Fig. 3i), whereas  $\text{Psu}^{\text{D49R}}$  was fully defective in inhibiting the  $\rho^{\text{P167L}}$  ATPase, and neither  $\text{Psu}$  variant significantly affected the ATPase activity of  $\rho^{\text{wt}}$  (Fig. 3i). Together, the above results confirm the importance of specific  $\text{Psu}$  residues in establishing contacts to  $\rho$  that inhibit the RNA-stimulated ATPase activity of  $\rho$ , fully consistent with our structures.

## **Psu effectively inhibits nucleotide trafficking on $\rho$**

Previously, a reduced ATP affinity of  $\rho$  was observed in the presence of  $\text{Psu}$  based on a cross-linking assay.<sup>38</sup> As  $\text{Psu}$  helix  $\alpha 7$  covers nucleotide-binding sites in our  $\rho^{\text{wt/P167L}}$ -ATPyS- $\text{Psu}$  structures (Fig. 3a,b),  $\text{Psu}$  may hinder nucleotide binding to and release from  $\rho$ . To examine this possibility, we monitored *mant*-ATPyS binding *via* Förster resonance energy transfer (FRET) from  $\rho$  W381 in the vicinity of the nucleotide-binding pockets in a stopped-flow device. To this end, we quickly mixed constant amounts of  $\rho$  variants or pre-assembled  $\rho^{\text{P167L}}$ - $\text{Psu}$  complex with increasing concentrations of *mant*-ATPyS at 18 °C. For each *mant*-ATPyS concentration, the apparent rate constants ( $k_{\text{app}}$ ) of binding were determined by fitting the time traces of fluorescence changes to a single exponential equation (Fig. 4a). The  $k_{\text{app}}$  values were plotted against the nucleotide concentrations and fitted to a linear regression function to estimate the association rate constant ( $k_{\text{on}}$ ; slope) and the dissociation rate constant ( $k_{\text{off}}$ , y-intercept; Fig. 4b). Dissociation constants ( $K_d$ ) were calculated as the ratios of the  $k_{\text{off}}$  and  $k_{\text{on}}$  values (Fig. 4b).

In the absence of Psu,  $\rho^{\text{wt}}$  and  $\rho^{\text{P167L}}$  exhibited similar  $K_d$ 's for *mant*-ATPyS (9.3  $\mu\text{M}$  and 14.0  $\mu\text{M}$ , respectively; Fig. 4b). The slightly lower *mant*-ATPyS affinity of  $\rho^{\text{P167L}}$  is due to a two-fold higher dissociation rate (Fig. 4b), which likely arises because the residue substitution renders the nucleotide-binding site more dynamic, as indicated by the lower thermal stability of  $\rho^{\text{P167L}}$  (Fig. 2f). The increased nucleotide dissociation constant may enable faster ATP turnover, consistent with the observed elevated ATPase rate of  $\rho^{\text{P167L}}$  compared to  $\rho^{\text{wt}}$  (Fig. 3h).

The time-dependent change in the FRET signal observed with  $\rho^{\text{P167L}}$ -Psu was almost five-fold stronger than for  $\rho^{\text{P167L}}$  alone (Fig. 4b). We attribute this effect to bound Psu, which contributes additional FRET donors (W20, W60) in the vicinity of the  $\rho^{\text{P167L}}$  nucleotide-binding pockets, close to or below the tryptophan-*mant* Förster radius of  $\sim 25 \text{ \AA}$ <sup>39</sup>, thus enhancing the energy transfer to the *mant*-moiety. In cryoEM reconstructions of  $\rho^{\text{wt/P167L}}$ -ATPyS-Psu complexes, we observed density features suggesting covalent attachment of ATPyS to C117 of Psu upon prolonged incubation. In principle, a similar covalent attachment of *mant*-ATPyS could contribute to the observed increase in FRET signal, as Psu residues W20, W60 and W133 reside within a distance of  $\sim 25 \text{ \AA}$  or less of C117. However, mixing of Psu and *mant*-ATPyS did not lead to a significant FRET signal in the timeframe of the experiment.

The  $\rho^{\text{P167L}}$ -Psu complex exhibited a  $\sim$ seven-fold higher *mant*-ATPyS affinity compared to  $\rho^{\text{P167L}}$  alone ( $K_d = 2.1 \mu\text{M}$ ; Fig. 4b). The *mant*-ATPyS association and dissociation rates of the complex were more than five-fold and  $\sim$ 40-fold lower, respectively, compared to  $\rho^{\text{P167L}}$  alone (Fig. 4b). The drastic decrease of the *mant*-ATPyS dissociation rate in the presence of Psu suggests that, once nucleotide is bound, Psu effectively hinders its release.

Consistent with our observations that  $\rho^{\text{wt}}$  does not form stable complexes with Psu in the absence of nucleotides (Extended Data Fig. 1a), we did not observe Psu-dependent changes in the FRET signal when quickly mixing *mant*-ATPyS with  $\rho^{\text{wt}}$ /Psu at 18 °C. We, therefore, increased the reaction temperature to 37 °C to accelerate binding events and acquired longer time traces to monitor possible slower interactions. Under these conditions, we observed an expected

temperature-dependent increase of the  $k_{app}$  for *mant*-ATPyS binding (about four-fold compared to 18 °C; Fig. 4c, first panel). In the presence of Psu, two phases could be clearly distinguished and the time-dependent FRET signal could be fitted to a double-exponential function (Fig. 4c, second panel). The apparent rate of the fast phase ( $k_{app1}$ ) was equivalent to that of *mant*-ATPyS binding to  $\rho^{wt}$  alone, indicating *mant*-ATPyS binding to unmodified  $\rho^{wt}$  (Fig. 4c, first and second panels). As the slow phase ( $k_{app2}$ ) was associated with a further increase in the FRET signal, similar to what was observed with  $\rho^{P167L}$  in the presence of Psu, it likely monitored the subsequent binding of Psu to the  $\rho^{wt}$ -*mant*-ATPyS complex. To test this notion, we pre-assembled  $\rho^{wt}$ -*mant*-ATPyS complex and quickly mixed it with Psu in the stopped-flow instrument. The resulting time-dependent FRET signal could be fitted with a single exponential equation, which indicated an apparent rate of Psu binding in perfect correspondence to the slow phase ( $k_{app2}$ ) of the  $\rho^{wt}$ /Psu vs. *mant*-ATPyS measurement (Fig. 4c, second and third panels). Thus, the  $k_{app}$  values derived from the individual  $\rho^{wt}$  to *mant*-ATPyS and  $\rho^{wt}$ -*mant*-ATPyS to Psu binding assays are equivalent to the fast rate constant,  $k_{app1}$ , and the slow rate constant,  $k_{app2}$ , respectively, derived from the double exponential fit of the  $\rho^{wt}$ /Psu to *mant*-ATPyS binding assay. Finally, we mixed an excess of non-labeled ATPyS with pre-formed  $\rho^{wt}$ -*mant*-ATPyS or  $\rho^{wt}$ -*mant*-ATPyS-Psu complexes to quantify the *mant*-ATPyS dissociation rates. Presence of Psu decreased the apparent dissociation rate constant more than 20-fold (Fig. 4d).

Collectively, these results confirm the higher affinity of  $\rho^{P167L}$  to Psu compared to  $\rho^{wt}$ ; corroborate that Psu binding to  $\rho^{wt}$ , but not to  $\rho^{P167L}$ , is dependent upon prior nucleotide binding to  $\rho^{wt}$ ; and support the notion that Psu helix  $\alpha 7$  obstructs access to, and exit from, the  $\rho$  nucleotide-binding pockets. The reduced nucleotide binding and release in the presence of Psu likely contributes to the observed Psu-mediated inhibition of the  $\rho$  ATPase.

## **Psu does not affect RNA binding at PBSes but inhibits SBS binding of $\rho^{P167L}$**

Stable RNA binding at the  $\rho$  SBS and RNA-stimulated ATPase activity of  $\rho$  require ring closure.<sup>22</sup> In our  $\rho^{wt/P167L}$ -ATPyS-Psu complex structures,  $\rho$  invariably adopts an open ring conformation, and we showed that Psu inhibits  $\rho$  ATPase (Fig. 3), suggesting that Psu may also interfere with RNA-SBS interactions. In contrast,  $\rho$  NTDs are invariably turned outwards in the structures, in principle granting RNAs access to the  $\rho$  PBSes. In agreement with the latter notion, Psu did not affect  $\rho$  binding to the PBS ligand poly(dC).<sup>38</sup>

To test potential effects of Psu on  $\rho$ -nucleic acid interactions in more detail, we used fluorescence anisotropy assays.<sup>40</sup>  $\rho$  PBSes bind pyrimidine-rich DNA or RNA while the  $\rho$  SBS exclusively engages RNA.<sup>41</sup> Thus, for probing PBS binding, we used a fluorescein-labeled DNA oligomer, dC<sub>15</sub>-FAM. For testing RNA binding to the SBS, we used a fluorescein-labeled RNA oligomer, rU<sub>12</sub>-FAM, after saturating the PBSes with unlabeled dC<sub>15</sub>. ADP-BeF<sub>3</sub> was included in the assays to stabilize  $\rho$ -Psu interactions. PBS and SBS affinities to the respective nucleic acid probes were very similar for  $\rho^{wt}$  and  $\rho^{P167L}$  (Extended Data Fig. 7a-d). Increasing concentrations of Psu did not affect dC<sub>15</sub>-FAM binding at the PBSes of either  $\rho$  variant (Extended Data Fig. 7a,b), consistent with our cryoEM structures revealing unobstructed PBSes in the Psu-bridged  $\rho$  complexes. However, the prior addition of Psu to  $\rho^{P167L}$ -dC<sub>15</sub> inhibited rU<sub>12</sub>-FAM binding at the SBS (Extended Data Fig. 7d;  $IC_{50} = 2.8 \mu M$ ), while Psu failed to displace rU<sub>12</sub>-FAM from pre-formed  $\rho^{P167L}$ -dC<sub>15</sub>-rU<sub>12</sub>-FAM complexes (Extended Data Fig. 7e). In contrast, Psu inhibited the  $\rho^{wt}$  SBS binding only very weakly (Extended Data Fig. 7c;  $IC_{50} = 299 \mu M$ ). Together with the effect of Psu on RNA-stimulated  $\rho$  ATPase, these observations suggest that while Psu stabilizes the open conformation of  $\rho^{wt}$  or  $\rho^{P167L}$ , it still allows binding of RNA at the center of the open  $\rho^{wt}$  complexes, but inhibits RNA binding to the center of open  $\rho^{P167L}$  complexes. Possibly, the refolded central  $\rho^{P167L}$  subunits observed in  $\rho^{P167L}$ -ATPyS-Psu complexes limit access of RNA to the centers of the corresponding Psu-bound  $\rho^{P167L}$  complexes.

### **Psu variants defective in $\rho$ binding exhibit anti-termination defects *in vivo***

To test the importance of interactions observed in our  $\rho^{\text{wt/P167L}}$ -ATP $\gamma$ S-Psu complex structures in cells, we monitored the effects of selected  $\rho$  and Psu variants under *in vivo* conditions. Apart from interaction-deficient Psu<sup>R43A</sup> and Psu<sup>D49R</sup> (Fig. 3d,g), we constructed Psu<sup>K180A</sup>, in which hydrogen bonds of K180 with E369 and T372 of  $\rho$  are abolished (Fig. 3e). We also generated  $\rho^{\text{Y197A}}$ , in which the cation- $\pi$  stacking of Y197 with R43 of Psu is not possible (Fig. 3d). Selected, previously investigated Psu and  $\rho$  variants<sup>33,37,38</sup> served as controls: Psu<sup>E56K</sup>, which cannot sustain an interaction with the backbone amide of Q374 (Fig. 3e); Psu<sup>F169V</sup>, in which the proper positioning of helix  $\alpha$ 7 is likely compromised; and  $\rho^{\text{R146E}}$ , in which a salt bridge to Psu D49 is destroyed (Fig. 3d).

To monitor potential interaction defects of  $\rho$  or Psu variants *in vivo*, we co-produced  $\rho$  variants and N-terminally His<sub>6</sub>-tagged Psu variants in the *E. coli* BL21(DE3) strain, prepared cell lysates of mid-log phase cultures, and conducted pull-down assays using Ni<sup>2+</sup>-NTA beads, as described previously<sup>33,38</sup>. About 45 % of  $\rho^{\text{wt}}$  was pulled down *via* Psu<sup>wt</sup>, while all tested Psu variants strongly reduced co-precipitation (< 8%; Fig. 5a). Thus, all Psu variants expected to be defective in  $\rho$  binding based on our structures indeed failed to bind  $\rho^{\text{wt}}$  efficiently *in vivo*. Likewise, only about 5 % or about 25 % of  $\rho^{\text{R146E}}$  or  $\rho^{\text{Y197A}}$ , respectively, were co-precipitated by Psu<sup>wt</sup> (Fig. 5b), consistent with  $\rho$ -Psu contacts observed in our structures.

Over-production of Psu<sup>wt</sup> or of Psu variants that bind and inhibit  $\rho$  leads to severe growth and  $\rho$ -dependent termination defects.<sup>38,42</sup> We, therefore, produced Psu<sup>wt</sup> and variants from an IPTG-inducible *P<sub>tac</sub>* promoter in *E. coli*, and spotted serial dilutions of saturated cultures in the absence or presence of 50  $\mu$ M or 100  $\mu$ M IPTG. Production of Psu<sup>wt</sup> caused severe growth defects, whereas the growth of most strains that produced Psu variants defective in  $\rho$  binding was unaffected (Fig. 5c). Only Psu<sup>R43A</sup> production led to mild growth defects in the presence of 100  $\mu$ M IPTG (Fig. 5c), consistent with only partial disruption of the  $\rho^{\text{P167L}}$  interaction in analytical SEC (Fig. 3g) and partial inhibition of  $\rho$  ATPase activity by this variant *in vitro* (Fig. 3i)



Next, we co-transformed the MG1655 strain that harbors a chromosomal deletion of the *rho* gene with the *P<sub>tac</sub>*-Psu<sup>wt</sup> plasmid and plasmids producing  $\rho$  variants. Upon induction of Psu<sup>wt</sup> by 50  $\mu$ M IPTG, strains producing  $\rho^{R146E}$  or  $\rho^{Y197A}$  showed no or reduced growth defects, respectively (Fig. 5d). These results are fully in line with the lost or partially reduced interactions of Psu<sup>wt</sup> with  $\rho^{R146E}$  or  $\rho^{Y197A}$  suggested by our structures and observed in pull-down assays (Fig. 5b).

We then investigated effects on  $\rho$ -dependent termination *in vivo* using a reporter system in which the *lacZ* gene was positioned downstream of a strong  $\rho$ -dependent terminator,  $\lambda t_{R1}$  (Fig. 5e). As  $\beta$ -galactosidase production from the reporter depends on the suppression of  $\rho$ -dependent termination, colonies will appear white or pale blue on LB-X-gal plates if  $\rho$  is functional, and will appear blue-green if  $\rho$  function is inhibited. When Psu<sup>wt</sup> was expressed in the reporter strain, colonies appeared blue-green, indicating inhibition of  $\rho$ . Consistent with partial defects in  $\rho$  binding observed *in vitro* (Fig. 3g) and in pull-down assays (Fig. 5a), Psu<sup>R43A</sup>-producing colonies were light blue, whereas production of all Psu variants defective in  $\rho$  binding led to white or pale blue colonies (Fig. 5e).

Finally, *E. coli* cells with the *lacZ* reporter were transformed with plasmids expressing  $\rho^{wt}$  or variants, and subsequently, the chromosomal *rho* gene was deleted. The resulting strains were then transformed with plasmids encoding Psu<sup>wt</sup> (or empty plasmid as control), and the colonies were streaked on LB-X-gal plates. Colonies producing  $\rho^{wt}$  in the presence of Psu<sup>wt</sup> appeared blue-green, indicating Psu<sup>wt</sup>-mediated inhibition of  $\rho^{wt}$ -dependent termination (Fig. 5f). In contrast, colonies producing  $\rho^{Y197A}$  or  $\rho^{R146E}$ , defective in Psu<sup>wt</sup> binding based on our structures, appeared pale blue (Fig. 5f), consistent with these  $\rho$  variants being partially resistant to inhibition by Psu<sup>wt</sup>.

#### **P4-like mobile elements encoding Psu are hubs for bacterial defense genes**

The role of Psu may not be limited to P4 infection. Recent studies revealed that P4-like mobile elements are hotspots for bacterial defense systems.<sup>43-45</sup> Almost half of all *E. coli* genomes were found to carry P4-like elements adjacent to anti-phage systems that included abortive infection



(Abi), restriction-modification (RM) and other defense modules.<sup>44</sup> In P4, *Psu* is encoded in the *sid- $\delta$ -*psu* operon and, consistent with *Psu* suppressing polarity, activates expression of  $\delta$  and *psu*, but not of *sid*.<sup>18</sup> As a general anti-terminator, *Psu* may also upregulate associated defense clusters by conferring protection against premature termination by  $\rho$ . To explore this possibility, we investigated 1,163 bacterial genomes encoding *Psu* (Extended Data Fig. 8) and used DefenseFinder<sup>46</sup> to detect putative defense systems in the *psu* neighborhood. This analysis revealed a known defense system inserted between the *psu* and *int* (integrase) genes of P4 in more than a quarter (247) of *psu*<sup>+</sup> genomes (Fig. 6a), consistent with acquisition through horizontal transfer<sup>44</sup>. These defense systems are compact (~3 kb), presumably reflecting the P4 genome size limitations, and astonishingly diverse, representing 37 known classes (Fig. 6b; Extended Data Fig. 8; Supplementary Dataset 1CD). Interestingly, one system, RM type II, is found in *Enterococcus faecalis*, a Firmicute (Fig. 6c); the latter example likely represents a very recent horizontal transfer event because of the near identity of the *Psu* protein to its P4 homolog.*

The catalog of known defense systems is far from complete. Bioinformatic screens based on gene syntax identified many novel putative defense gene clusters, which have been shown to encode active defense elements<sup>47-49</sup>. We thus used analogous guilt-by-association criteria to search for putative defense systems in the *psu-int* region. This analysis revealed 659 inserts that were not identified by DefenseFinder (Fig. 6a; Supplementary Dataset 1E) but share the compact size of known defense systems (Fig. 6c). An HMM search against the Pfam database for the proteins encoded in these loci identified 767 hits from 107 Pfam families that contain nucleases, methylases, toxin-antitoxin and Abi systems, and other components commonly present in phage-defense systems (Supplemental Dataset 1). This analysis suggests that a considerable fraction of the P4 "dark matter" is also involved in transferring defense genes across host genomes, and that new P4-linked defense systems can be found *in silico*, followed by experimental validation. For example, our analysis identified a sirtuin (Sir2)-HerA helicase module adjacent to *Psu* in *Klebsiella pneumoniae* (Fig. 6c). Although this module was not recognized by DefenseFinder, an

analogous Sir2-HerA system from *Paenibacillus* sp. has been recently shown to trigger abortive infection and cell death<sup>50</sup> via the Sir2-mediated depletion of the cellular pool of NAD<sup>+</sup>, a signature activity of Sir2 in different defense contexts<sup>50-52</sup>.

Many defense genes are codirectional with *Psu* (Fig. 6c), potentially extending the *psu* operon and falling under its cognate polarity control<sup>18</sup>. However, given that *Psu* works in *trans*, including on chromosomal genes, it may activate any gene that is subject to  $\rho$ -mediated termination, a common mechanism of xenogeneic silencing<sup>8</sup>.

## Discussion

*P4* is a paradigmatic satellite phage that relies on its helper phage, *P2*, to produce functional virions. The *P4*-encoded *Psu* plays two vital roles in *P4* infection: it stabilizes the Sid-hijacked *P2* capsid proteins to selectively package the *P4* DNA<sup>53</sup>, and it inhibits  $\rho$ -mediated silencing of phage genes<sup>18</sup>. Here, we elucidated the structural basis and molecular mechanisms underlying the *Psu*-mediated inhibition of  $\rho$ . Our cryoEM structures revealed that *Psu* instigates hyper-oligomerization of  $\rho$  on two levels. First, multiple *Psu* dimers cross-strut two open  $\rho$  complexes; and second, lower-level  $\rho$ -*Psu* complexes can incorporate additional  $\rho$  subunits, expanding  $\rho$  to at least the nonamer stage, in part aided by the concomitant incorporation of additional *Psu* dimers.

During capsid morphogenesis, two *P4* proteins, *Psu* and its distant paralog *Sid*, constrict the *P2* capsid<sup>53</sup>; the *Psu*/*Sid*-modified capsid can only accommodate a much smaller *P4* genome, one-third the length of *P2* (11 kb versus 33 kb). In a high-resolution reconstruction of the *P4* pro-capsid, *Sid* dimers are seen bound across hexamers of gpN, the *P2* capsid protein.<sup>54</sup> A low-resolution structure of the mature *P4* capsid, in which *Psu* replaces *Sid*, suggested that the structurally similar *Psu* dimer likewise caps gpN hexamers.<sup>55</sup> Guided by the *Psu*-gpN interaction as well as by mutational and cross-linking studies, a model of a  $\rho$ -*Psu* complex has been proposed, in which a *Psu* dimer also caps a closed  $\rho$  ring.<sup>33</sup> Our cryoEM structures of  $\rho$ -ATPyS-

Psu complexes revealed a decisively different mode of  $\rho$ -Psu complex formation, fully in line with previous (Supplementary Table 2) and present mutational analyses of the  $\rho$ -Psu interaction, the estimated ~1:1 subunit stoichiometry in  $\rho$ -Psu complexes<sup>33,42</sup>, and additional biochemical data presented here. First, Psu blocks the nucleotide-binding pockets on  $\rho$  – we show that Psu reduces the binding and release of nucleotides to and from  $\rho$  (Fig. 4). Second,  $\rho$  adopts an open, inactive conformation in all complex structures with Psu – we demonstrate that Psu efficiently inhibits  $\rho$  ATPase activity (Fig. 3). Third, in all our structures the NTDs of the two Psu-bridged  $\rho$  complexes are facing outwards – we show that Psu does not inhibit nucleic acid binding to the  $\rho$  PBSes (Extended Data Fig. 7a). Finally, our functional analyses addressing the effects of Psu and  $\rho$  variants on  $\rho$ -Psu interaction in cells, cell viability, and  $\rho$ -mediated termination *in vivo* (Fig. 5) are fully consistent with our cryoEM structures.

Psu-mediated inhibition of  $\rho$  ring closure would inhibit the  $\rho$  translocase activity, and thus  $\rho$ -mediated termination *via* the tethered tracking pathway<sup>25</sup>. Open and expanded  $\rho$  complexes bridged by Psu would likely also be incompatible with  $\rho$  transferring NusA/NusG-modified ECs into a pre-termination state, as recently imaged<sup>27,28</sup>. While  $\rho$ -Psu complexes might still engage NusA/NusG-ECs, Psu would most likely hinder the conformational changes needed to inactivate RNAP or inhibit late stages of termination/hybrid unwinding, when  $\rho$  is expected to engage and translocate on RNA<sup>27</sup>. We posit that the Psu-mediated blockade of the nucleotide-binding pockets of  $\rho$  contributes to the inhibition of the  $\rho$  ATPase activity. Interestingly, ATPyS (and thus presumably ATP) supports Psu binding to  $\rho$  (Supplementary Fig. 1). We attribute this effect of ATPyS (ATP) to the stabilization of the Psu contact surface on the  $\rho$  CTDs and direct contacts of Psu to residues surrounding the nucleotide-binding pockets (Fig. 3f). Once bound, however, Psu inhibits the  $\rho$  ATPase activity. Thus, Psu exhibits typical features of an uncompetitive inhibitor of the  $\rho$  ATPase.

$\rho$  activity is regulated on multiple levels. On the one hand, interactions with nucleotides, RNAs, or NusG shift the equilibrium between open and closed hexameric states and thereby regulate  $\rho$

ATPase and RNA translocase activities.<sup>22,56</sup> On the other hand, RNAP-associated NusA<sup>57,58</sup>, an RNAP-trailing ribosome<sup>59-61</sup>, RNAP-bound elements of nascent RNA<sup>14,15</sup>, RNAP-bound specialized transcription factors<sup>11-13</sup>, or multi-factorial RNAP-bound RNA-protein complexes<sup>9,10</sup> can delay or prevent the attack of ECs by  $\rho$ . Furthermore, the RNA chaperone Hfq can counteract  $\rho$  by depositing  $\rho$ -inhibitory sRNAs on nascent transcripts.<sup>4,62</sup> Results presented here underscore an additional mechanism *via* which  $\rho$  can be regulated, *i.e.*, factor-dependent modulation of its oligomerization state. This mechanism capitalizes on the intrinsic oligomerization dynamics of  $\rho$ , with the hexamer representing the only state in which  $\rho$  can terminate transcription. Depending on the concentration and conditions,  $\rho$  can dissociate into non-functional, lower oligomeric states, down to the monomer level, or form dodecamers, albeit exhibiting a different configuration, as seen in complexes with Psu.<sup>63-65</sup> While some proteins, such as *Vibrio cholerae* YaeO/Rof, may stabilize  $\rho$  in a hypo-oligomerized state<sup>66</sup>, Psu stabilizes hyper-oligomerized, inactive states. Apart from interacting proteins, some  $\rho$  orthologs have acquired additional regions or domains that facilitate  $\rho$  regulation *via* modulation of its oligomeric state: *Clostridium botulinum*  $\rho$  has an additional, N-terminal prion-like domain that promotes the formation of amyloid-like aggregates<sup>67</sup>, while an intrinsically disordered region in *Bacteroides thetaiotaomicron*  $\rho$  can promote phase separation associated with increased termination activity<sup>68</sup>. We posit that condition-dependent activation or inactivation of  $\rho$  through changes in its oligomeric or aggregation state may be a widespread regulatory strategy.

Modulation of the oligomerization state as a regulatory mechanism has been observed in other molecular systems, including in eukaryotes, and is of growing interest among drug developers.<sup>69</sup> For example, plant or mammalian ASPL/PUX1 proteins can disassemble the essential and multi-functional AAA+ ATPase CDC48/p97<sup>70-72</sup>, whereas biparatopic designed ankyrin repeat proteins can induce apoptosis by oligomerization of human epidermal growth factor receptor-2 in nanoscopic membrane domains<sup>73</sup>. The development of oligomerization-modulating compounds can be fueled by structural knowledge of the target protein in different oligomeric states dependent

on a cellular interaction partner. Inhibitory peptides (shiftides) of HIV-1 integrase, for instance, have been designed based on the interacting lens epithelium-derived growth factor/p75, which shift HIV-1 integrase into an inactive, tetrameric state.<sup>74</sup> Results presented here could thus inspire the development of novel p-inhibitory substances. Peptides (and eventually peptide-derived compounds) designed based on p-contacting regions of Psu might not only block the nucleotide-binding pockets on p but may also facilitate p hyper-oligomerization. While short peptides or low-molecular-weight compounds will not support the dimerization of two open p rings akin to Psu, substances that bind to and increase the pitch of open p hexamers may conceivably allow expansion into inactive p filaments. Notably, Psu-derived, p-inhibiting peptides that are active *in vivo* have already been designed<sup>75</sup>, but the structural basis of their modes of action remains to be explored.

Our findings also suggest that Psu may play a hitherto unknown role in antiviral defense. Phage-defense systems encoded on mobile elements are pan-genomic “guns for hire”<sup>76</sup> that bacteria deploy when under attack by phages<sup>77</sup>. In rapid exchange through horizontal transfer, these elements shape bacterial adaptation and evolution, and can inform the development of new phage therapies. P4-like elements are integrated into about 30 % of enterobacterial genomes with a high probability for retained functionality, and many of these elements carry phage-defense systems.<sup>43-45</sup> Several of these systems have been shown to protect *E. coli* from lytic phages, including  $\lambda$ , P1, and T7, while not restricting P2<sup>44</sup>, implying the existence of mutualistic relationships between parasitic satellites and their helpers. Our *in silico* analysis suggests that Psu proteins may serve as built-in anti-terminators that insulate the associated defense genes from p-mediated silencing. Curiously, the RM systems, which dominate the bacterial defense arsenal,<sup>46</sup> are underrepresented among Psu-associated defense genes (Fig. 6b), which are instead enriched in Abi modules, such as Gabija<sup>78</sup>, retrons<sup>79</sup>, and PD-T7-2<sup>45</sup>. In fact, Psu itself is a potent toxin that, if present at high levels, triggers cell death in diverse bacteria<sup>80</sup>.

While supplying the cell with antiviral weapons, a *Psu*/defense cassette also poses a threat to its very existence. The expression of *Abi* systems must be tightly controlled to ensure that a suicide program is triggered only when the first line of cellular defense has been breached.<sup>81</sup>  $\rho$  limits the expression of the *psu* gene<sup>18</sup> and could be a key part of the silencing mechanism. Furthermore,  $\rho$  may have evolved to limit its sensitivity to *Psu*: P167L substitution stabilizes  $\rho$ -*Psu* interactions<sup>33,34</sup>, and we traced this enhanced stability to a tendency of  $\rho^{P167L}$  to refold and form domain-swapped dimers at the center of  $\rho^{P167L}$ -*Psu* complexes. Interestingly, we find an invariant proline at the 167-equivalent position in bacterial  $\rho$  proteins (Extended Data Fig. 4a). Notably,  $\rho^{P167L}$  does not exhibit defects in ATPase activity or termination *in vitro* and supports *E.coli* growth<sup>33</sup>; thus, the invariance of P167 could hint at an evolutionary pressure to limit  $\rho$ -*Psu* affinity and thereby balance *Psu*-dependent expression of phage defense genes and cellular toxicity of *Psu*.

# Methods

## Recombinant protein production and purification

Primers used for molecular cloning are listed in Supplementary Table 3. Plasmids used for protein production or as templates for *in vitro* transcription are listed in Supplementary Table 4. Genes encoding Psu variants (R43A and D49R) were produced by the single-primer quick mutagenesis method<sup>82</sup>. For protein production, *E. coli* BL21 RIL cells were transformed with corresponding plasmids.  $\rho^{wt}$ ,  $\rho^{P167L}$  and  $Psu^{wt}$  were produced and purified as previously described<sup>27,33,83</sup>. Briefly, lysates containing non-tagged  $\rho^{wt}$  or  $\rho^{P167L}$  were loaded on a heparin column equilibrated in 50 mM Tris-HCl, pH 7.5, 50 mM NaCl, 5 % (v/v) glycerol, 1 mM DTT. The proteins were eluted with a gradient to 1 M NaCl, and the fractions containing  $\rho^{wt}$  or  $\rho^{P167L}$  were pooled, concentrated and loaded on a Superdex 200 column for SEC in 10 mM Tris-HCl, pH 7.5, 150 mM NaCl, 1 mM DTT.  $Psu^{wt}$  or variants were precipitated with 25 % final concentration of ammonium sulfate for 1 h at RT, and pelleted by 15,000 x g centrifugation at 4 °C for 20 min. The pellet was resuspended in 20 mM Tris-HCl, pH 8.9, 150 mM NaCl, 5 % (v/v) glycerol, 0.1 mM EDTA, 1 mM DTT and passed through heparin and MonoQ chromatography columns to clear contaminants. The flow-through was subjected to MonoS cation-exchange chromatography, eluted with a gradient to 1 M NaCl and the fractions containing Psu variants were pooled, concentrated and loaded on a Superdex 75 column for SEC in 20 mM Tris-HCl, pH 8.0, 100 mM NaCl, 10 % (v/v) glycerol, 0.1 mM EDTA, 1 mM DTT.

## Rut RNA production

A DNA template for *in vitro* transcription, harboring a T7 RNAP promoter and the sequence encoding the  $\lambda tR1$  *rut* site, was produced from plasmid pUC18- $\lambda tR1$ -*rut* (Supplementary Table 4) by PCR with a reverse primer containing two 2'-O-methylated nucleotides at the 5'-terminus, for decreasing possible 3'-end run-off products.<sup>84,85</sup> T7 RNAP-based *in vitro* transcription was



performed as described previously.<sup>10</sup> The produced *ltR1 rut* site RNA was treated with DNase I and purified by strong anion-exchange (MonoQ, GE Healthcare) and size-exclusion chromatography (Superdex 75, GE Healthcare) in 10 mM HEPES-NaOH pH 7.5, 50 mM NaCl. The purity of the RNA was assessed by urea-PAGE, stained with ethidium bromide.

### Cryogenic electron microscopy

$\rho^{\text{wt}}$ -ATPyS-Psu and  $\rho^{\text{P167L}}$ -ATPyS-Psu complexes were freshly prepared in 10 mM TRIS-HCl, pH 7.5, 150 mM NaCl, 5 mM  $\text{MgCl}_2$  ( $\rho^{\text{wt}}$ -ATPyS-Psu), or 10 mM TRIS-HCl, pH 8.0, 120 mM KOAc, 5 mM  $\text{Mg}(\text{OAc})_2$ , 5  $\mu\text{M}$   $\text{ZnCl}_2$  ( $\rho^{\text{P167L}}$ -ATPyS-Psu), concentrated to 5.5 mg/ml using a 100 kDa ultra-centrifugal filter (Merck) and supplemented with 1 mM ATPyS. 3.8  $\mu\text{l}$  of the samples were applied to glow-discharged holey carbon R1.2/1.3 copper grids (Quantifoil Microtools) and plunge-frozen in liquid ethane using a Vitrobot Mark IV (Thermo Fisher) equilibrated at 10 °C and 100 % humidity. Micrographs were recorded on a FEI Titan Krios G3i transmission electron microscope operated at 300 kV with a Falcon 3EC detector. Movies were recorded for 40.57 s, accumulating a total electron flux of  $\sim 40 \text{ e}/\text{\AA}^2$  in counting mode at a calibrated pixel size of 0.832  $\text{\AA}/\text{px}$  distributed over 33 fractions.

### CryoEM data analysis

All image analysis steps were done with cryoSPARC (version 3.2.2)<sup>86</sup>. Movie alignment was done with patch motion correction generating Fourier-cropped micrographs (pixel size 1.664  $\text{\AA}/\text{px}$ ), and CTF estimation was conducted by Patch CTF. Class averages of manually selected particle images were used to generate an initial template for reference-based particle picking from 5,986/2,723 micrographs ( $\rho^{\text{wt}}$ -ATPyS-Psu/ $\rho^{\text{P167L}}$ -ATPyS-Psu). 1,815,462/823,078 particle images were extracted with a box size of 224 px and Fourier-cropped to 112 px for initial analysis. Reference-free 2D classification was used to select 747,325/763,878 particle images for further analysis. *Ab initio* reconstruction using a small subset of particles was conducted to generate an



initial 3D reference for subsequent classification by 3D heterogeneous refinement or 3D variability analysis. Local motion correction was applied to extract particle images at full spatial resolution with a box size of 448 px followed by global and local CTF refinement. Due to considerable structural and compositional flexibility of the complexes, classification was not straightforward. Instead, similar structures appeared multiple times during classification, which were then combined for final homogeneous refinement. For the  $\rho^{\text{wt}}$ -ATPyS-Psu sample, classes of 15,407 particle images (4.25 Å) and 73,056 particle images (3.65 Å) were finally reconstructed by non-uniform (NU) refinement. For  $\rho^{\text{P167L}}$ -ATPyS-Psu, 317,635 and 273,935 particle images were selected for final NU refinement, yielding reconstructions at global resolutions of 2.95 Å and 3.09 Å, respectively.

### Model building, refinement and analysis

Crystal structures of  $\rho$  (PDB ID: 1PV4)<sup>32</sup> and Psu (PDB ID: 3RX6)<sup>83</sup> were manually placed in the cryoEM reconstructions and each protomer was adjusted by rigid body fitting and segmental real-space refinement using Coot (version 0.9.6)<sup>87,88</sup>. For higher-oligomeric complexes, additional  $\rho$  subunits and Psu dimers were added to unoccupied regions of the cryoEM reconstructions and adjusted as above. The models were refined by iterative rounds of real space refinement in PHENIX (version 1.20\_4459)<sup>89,90</sup> and manual adjustment in Coot. The structural models were evaluated with MolProbity (version 4.5.1)<sup>91</sup>. Structure figures were prepared with PyMOL (version 2.4.0, Schrödinger) and ChimeraX<sup>92</sup>.

### Analytical size-exclusion chromatography

Analytical SEC-based interaction analyses were performed in 10 mM TRIS-HCl, pH 7.5, 150 mM NaCl, 5 mM MgCl<sub>2</sub>, 1 mM DTT. 300 pmol  $\rho^{\text{wt/P167L}}$  hexamer and 1200 pmol Psu dimer variants were mixed in a final reaction volume of 50  $\mu$ l. After incubation of the mixtures at 32 °C for 10 min, the samples were loaded on a Superose 6 increase 3.2/300 analytical size exclusion column

(Cytiva). 50  $\mu$ l fractions were collected and analyzed by SDS-PAGE and visualized by Coomassie staining.

### Differential scanning fluorimetry

Differential scanning fluorimetry was conducted in a 96-well plate in a plate reader combined with a Mx3005P thermocycler (Stratgene).  $\rho$  variants were prepared at a final concentration of 2  $\mu$ M in 10 mM Tris-HCl, pH 7.5, 150 mM NaCl, and 1 mM DTT, supplemented with 10x SYPRO orange in a final volume of 20  $\mu$ l. The temperature was increased at a rate of 1  $^{\circ}$ C/min from 25  $^{\circ}$ C to 95  $^{\circ}$ C, and the fluorescence emission was monitored at each step. The fluorescence intensity was plotted as a function of temperature. The thermal melting temperature was determined from the melting profile as the temperature at which the fluorescence intensity reached 50 % of the highest fluorescence intensity.

### ATPase assays

TLC-based ATPase assays were performed using [ $\alpha$ - $^{32}$ P]ATP (Hartmann Analytic). To quantify RNA-stimulated ATPase activity, 100 nM  $\rho$  variants, in isolation or supplemented with 1  $\mu$ M or 5  $\mu$ M Psu, were mixed with 20  $\mu$ M *rut* RNA in 10 mM Tris-HCl pH 7.5, 100 mM NaCl, 5 mM MgCl<sub>2</sub>, 1 mM DTT. The mixtures were incubated with 1 mM ATP, supplemented with [ $\alpha$ - $^{32}$ P]ATP, at 37  $^{\circ}$ C for up to 2 h. 2  $\mu$ l of the samples were withdrawn at selected time points and the reactions were quenched with 6  $\mu$ l of 100 mM EDTA, pH 8.0. 3 x 1  $\mu$ l of samples were spotted on a PEI-cellulose TLC membrane and chromatographed with 1 M acetic acid, 0.5 M LiCl, 20 % (v/v) ethanol. The corresponding ATP and ADP spots were visualized using a Storm 860 phosphorimager (GMI) and quantified using ImageQuant software (version 5.2; Cytiva). To test the effect of Psu variants on  $\rho$  ATPase activity, increasing concentrations of Psu variants, up to 15  $\mu$ M, were mixed with 100 nM  $\rho$  variants, and subsequently with 20  $\mu$ M *rut* RNA, in 10 mM Tris-HCl pH 7.5, 100 mM NaCl, 5 mM MgCl<sub>2</sub>, 1 mM DTT. The mixtures were incubated with 1 mM

ATP, supplemented with [ $\alpha$ - $^{32}$ P]ATP, at 37 °C for 1 h. Subsequently, the experiments were conducted as above. Data from two biological replicates were plotted and analyzed using Prism software (version 9.0.2; GraphPad).  $\rho^{\text{wt}}$  or  $\rho^{\text{P167L}}$  ATPase activity was calculated as the percentage of hydrolyzed ATP over time, by fitting quantified data to the equation  $A = A_0 + (A_{\text{max}} - A_0) * (1 - \exp(-k * t))$ , in which  $A_0$  is the fraction of ATP hydrolyzed at time zero;  $A_{\text{max}}$  is the the fraction of ATP hydrolyzed at infinite time; and  $k$  is the rate constant. To quantify the effect of Psu on  $\rho^{\text{wt}}$  or  $\rho^{\text{P167L}}$  ATPase activity, data were fitted to an inhibitor-vs.-response function:  $A = A_{\text{min}} + (A_{\text{max}} - A_{\text{min}}) / (1 + ([\text{inhibitor}] / \text{IC}_{50}))$ , in which  $A$  is the fraction of ATP hydrolyzed at a given Psu (inhibitor) concentration; and  $A_{\text{min}}$  and  $A_{\text{max}}$  are the fitted minimum and maximum fractions of ATP hydrolyzed.

### Stopped-flow/FRET-based analysis of nucleotide binding

Nucleotide binding studies were conducted on an SX-20MV stopped-flow/fluorescence spectrometer (Applied Photophysics). Tryptophan residues near the ATP binding pockets served as FRET donors and were excited at 280 nm. A 420 nm cut-off filter was used for detecting the emission by the acceptor, *mant*-ATPyS. All experiments were performed with 100 nM  $\rho^{\text{wt}}$  or  $\rho^{\text{P167L}}$ , 500 nM Psu and/or 2  $\mu$ M *mant*-ATPyS in 10 mM Tris-HCl, pH 7.5, 150 mM NaCl, 5 mM MgCl<sub>2</sub>. For quantifying nucleotide association, data was acquired at 18 °C or 37 °C by fast mixing of equal volumes (60  $\mu$ l) of the reactants (syringe 1,  $\rho^{\text{wt}}$  or  $\rho^{\text{P167L}}$  in isolation or in the presence of Psu; syringe 2, *mant*-ATPyS). FRET signals were recorded for up to 30 min. For detecting the binding of Psu to  $\rho^{\text{wt}}$  or  $\rho^{\text{P167L}}$  pre-bound to *mant*-ATPyS, the experiment was performed at 37 °C by rapidly mixing equal volumes (60  $\mu$ l) of the reactants (syringe 1,  $\rho^{\text{wt}}$  or  $\rho^{\text{P167L}}$  and *mant*-ATPyS; syringe 2, Psu) and the monitoring fluorescence changes over 30 min. Data were acquired in logarithmic sampling mode and visualized using the Pro-Data Viewer software package (version 2.5; Applied Photophysics). The final curves were obtained by averaging 3 to 6 individual traces and analyzed in Prism software (version 9.0.2; GraphPad). Unless stated otherwise, data were

fitted to a single exponential function:  $A = A_0 + (A_{max} - A_0) * (1 - \exp(-k_{app} * t))$ ;  $A$ , fluorescence at time  $t$ ,  $A_{max}$ , final signal;  $A_0$ , initial fluorescence signal;  $k_{app}$ , characteristic time constant. For  $\rho^{wt}$  binding to *mant*-ATP $\gamma$ S and the subsequent association of Psu, a double exponential function was used:  $A = A_0 + A_1(1 - \exp(-k_{app1} * t)) + A_2(1 - \exp(-k_{app2} * t))$ ;  $A$ , fluorescence at time  $t$ ,  $A_0$ , initial fluorescence signal;  $A_1$  and  $A_2$ , amplitudes of the signal change;  $k_{app1}$  and  $k_{app2}$ , characteristic time constants. Dependencies of the apparent rate constants on nucleotide concentration were fitted by a linear equation,  $k_{app} = k_1[*mant*-ATP $\gamma$ S] + k<sub>-1</sub>;  $k_1$ , nucleotide association rate constant (derived from the slope);  $k_{-1}$ , nucleotide dissociation rate constant (derived from the Y-axis intercept). For *mant*-ATP $\gamma$ S dissociation, data were fitted to a single exponential function:  $A = A_{min} + (A_0 - A_{min})\exp(-k_{app} * t)$ ;  $A$ , fluorescence at time  $t$ ,  $A_{min}$ , final signal;  $A_0$ , initial fluorescence signal;  $k_{app}$ , characteristic time constant.$

## Nucleic acid binding assays

Nucleic acid binding to  $\rho^{wt}$  or  $\rho^{P167L}$  PBS or SBS was tested *via* fluorescence depolarization-based assays.<sup>40</sup> For PBS binding, 5  $\mu$ M FAM-labeled dC<sub>15</sub> oligo (Eurofins) were mixed with increasing amounts of  $\rho^{wt}$  or  $\rho^{P167L}$  (0 to 8  $\mu$ M final hexamer concentration) in 20 mM HEPES pH 7.5, 150 mM KCl, 5 % (v/v) glycerol, 5 mM MgCl<sub>2</sub>, 0.5 mM TCEP. For SBS binding,  $\rho^{wt}$  or  $\rho^{P167L}$  PBSeS were first saturated with 10  $\mu$ M non-labeled dC<sub>15</sub>. Increasing amounts of PBS-blocked  $\rho^{wt}$  or  $\rho^{P167L}$  (0 to 16  $\mu$ M final hexamer concentration) were then mixed with 2 mM ADP-BeF<sub>3</sub> and 2  $\mu$ M FAM-labeled rU<sub>12</sub> oligo (Eurofins). For examining the effect of Psu on nucleic acid binding to the  $\rho^{wt}$  or  $\rho^{P167L}$  PBS or SBS, 1  $\mu$ M of  $\rho^{wt}$  or  $\rho^{P167L}$  hexamers were mixed with increasing amounts of Psu (0 to 30  $\mu$ M final dimer concentration). Subsequently, experiments were conducted as above. The fluorescence anisotropy was recorded in OptiPlate™ 384-well plates (PerkinElmer) using a Spark Multimode Microplate reader (Tecan; excitation wavelength, 485 nm; detected emission wavelength, 530 nm). Two technical replicates were averaged for each sample and the data were analyzed with Prism software (version 9.0.2; GraphPad). To quantify  $\rho^{wt}$  or  $\rho^{P167L}$  PBS

or SBS binding, data were fitted to a single exponential Hill function;  $A = A_{max}[\text{protein}]^h / (K_d^h + [\text{protein}]^h)$ ; in which  $A_{max}$  is the fitted maximum of nucleic acid bound;  $K_d$  is the dissociation constant; and  $h$  is the Hill coefficient. For quantification of the effects of Psu on  $\rho^{wt}$  or  $\rho^{P167L}$  PBS or SBS binding, data were fitted to an inhibitor vs. response function:  $A = A_{min} + (A_{max} - A_{min}) / (1 + ([\text{inhibitor}] / IC_{50}))$ ; in which  $A$  is the anisotropy signal at a given concentration of Psu (inhibitor); and  $A_{min}$  and  $A_{max}$  are the fitted minimum and maximum of nucleic acid bound.

### Generation of *rho* and *psu* mutants for *in vivo* assays

A low-copy-number pCL1920 plasmid guiding the production of  $\rho^{Y197A}$  was constructed by site-directed mutagenesis of a pCL1920 plasmid encoding  $\rho^{wt}$  (pRS649)<sup>93</sup>. pNL150 plasmids guiding the production of  $\text{Psu}^{R43A}$ ,  $\text{Psu}^{D49R}$  or  $\text{Psu}^{K180A}$  were constructed by site-directed mutagenesis of a pNL150 plasmid encoding  $\text{Psu}^{wt}$  (pRS1117)<sup>37</sup>. Mutations were confirmed by sequencing.

### Bacterial growth assays

To monitor possible growth defects of *rho* mutants in the absence and presence of Psu, *E. coli* RS1309 (MG1655 $\Delta\rho\Delta rac$ ) strain was transformed with pCL1920 plasmids expressing wt or mutant *rho* and were plated on LB in the absence of IPTG. After removal of the shelter plasmid expressing wt *rho*, the strains were transformed with empty pNL150 or with pNL150 expressing wt *psu*. The transformed colonies were streaked on LB agar plates supplemented with 0  $\mu\text{M}$  or 50  $\mu\text{M}$  IPTG and incubated at 37 °C. Serial dilutions of the overnight cultures of each transformed strain were spotted onto LB agar plates supplemented with 0, 50 or 100  $\mu\text{M}$  IPTG.

To monitor possible growth defects of the cells expressing *psu* mutants, *E. coli* MG1655 (RS1263) strain was transformed with pNL150 plasmids expressing either wt or mutant *psu*. The transformed colonies were streaked on LB agar plates supplemented with 0  $\mu\text{M}$  or 50  $\mu\text{M}$  IPTG.

To record growth curves, overnight cultures of the transformed strains were cultured in 96-well microtiter plates in the absence of IPTG or in the presence of 50  $\mu\text{M}$  IPTG. The growth was

monitored by measuring the absorbance at 660 nm using a Spectramax M5 plate reader. The curves were plotted using SigmaPlot software (version 13).

## **Pull-down assays**

The pET28b (*kan<sup>R</sup>*) plasmids guiding the production of Psu variants with an N-terminal His<sub>6</sub>-tag, and pET21b (*amp<sup>R</sup>*) plasmids guiding the production of p variants were co-transformed in *E. coli* BL21(DE3) cells. The transformed colonies were used to inoculate 5 ml of LB medium, and cultures were grown at 37 °C for ~3 h. The 5 ml cultures were then added to 100 ml LB medium and further incubated until the OD<sub>600</sub> reached ~0.3. The cultures were then induced with 0.1 mM IPTG for 3 h. Cells were then lysed in 100 mM NaH<sub>2</sub>PO<sub>4</sub>, pH 7.0, 100 mM NaCl, 10 mM imidazole, 1 mg/ml lysozyme, 10 mg/ml PMSF. The lysates were passed through Ni<sup>2+</sup>-NTA columns (Qiagen), washed with 100 mM NaH<sub>2</sub>PO<sub>4</sub>, pH 7.0, 100 mM NaCl, 20 mM imidazole, and the bound protein was eluted with 100 mM NaH<sub>2</sub>PO<sub>4</sub>, pH 7.0, 100 mM NaCl, 500 mM imidazole. Identical volumes of all buffers were used in each step. Identical volumes of the flow-through (FT), wash (W) and eluted fractions (E) were separated *via* SDS-PAGE. The amounts of p and Psu proteins in each of the FT, W and E fractions were quantified by ImageQuant (version 5.2; Cytiva) of the corresponding bands of the Coomassie-stained gels.

## ***In vivo* p-dependent transcription termination assays**

A *lacZ* reporter system was used for studying *in vivo* p-dependent transcription termination. *E. coli* RS2047 strain, containing the reporter cassette, *P<sub>lac</sub>-t<sub>R1</sub>-lacZ<sub>YA</sub>*, in the form of a λRS45 lysogen, was transformed with pCL1920 plasmids that guide the production of p variants. Subsequently, cells were transduced with *rho::kan<sup>R</sup>* lysate to remove the chromosomal copy of *rho*. Finally, the modified strains were transformed with an empty pNL150 vector or with pNL150 expressing *psu*. Transformed colonies were streaked on LB-X-gal agar plates supplemented with 0 or 5 μM IPTG, and were incubated at 37 °C for 14 to 16 h.

To study the effect of *psu* mutants on  $\rho$ -dependent transcription termination *in vivo*, the RS2047 strain was transformed with empty pNL150 or with pNL150 expressing wt or mutant *psu*. The transformed colonies were streaked on LB-X-gal agar plates supplemented with 0 or 15  $\mu$ M IPTG and were incubated at 37 °C for 12 to 14 hrs.

### **$\rho$ conservation analysis**

To analyze the conservation of  $\rho$ , a representative database was compiled (Supplementary Dataset 1A). The GTDB<sup>94</sup> bacterial taxonomy list (Released April 08, 2022) was downloaded. Genomes labeled as gtdb\_type\_species\_of\_genus and gtdb\_representative were kept. Then the genome assembly level was inspected. Only complete genomes were used to build the database. The new dataset covers 41 phyla including 649 genus. The protein sequences of all genomes were downloaded from NCBI. To identify  $\rho$ , Pfam model Rho\_RNA\_bind (PF07497) was searched against the representative database using hmmsearch (v. 3.3)<sup>95</sup> with a bit score of 27. The identified  $\rho$ -like proteins were further confirmed by searching  $\rho$  NCBI HMM model (NF006886.1) against them using hmmsearch with cutoff 315. Finally, 632  $\rho$  sequences are collected.

The length of  $\rho$  sequences ranges from 355 to 857 residues. Many of them have long insertions. To detect local homologies, multiple sequence alignment (MSA) was done by Dialign2<sup>96</sup> with the default settings. The MSA was indexed according to *E. coli* MG1655  $\rho$  (NP\_418230.1). The sequence conservation score was calculated using Protein Residue Conservation Prediction<sup>97</sup> with property entropy option. Sequence logo was generated using WebLogo (v. 3.7.8)<sup>98</sup>.

### **Compiling Psu sequences**

The Pfam model (PF07455) of Psu was searched against Reference Proteomes (v.2021\_04) using HmmerWeb (v. 2.41.2)<sup>99</sup> with bit scores 21.2 and E-values  $10^{-4}$ . However, only 28 sequences were found, one from Enterobacteria phage P4 and 27 from Enterobacterales. To



obtain a better picture of Psu diversity, we launched a 'greedy search': 1) searching Psu Pfam model against uniprotKB (v.2021\_04) using HmmerWeb<sup>99</sup> with bit score 21.2 and e-value  $10^{-4}$ ; 2) Jackhmmer searching with Enterobacteria phage P4 Psu reference sequence (NCBI ID: NP\_042044.1) as a query against uniprotKB (v.2021\_04) using HmmerWeb<sup>99</sup>. A total of seven iterations were done; 3) blastp searching with Psu reference sequence as a query against uniprotKB (v.2022\_04) using UniProt website<sup>100</sup> with e-value  $10^{-4}$ ; 4) blastp searching Psu reference sequence against Reference proteins using NCBI website<sup>101</sup> with e-value  $10^{-4}$ . Max target sequences option was set to 5000; 5) blastp searching Psu reference sequence against Non-redundant protein sequences using NCBI website<sup>101</sup> with e-value  $10^{-4}$ . The search was limited to the Virus superkingdom (Taxonomy ID: 10239). Max target sequences option was set to 5000.

All sequences from the 'greedy search' were downloaded from NCBI and pooled together. An all-vs-all blastp was performed using BLAST+ (v. 2.9.0)<sup>102</sup> with e-value  $10^{-9}$ . A Psu similarity network was built based on e-value and visualized in Cytoscape (v. 3.9.1)<sup>103</sup>. Sequences from two Psu-like groups in the network (Extended Data Fig. 8) were fed into Batch CD-Search<sup>104</sup> to confirm the presence of the Psu superfamily (NCBI ID: cl06476). Sequences with > 80% coverage of the Psu superfamily and sequences with coverage < 80% but having no 'incomplete' flags were selected. Finally, 2625 Psu sequences were kept (Supplementary Dataset 1B). NCBI taxonomy ID was first assigned to the Psu sequences using accession2taxid file (<https://ftp.ncbi.nlm.nih.gov/pub/taxonomy/accession2taxid/>). Then NCBI taxonomy was replaced with GTDB taxonomy<sup>94</sup>. As the reference sequences are not organism-specific, the taxonomy was assigned to the genus level.

## **Psu conservation**

The compiled Psu sequences from Pseudomonadota were clustered at 90 % (-c 0.9 option) using cd-hit (v. 4.8.1)<sup>105</sup>. Then two sequences from Firmicutes and three sequences from the virus



were added into the pool. A total of 124 sequences were aligned using MUSCLE (v. 5.1) with default settings (<https://doi.org/10.1101/2021.06.20.449169>). The multiple sequence alignment of *Psu* was indexed according to *Psu* reference sequence before generating sequence logo using WebLogo (v. 3.7.8)<sup>98</sup>.

## Identification of anti-phage defense systems

The accession numbers of ten neighboring genes on both sides of *psu* were fetched by TREND<sup>106</sup> and the sequences were downloaded from NCBI. A total of 1,163 *psu*-containing genomes were investigated. The presence of Enterobacteria phage P4-like genes among the neighboring genes was determined using blastp. The proteins from phage P4 reference genome (NCBI ID: GCF\_000846325.1) were used as queries, and e-value threshold was set to  $10^{-5}$ . To find known anti-phage defense system, all neighboring proteins were submitted to DefenseFinder<sup>46</sup>. The results were checked manually. Gene clusters located between *psu* and *int* genes that were not identified by DefenseFinder were considered putative defense systems (Supplementary Dataset 1E) and genes from putative defense systems were annotated (Supplementary Dataset 1G,H) by searching against Pfam-A database (v. 35.0) using hmmscan (v. 3.2.1)<sup>95</sup>.

## Data availability

CryoEM reconstructions have been deposited in the Electron Microscopy Data Bank (<https://www.ebi.ac.uk/pdbe/emdb>) under accession codes EMD-17637 (<https://www.ebi.ac.uk/pdbe/entry/emdb/EMD-17637>;  $\rho^{\text{wt}}$ -ATP $\gamma$ S-*Psu* complex I), EMD-17639 (<https://www.ebi.ac.uk/pdbe/entry/emdb/EMD-17639>;  $\rho^{\text{wt}}$ -ATP $\gamma$ S-*Psu* complex II), EMD-17640 (<https://www.ebi.ac.uk/pdbe/entry/emdb/EMD-17640>;  $\rho^{\text{P167L}}$ -ATP $\gamma$ S-*Psu* complex I) and EMD-17641 (<https://www.ebi.ac.uk/pdbe/entry/emdb/EMD-17641>;  $\rho^{\text{P167L}}$ -ATP $\gamma$ S-*Psu* complex II).

815 Structure coordinates have been deposited in the RCSB Protein Data Bank (<https://www.rcsb.org>)  
816 with accession codes 8PEU (<https://www.rcsb.org/structure/8PEU>;  $\rho^{\text{wt}}$ -ATPyS-Psu complex I),  
817 8PEW (<https://www.rcsb.org/structure/8PEW>;  $\rho^{\text{wt}}$ -ATPyS-Psu complex II), 8PEX  
818 (<https://www.rcsb.org/structure/8PEX>;  $\rho^{\text{P167L}}$ -ATPyS-Psu complex I) and 8PEY  
819 (<https://www.rcsb.org/structure/8PEY>;  $\rho^{\text{P167L}}$ -ATPyS-Psu complex II). All other data are contained  
820 in the manuscript or the Supplementary Information. Source data are provided with this paper.  
821 Structure coordinates used in this study are available from the RCSB Protein Data Bank  
822 (<https://www.rcsb.org>) under accession codes 1PV4 (<https://www.rcsb.org/structure/1PV4>) and  
823 3RX6 (<https://www.rcsb.org/structure/3RX6>).

824

# References

1. D'Heygere, F., Rabhi, M. & Boudvillain, M. Phyletic distribution and conservation of the bacterial transcription termination factor Rho. *Microbiology (Reading)* **159**, 1423-1436 (2013).
2. Roberts, J.W. Termination factor for RNA synthesis. *Nature* **224**, 1168-1174 (1969).
3. Peters, J.M. et al. Rho directs widespread termination of intragenic and stable RNA transcription. *Proc Natl Acad Sci U S A* **106**, 15406-15411 (2009).
4. Sedlyarova, N. et al. sRNA-Mediated Control of Transcription Termination in E. coli. *Cell* **167**, 111-121 e113 (2016).
5. Sedlyarova, N. et al. Natural RNA Polymerase Aptamers Regulate Transcription in E. coli. *Molecular Cell* **67**, 30-43 (2017).
6. Peters, J.M. et al. Rho and NusG suppress pervasive antisense transcription in Escherichia coli. *Genes Dev* **26**, 2621-2633 (2012).
7. Leela, J.K., Syeda, A.H., Anupama, K. & Gowrishankar, J. Rho-dependent transcription termination is essential to prevent excessive genome-wide R-loops in Escherichia coli. *Proc Natl Acad Sci U S A* **110**, 258-263 (2013).
8. Cardinale, C.J. et al. Termination factor Rho and its cofactors NusA and NusG silence foreign DNA in E. coli. *Science* **320**, 935-938 (2008).
9. Krupp, F. et al. Structural Basis for the Action of an All-Purpose Transcription Anti-Termination Factor. *Mol Cell* **74**, 143-157 (2019).
10. Said, N. et al. Structural basis for lambdaN-dependent processive transcription antitermination. *Nat Microbiol* **2**, 17062 (2017).
11. Yin, Z., Bird, J.G., Kaelber, J.T., Nickels, B.E. & Ebright, R.H. In transcription antitermination by Qlambda, NusA induces refolding of Qlambda to form a nozzle that extends the RNA polymerase RNA-exit channel. *Proc Natl Acad Sci U S A* **119**, e2205278119 (2022).
12. Yin, Z., Kaelber, J.T. & Ebright, R.H. Structural basis of Q-dependent antitermination. *Proc Natl Acad Sci U S A* **116**, 18384-18390 (2019).
13. Shi, J. et al. Structural basis of Q-dependent transcription antitermination. *Nat Commun* **10**, 2925 (2019).
14. Hwang, S. et al. Structural basis of transcriptional regulation by a nascent RNA element, HK022 putRNA. *Nat Commun* **13**, 4668 (2022).
15. Dey, S. et al. Structural insights into RNA-mediated transcription regulation in bacteria. *Mol Cell* **82**, 3885-3900 e3810 (2022).
16. Isaksen, M.L., Rishovd, S.T., Calendar, R. & Lindqvist, B.H. The polarity suppression factor of bacteriophage P4 is also a decoration protein of the P4 capsid. *Virology* **188**, 831-839 (1992).
17. Linderth, N.A. & Calendar, R.L. The Psi protein of bacteriophage P4 is an antitermination factor for rho-dependent transcription termination. *J Bacteriol* **173**, 6722-6731 (1991).
18. Lagos, R., Jiang, R.Z., Kim, S. & Goldstein, R. Rho-dependent transcription termination of a bacterial operon is antagonized by an extrachromosomal gene product. *Proc Natl Acad Sci U S A* **83**, 9561-9565 (1986).

- 866 19. Sauer, B., Ow, D., Ling, L. & Calendar, R. Mutants of satellite bacteriophage P4 that are  
867 defective in the suppression of transcriptional polarity. *J Mol Biol* **145**, 29-46 (1981).
- 868 20. Mitra, P., Ghosh, G., Hafeezunnisa, M. & Sen, R. Rho Protein: Roles and Mechanisms.  
869 *Annu Rev Microbiol* **71**, 687-709 (2017).
- 870 21. Ray-Soni, A., Bellecourt, M.J. & Landick, R. Mechanisms of Bacterial Transcription  
871 Termination: All Good Things Must End. *Annu Rev Biochem* **85**, 319-347 (2016).
- 872 22. Thomsen, N.D., Lawson, M.R., Witkowsky, L.B., Qu, S. & Berger, J.M. Molecular  
873 mechanisms of substrate-controlled ring dynamics and substepping in a nucleic acid-  
874 dependent hexameric motor. *Proc Natl Acad Sci U S A* **113**, E7691-E7700 (2016).
- 875 23. Richardson, J.P. Rho-dependent termination and ATPases in transcript termination.  
876 *Biochim Biophys Acta* **1577**, 251-260 (2002).
- 877 24. Park, J.S. & Roberts, J.W. Role of DNA bubble rewinding in enzymatic transcription  
878 termination. *Proc Natl Acad Sci U S A* **103**, 4870-4875 (2006).
- 879 25. Molodtsov, V., Wang, C., Firlar, E., Kaelber, J.T. & Ebright, R.H. Structural basis of Rho-  
880 dependent transcription termination. *Nature* **614**, 367-374 (2023).
- 881 26. Epshtein, V., Dutta, D., Wade, J. & Nudler, E. An allosteric mechanism of Rho-dependent  
882 transcription termination. *Nature* **463**, 245-249 (2010).
- 883 27. Said, N. et al. Steps toward translocation-independent RNA polymerase inactivation by  
884 terminator ATPase rho. *Science* **371**(2021).
- 885 28. Hao, Z. et al. Pre-termination Transcription Complex: Structure and Function. *Mol Cell* **81**,  
886 281-292 e288 (2021).
- 887 29. Mooney, R.A. et al. Regulator Trafficking on Bacterial Transcription Units In Vivo. *Molecular*  
888 *Cell* **33**, 97-108 (2009).
- 889 30. Song, E. et al. Rho-dependent transcription termination proceeds via three routes. *Nat*  
890 *Commun* **13**, 1663 (2022).
- 891 31. Song, E. et al. Transcriptional pause extension benefits the stand-by rather than catch-up  
892 Rho-dependent termination. *Nucleic Acids Res* (2023).
- 893 32. Skordalakes, E. & Berger, J.M. Structure of the Rho transcription terminator: mechanism of  
894 mRNA recognition and helicase loading. *Cell* **114**, 135-146 (2003).
- 895 33. Ranjan, A., Sharma, S., Banerjee, R., Sen, U. & Sen, R. Structural and mechanistic basis  
896 of anti-termination of Rho-dependent transcription termination by bacteriophage P4 capsid  
897 protein Psu. *Nucleic Acids Res* **41**, 6839-6856 (2013).
- 898 34. Ranjan, A., Banerjee, R., Pani, B., Sen, U. & Sen, R. The moonlighting function of  
899 bacteriophage P4 capsid protein, Psu, as a transcription antiterminator. *Bacteriophage* **3**,  
900 e25657 (2013).
- 901 35. Ding, F., Prutzman, K.C., Campbell, S.L. & Dokholyan, N.V. Topological determinants of  
902 protein domain swapping. *Structure* **14**, 5-14 (2006).
- 903 36. Fan, C. & Rees, D.C. Crystal structure of the Escherichia coli transcription termination factor  
904 Rho. *Acta Crystallogr F Struct Biol Commun* **76**, 398-405 (2020).
- 905 37. Linderoth, N.A., Tang, G. & Calendar, R. In vivo and in vitro evidence for an anti-Rho activity  
906 induced by the phage P4 polarity suppressor protein Psu. *Virology* **227**, 131-141 (1997).

- 907 38. Pani, B. et al. Mechanism of inhibition of Rho-dependent transcription termination by  
908 bacteriophage P4 protein Psi. *J Biol Chem* **281**, 26491-26500 (2006).
- 909 39. Yengo, C.M., Chrin, L.R., Rovner, A.S. & Berger, C.L. Tryptophan 512 is sensitive to  
910 conformational changes in the rigid relay loop of smooth muscle myosin during the  
911 MgATPase cycle. *J Biol Chem* **275**, 25481-25487 (2000).
- 912 40. Lawson, M.R. & Berger, J.M. A Fluorescent Assay to Monitor Ligand-Dependent Closure of  
913 the Hexameric Rho Helicase Ring. *Methods Mol Biol* **2209**, 133-142 (2021).
- 914 41. Richardson, J.P. Activation of rho protein ATPase requires simultaneous interaction at two  
915 kinds of nucleic acid-binding sites. *J Biol Chem* **257**, 5760-5766 (1982).
- 916 42. Pani, B., Ranjan, A. & Sen, R. Interaction surface of bacteriophage P4 protein Psi required  
917 for complex formation with the transcription terminator Rho. *J Mol Biol* **389**, 647-660 (2009).
- 918 43. Moura de Sousa, J.A. & Rocha, E.P.C. To catch a hijacker: abundance, evolution and  
919 genetic diversity of P4-like bacteriophage satellites. *Philos Trans R Soc Lond B Biol Sci*  
920 **377**, 20200475 (2022).
- 921 44. Rousset, F. et al. Phages and their satellites encode hotspots of antiviral systems. *Cell Host*  
922 *Microbe* **30**, 740-753 e745 (2022).
- 923 45. Vassallo, C.N., Doering, C.R., Littlehale, M.L., Teodoro, G.I.C. & Laub, M.T. A functional  
924 selection reveals previously undetected anti-phage defence systems in the E. coli  
925 pangenome. *Nat Microbiol* **7**, 1568-1579 (2022).
- 926 46. Tesson, F. et al. Systematic and quantitative view of the antiviral arsenal of prokaryotes.  
927 *Nat Commun* **13**, 2561 (2022).
- 928 47. Doron, S. et al. Systematic discovery of antiphage defense systems in the microbial  
929 pangenome. *Science* **359**(2018).
- 930 48. Gao, L. et al. Diverse enzymatic activities mediate antiviral immunity in prokaryotes. *Science*  
931 **369**, 1077-1084 (2020).
- 932 49. Millman, A. et al. An expanded arsenal of immune systems that protect bacteria from  
933 phages. *Cell Host Microbe* **30**, 1556-1569 e1555 (2022).
- 934 50. Garb, J. et al. Multiple phage resistance systems inhibit infection via SIR2-dependent  
935 NAD(+) depletion. *Nat Microbiol* **7**, 1849-1856 (2022).
- 936 51. Zaremba, M. et al. Short prokaryotic Argonautes provide defence against incoming mobile  
937 genetic elements through NAD(+) depletion. *Nat Microbiol* **7**, 1857-1869 (2022).
- 938 52. Ofir, G. et al. Antiviral activity of bacterial TIR domains via immune signalling molecules.  
939 *Nature* **600**, 116-120 (2021).
- 940 53. Christie, G.E. & Dokland, T. Pirates of the Caudovirales. *Virology* **434**, 210-221 (2012).
- 941 54. Kizziah, J.L., Rodenburg, C.M. & Dokland, T. Structure of the Capsid Size-Determining  
942 Scaffold of "Satellite" Bacteriophage P4. *Viruses* **12**(2020).
- 943 55. Dokland, T., Isaksen, M.L., Fuller, S.D. & Lindqvist, B.H. Capsid localization of the  
944 bacteriophage P4 Psi protein. *Virology* **194**, 682-687 (1993).
- 945 56. Lawson, M.R. et al. Mechanism for the Regulated Control of Bacterial Transcription  
946 Termination by a Universal Adaptor Protein. *Mol Cell* **71**, 911-922 e914 (2018).
- 947 57. Lau, L.F. & Roberts, J.W. Rho-dependent transcription termination at lambda R1 requires  
948 upstream sequences. *J Biol Chem* **260**, 574-584 (1985).

- 949 58. Qayyum, M.Z., Dey, D. & Sen, R. Transcription Elongation Factor NusA Is a General  
950 Antagonist of Rho-dependent Termination in Escherichia coli. *J Biol Chem* **291**, 8090-8108  
951 (2016).
- 952 59. Wang, C. et al. Structural basis of transcription-translation coupling. *Science* **369**, 1359-  
953 1365 (2020).
- 954 60. Webster, M.W. et al. Structural basis of transcription-translation coupling and collision in  
955 bacteria. *Science* **369**, 1355-1359 (2020).
- 956 61. O'Reilly, F.J. et al. In-cell architecture of an actively transcribing-translating expressome.  
957 *Science* **369**, 554-557 (2020).
- 958 62. Rabhi, M. et al. The Sm-like RNA chaperone Hfq mediates transcription antitermination at  
959 Rho-dependent terminators. *EMBO J* **30**, 2805-2816 (2011).
- 960 63. Gogol, E.P., Seifried, S.E. & von Hippel, P.H. Structure and assembly of the Escherichia  
961 coli transcription termination factor rho and its interaction with RNA. I. Cryoelectron  
962 microscopic studies. *J Mol Biol* **221**, 1127-1138 (1991).
- 963 64. Seifried, S.E., Bjornson, K.P. & von Hippel, P.H. Structure and assembly of the Escherichia  
964 coli transcription termination factor rho and its interactions with RNA. II. Physical chemical  
965 studies. *J Mol Biol* **221**, 1139-1151 (1991).
- 966 65. Geiselmann, J., Seifried, S.E., Yager, T.D., Liang, C. & von Hippel, P.H. Physical properties  
967 of the Escherichia coli transcription termination factor rho. 2. Quaternary structure of the rho  
968 hexamer. *Biochemistry* **31**, 121-132 (1992).
- 969 66. Pal, K. et al. Vibrio cholerae YaeO is a Structural Homologue of RNA Chaperone Hfq that  
970 Inhibits Rho-dependent Transcription Termination by Dissociating its Hexameric State. *J*  
971 *Mol Biol* **431**, 4749-4766 (2019).
- 972 67. Yuan, A.H. & Hochschild, A. A bacterial global regulator forms a prion. *Science* **355**, 198-  
973 201 (2017).
- 974 68. Kryptou, E. et al. Bacteria require phase separation for fitness in the mammalian gut.  
975 *Science* **379**, 1149-1156 (2023).
- 976 69. Gabizon, R. & Friedler, A. Allosteric modulation of protein oligomerization: an emerging  
977 approach to drug design. *Front Chem* **2**, 9 (2014).
- 978 70. Banchenko, S. et al. Common Mode of Remodeling AAA ATPases p97/CDC48 by Their  
979 Disassembling Cofactors ASPL/PUX1. *Structure* **27**, 1830-1841 e1833 (2019).
- 980 71. Arumughan, A. et al. Quantitative interaction mapping reveals an extended UBX domain in  
981 ASPL that disrupts functional p97 hexamers. *Nat Commun* **7**, 13047 (2016).
- 982 72. Rancour, D.M., Park, S., Knight, S.D. & Bednarek, S.Y. Plant UBX domain-containing  
983 protein 1, PUX1, regulates the oligomeric structure and activity of arabidopsis CDC48. *J*  
984 *Biol Chem* **279**, 54264-54274 (2004).
- 985 73. Stuber, J.C. et al. Apoptosis-inducing anti-HER2 agents operate through oligomerization-  
986 induced receptor immobilization. *Commun Biol* **4**, 762 (2021).
- 987 74. Hayouka, Z. et al. Inhibiting HIV-1 integrase by shifting its oligomerization equilibrium. *Proc*  
988 *Natl Acad Sci U S A* **104**, 8316-8321 (2007).
- 989 75. Ghosh, G., Sharma, P.V., Kumar, A., Jain, S. & Sen, R. Design of novel peptide inhibitors  
990 against the conserved bacterial transcription terminator, Rho. *J Biol Chem* **296**, 100653  
991 (2021).



992 76. Koonin, E.V., Makarova, K.S., Wolf, Y.I. & Krupovic, M. Evolutionary entanglement of mobile  
993 genetic elements and host defence systems: guns for hire. *Nat Rev Genet* **21**, 119-131  
994 (2020).

995 77. Bernheim, A. & Sorek, R. The pan-immune system of bacteria: antiviral defence as a  
996 community resource. *Nat Rev Microbiol* **18**, 113-119 (2020).

997 78. Cheng, R. et al. A nucleotide-sensing endonuclease from the Gabija bacterial defense  
998 system. *Nucleic Acids Res* **49**, 5216-5229 (2021).

999 79. Bobonis, J. et al. Bacterial retrons encode phage-defending tripartite toxin-antitoxin  
1000 systems. *Nature* **609**, 144-150 (2022).

1001 80. Ghosh, G., Reddy, J., Sambhare, S. & Sen, R. A Bacteriophage Capsid Protein Is an  
1002 Inhibitor of a Conserved Transcription Terminator of Various Bacterial Pathogens. *J*  
1003 *Bacteriol* **200**(2018).

1004 81. Rousset, F. & Sorek, R. The evolutionary success of regulated cell death in bacterial  
1005 immunity. *Curr Opin Microbiol* **74**, 102312 (2023).

1006 82. Edelheit, O., Hanukoglu, A. & Hanukoglu, I. Simple and efficient site-directed mutagenesis  
1007 using two single-primer reactions in parallel to generate mutants for protein structure-  
1008 function studies. *BMC Biotechnol* **9**, 61 (2009).

1009 83. Banerjee, R. et al. The first structure of polarity suppression protein, Psu from  
1010 enterobacteria phage P4, reveals a novel fold and a knotted dimer. *J Biol Chem* **287**, 44667-  
1011 44675 (2012).

1012 84. Sherlin, L.D. et al. Chemical and enzymatic synthesis of tRNAs for high-throughput  
1013 crystallization. *Rna* **7**, 1671-1678 (2001).

1014 85. Kao, C., Zheng, M. & Rüdiger, S. A simple and efficient method to reduce nontemplated  
1015 nucleotide addition at the 3' terminus of RNAs transcribed by T7 RNA polymerase. *RNA* **5**,  
1016 1268-1272 (1999).

1017 86. Punjani, A., Rubinstein, J.L., Fleet, D.J. & Brubaker, M.A. cryoSPARC: algorithms for rapid  
1018 unsupervised cryo-EM structure determination. *Nat Methods* **14**, 290-296 (2017).

1019 87. Casan, A., Lohkamp, B. & Emsley, P. Current developments in Coot for macromolecular  
1020 model building of Electron Cryo-microscopy and Crystallographic Data. *Protein Sci* **29**,  
1021 1069-1078 (2020).

1022 88. Emsley, P. & Cowtan, K. Coot: model-building tools for molecular graphics. *Acta Crystallogr*  
1023 *D Biol Crystallogr* **60**, 2126-2132 (2004).

1024 89. Adams, P.D. et al. The Phenix software for automated determination of macromolecular  
1025 structures. *Methods* **55**, 94-106 (2011).

1026 90. Afonine, P.V. et al. Real-space refinement in PHENIX for cryo-EM and crystallography. *Acta*  
1027 *Crystallogr D Struct Biol* **74**, 531-544 (2018).

1028 91. Williams, C.J. et al. MolProbity: More and better reference data for improved all-atom  
1029 structure validation. *Protein Sci* **27**, 293-315 (2018).

1030 92. Pettersen, E.F. et al. UCSF ChimeraX: Structure visualization for researchers, educators,  
1031 and developers. *Protein Sci* **30**, 70-82 (2021).

1032 93. Muteeb, G., Dey, D., Mishra, S. & Sen, R. A multipronged strategy of an anti-terminator  
1033 protein to overcome Rho-dependent transcription termination. *Nucleic Acids Res* **40**, 11213-  
1034 11228 (2012).

94. Parks, D.H. et al. GTDB: an ongoing census of bacterial and archaeal diversity through a phylogenetically consistent, rank normalized and complete genome-based taxonomy. *Nucleic Acids Res* **50**, D785-D794 (2022).
95. Mistry, J., Finn, R.D., Eddy, S.R., Bateman, A. & Punta, M. Challenges in homology search: HMMER3 and convergent evolution of coiled-coil regions. *Nucleic Acids Res* **41**, e121 (2013).
96. Al Ait, L., Yamak, Z. & Morgenstern, B. DIALIGN at GOBICS--multiple sequence alignment using various sources of external information. *Nucleic Acids Res* **41**, W3-7 (2013).
97. Capra, J.A. & Singh, M. Predicting functionally important residues from sequence conservation. *Bioinformatics* **23**, 1875-1882 (2007).
98. Crooks, G.E., Hon, G., Chandonia, J.M. & Brenner, S.E. WebLogo: a sequence logo generator. *Genome Res* **14**, 1188-1190 (2004).
99. Potter, S.C. et al. HMMER web server: 2018 update. *Nucleic Acids Res* **46**, W200-W204 (2018).
100. UniProt, C. UniProt: the universal protein knowledgebase in 2021. *Nucleic Acids Res* **49**, D480-D489 (2021).
101. Johnson, M. et al. NCBI BLAST: a better web interface. *Nucleic Acids Res* **36**, W5-9 (2008).
102. Camacho, C. et al. BLAST+: architecture and applications. *BMC Bioinformatics* **10**, 421 (2009).
103. Shannon, P. et al. Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res* **13**, 2498-2504 (2003).
104. Lu, S. et al. CDD/SPARCLE: the conserved domain database in 2020. *Nucleic Acids Res* **48**, D265-D268 (2020).
105. Li, W. & Godzik, A. Cd-hit: a fast program for clustering and comparing large sets of protein or nucleotide sequences. *Bioinformatics* **22**, 1658-1659 (2006).
106. Gumerov, V.M. & Zhulin, I.B. TREND: a platform for exploring protein function in prokaryotes based on phylogenetic, domain architecture and gene neighborhood analyses. *Nucleic Acids Res* **48**, W72-W76 (2020).
107. Valabhoju, V., Agrawal, S. & Sen, R. Molecular Basis of NusG-mediated Regulation of Rho-dependent Transcription Termination in Bacteria. *J Biol Chem* **291**, 22386-22403 (2016).



# Acknowledgements

We acknowledge the assistance of the core facility BioSupraMol supported by the Deutsche Forschungsgemeinschaft in electron microscopic analyses. This work was supported by grants from the Deutsche Forschungsgemeinschaft (INST 130/1064-1 FUGG to Freie Universität Berlin; GRK 2473 "Bioactive Peptides", project number 392923329, to M.C.W.; WA 1126/11-1, project number 433623608, to M.C.W.), the Bundesministerium für Bildung und Forschung (ICMR2019-016 to M.C.W.), the Berlin University Alliance (501\_BIS-CryoFac to M.C.W.), the Indian Council of Medical Research (AMR/INDO/GER/219/2019-ECD-IIG to R.S.), and the National Institutes of Health (GM067153 to I.A.).

# Author contributions

D.G. produced recombinant proteins and assembled complexes with help by N.S.. T.H. acquired, processed and refined cryoEM data. D.G. built atomic models with the help of M.C.W. and B.L. and refined models with the help of B.L.. D.G. conducted *in vitro* biochemical and biophysical experiments. N.K. cloned constructs for *in vivo* studies and conducted the *in vivo* pull-down, bacterial growth and termination assays. B.W. performed bioinformatic analyses. All authors contributed to the analysis of the data and the interpretation of the results. D.G. wrote the original draft with contributions from I.A. and R.S.. M.C.W. revised the manuscript with contributions from D.G., I.A. and R.S.. I.A., R.S. and M.C.W. supervised work in their respective groups and coordinated the collaboration.

# Competing interests

The authors declare no competing interests.



**Fig. 1: Structures of  $\rho^{\text{wt}}$ -ATPyS-Psu complexes.**

**a**, CryoEM reconstruction of  $\rho^{\text{wt}}$ -ATPyS-Psu complex I. Six Psu dimers (violet and purple) bridge two open  $\rho^{\text{wt}}$  hexamers,  $\rho^{\text{wt}}$ (A) and  $\rho^{\text{wt}}$ (B). Individual subunits of  $\rho^{\text{wt}}$ (A) and  $\rho^{\text{wt}}$ (B) are colored in increasingly darker shades of gray from the center to the peripheries of the complex.

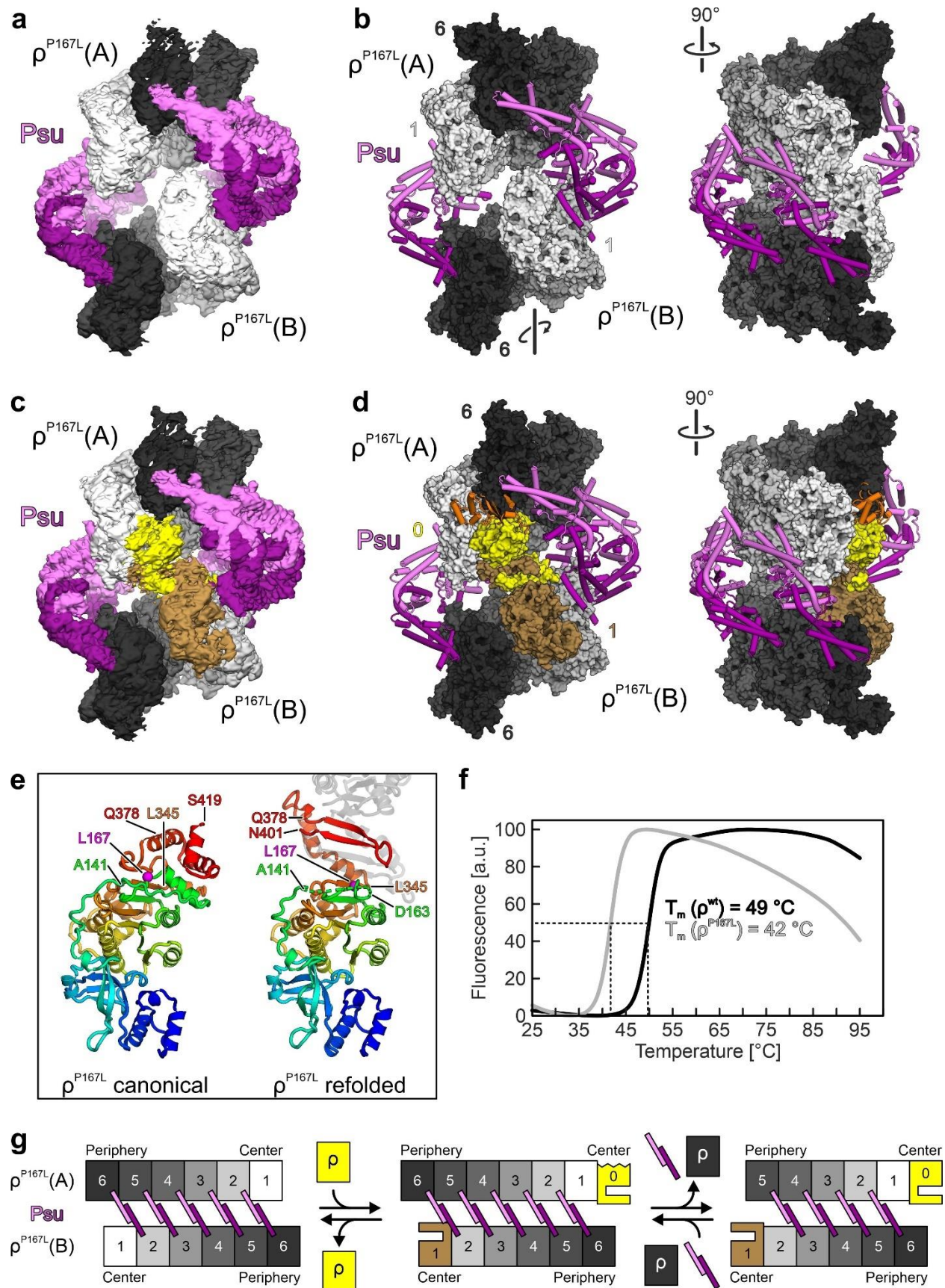
**b**, Orthogonal views of  $\rho^{\text{wt}}$ -ATPyS-Psu complex I. Psu dimers, cartoon representation;  $\rho^{\text{wt}}$ , surface representation. Subunit coloring as in (a).  $\rho^{\text{wt}}$  subunits are designated by increasing Arabic numerals from the centers to the peripheries of the complex.

**c**, CryoEM reconstruction of  $\rho^{\text{wt}}$ -ATPyS-Psu complex II. Eight Psu dimers (violet and purple) bridge two open  $\rho^{\text{wt}}$  nonamers,  $\rho^{\text{wt}}$ (A) and  $\rho^{\text{wt}}$ (B). Portions of  $\rho^{\text{wt}}$ (A) and  $\rho^{\text{wt}}$ (B) equivalent to  $\rho^{\text{wt}}$ -ATPyS-Psu complex I are shown in the same colors.  $\rho^{\text{wt}}$ -ATPyS-Psu complex II can be envisaged to emerge from complex I by the additional of  $\rho^{\text{wt}}$  subunits and Psu dimers. Additional  $\rho^{\text{wt}}$  subunits added at the center of complex I, yellow; additional  $\rho^{\text{wt}}$  subunits added at the peripheries of complex I, light and dark green.

**d**, Orthogonal views of  $\rho^{\text{wt}}$ -ATPyS-Psu complex II. Psu dimers, cartoon representation;  $\rho^{\text{wt}}$ , surface representation. Subunit coloring as in (c). The portion of complex II that is equivalent to complex I is shown in semi-transparent cartoon/surface representation, additional Psu dimers and  $\rho^{\text{wt}}$  subunits in complex II are shown in solid cartoon/surface representation.

**e**, Scheme illustrating  $\rho^{\text{wt}}$ -Psu interaction patterns in  $\rho^{\text{wt}}$ -ATPyS-Psu complexes I and II. A dynamic ensemble of  $\rho^{\text{wt}}$ -ATPyS-Psu complexes is supported by additional complexes observed in the cryoEM analysis (Extended Data Fig. 3).





1116

**Fig. 2: Structures of  $\rho^{P167L}$ -ATPyS-Psu complexes.**

**a**, CryoEM reconstruction of  $\rho^{P167L}$ -ATPyS-Psu complex I. Five Psu dimers (violet and purple) bridge two open  $\rho^{P167L}$  hexamers,  $\rho^{P167L}(A)$  and  $\rho^{P167L}(B)$ . Subunit coloring as for  $\rho^{wt}$ -ATPyS-Psu complex I (Fig. 1a). Rotation symbol, different rotational arrangement of  $\rho^{P167L}(B)$  relative to  $\rho^{P167L}(A)$ , compared to  $\rho^{wt}$ -ATPyS-Psu complex I.

**b**, Orthogonal views of  $\rho^{P167L}$ -ATPyS-Psu complex I. Psu dimers, cartoon representation;  $\rho^{P167L}$ , surface representation. Subunit coloring as in (a).  $\rho^{P167L}$  subunits are designated by increasing Arabic numerals from the centers to the peripheries of the complex.

**c**, CryoEM reconstruction of  $\rho^{P167L}$ -ATPyS-Psu complex II. Compared to  $\rho^{P167L}$ -ATPyS-Psu complex I, an additional  $\rho^{P167L}$  subunit (yellow) joined the complex, expanding  $\rho^{P167L}(A)$  into a heptamer. The additional  $\rho^{P167L}$  subunit is refolded and engages in a domain-swapped interaction with  $\rho^{P167L}$  subunit 1 of the  $\rho^{P167L}(B)$  hexamer (light brown).

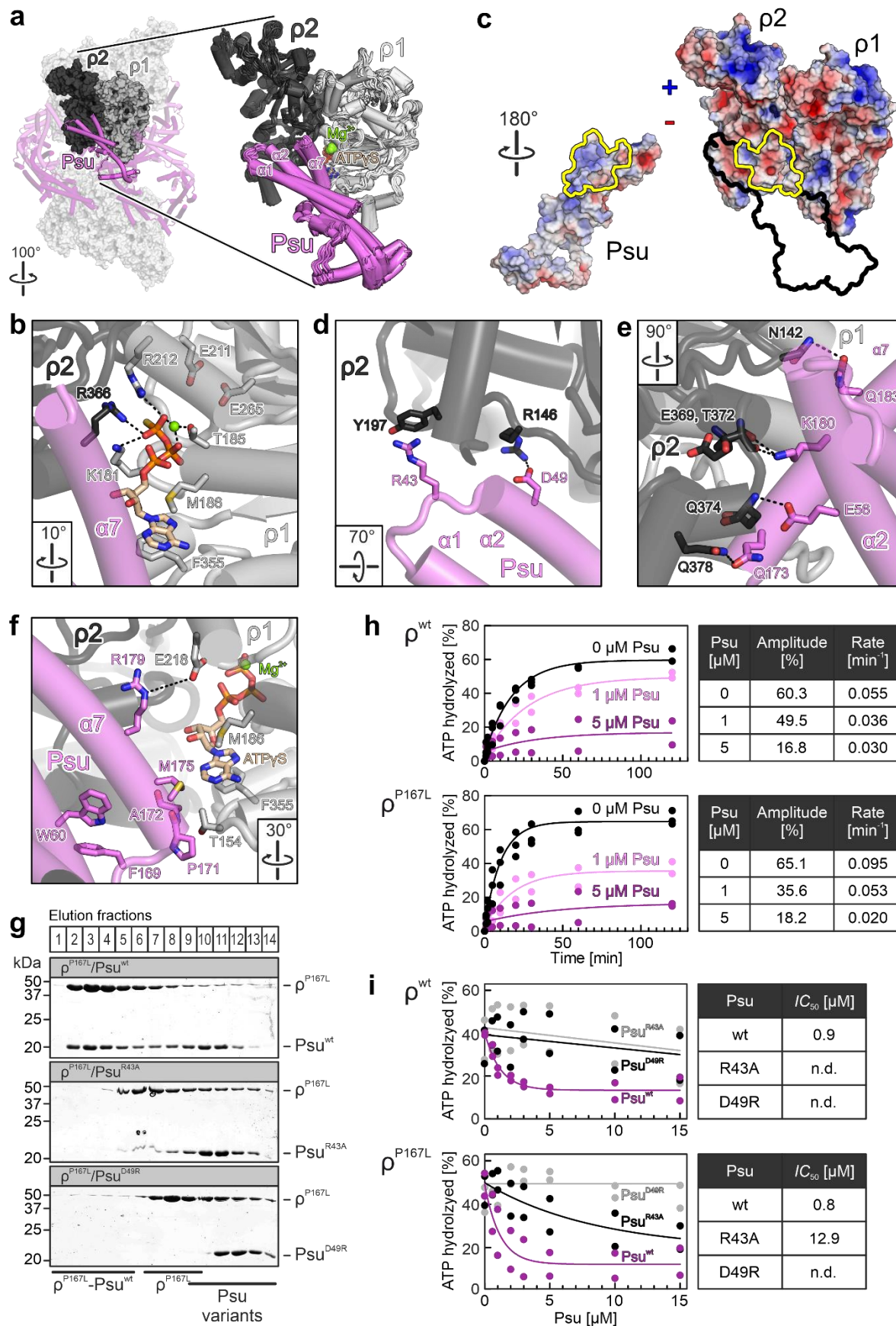
**d**, Orthogonal views of  $\rho^{P167L}$ -ATPyS-Psu complex II. Psu dimers, cartoon representation;  $\rho^{P167L}$ , surface representation. Subunit coloring as in (c). An N-terminal domain modeled in canonical conformation onto the additional  $\rho^{P167L}$  subunit (orange cartoon) would clash with the peripheral subunit of  $\rho^{P167L}(A)$ .

**e**, Side-by-side comparison of a  $\rho^{P167L}$  subunit in canonical conformation (left) and in the refolded conformation (right). The models are colored blue to red from N- to C-terminus. For the refolded model, a second refolded  $\rho^{P167L}$  subunit, interacting *via* domain-swap in  $\rho^{P167L}$ -ATPyS-Psu complex II, is shown as a gray, semi-transparent cartoon. C $\alpha$  atom of L167, magenta sphere. Positions of landmark residues used to describe the rearrangement are indicated.

**f**, Differential scanning fluorimetry monitoring fold stabilities of  $\rho^{wt}$  (black curve) and  $\rho^{P167L}$  (gray curve).  $T_m$  values were determined as the maxima of the first derivatives of the melting curves. Experiments were repeated independently at least three times with similar results.

1142 **g**, Scheme illustrating  $\rho^{P167L}$ -Psu interaction patterns in  $\rho^{P167L}$ -ATPyS-Psu complexes I (left) and  
 1143 II (center), and how an additional complex observed in the cryoEM analysis (right; Extended Data  
 1144 Fig. 6) could emerge from complex II.

1145





**Fig. 3: p-Psu interaction details.**

**a**, Left,  $\rho^{\text{wt}}$ -ATPyS-Psu complex I with a building block comprising two neighboring  $\rho^{\text{wt}}$  subunits (p1 and p2) and an interacting Psu molecule highlighted in a solid representation. Right, overlay of all corresponding building blocks of the  $\rho^{\text{wt}}$ -ATPyS-Psu complexes I and II. p-interacting helices of Psu are labeled. Bound ATPyS molecules and  $\text{Mg}^{2+}$  ions are shown as sticks and spheres, respectively, colored by atom type. In this and the following panels: ATPyS carbon, beige; nitrogen, blue; oxygen, red; sulfur, yellow; phosphorus, orange; magnesium, green. Rotation symbol, view relative to Fig. 1b, left.

**b**, Details of ATPyS binding at the building blocks shown in (a). Relevant protein residues are shown as sticks and colored by atom type. In this and the following panels: carbon, as the respective protein subunit; black dashed lines, hydrogen bonds or salt bridges; rotation symbols, view relative to (a).

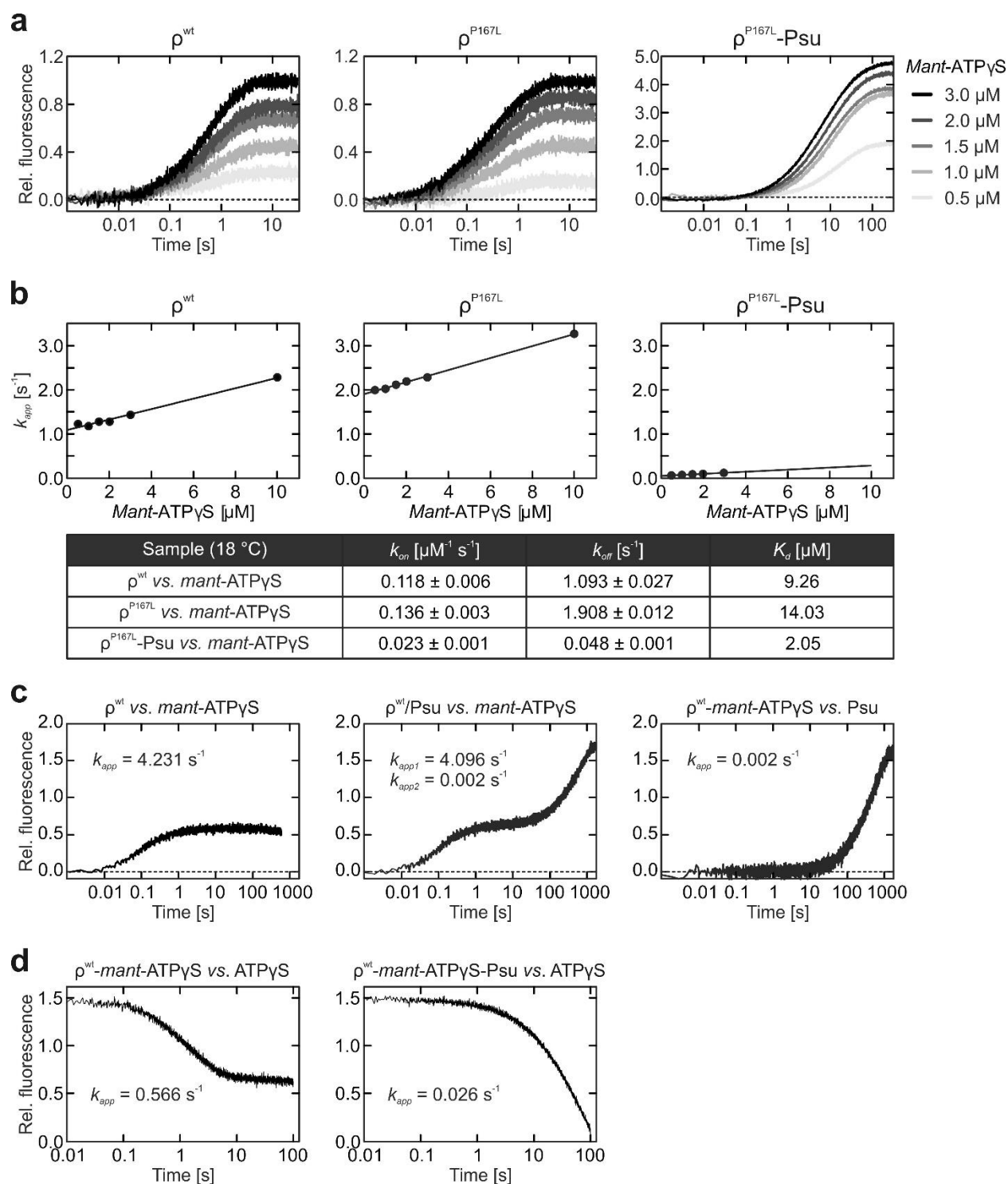
**c**, Book-view on the interacting surfaces (yellow outlines) of Psu (left) and of two neighboring  $\rho^{\text{wt}}$  subunits (right), with the electrostatic potential (blue, positive; red, negative) mapped onto the protein surfaces. Black outline, border of the bound Psu molecule.

**d,e,f**, Details of the p-Psu interaction within the building blocks shown in (a).

**g**, SDS-PAGE analyses of analytical SEC elution fractions monitoring the interaction of  $\rho^{\text{P167L}}$  with Psu variants. Elution fractions are indicated at the top. The same elution fractions were analyzed for each run and were aligned below each other. Proteins and protein mixtures analyzed are indicated above each gel. Molecular mass markers are indicated on the left. Protein bands are identified on the right. Fractions containing isolated proteins or complexes are identified at the bottom. Experiments were repeated independently at least two times with similar results.

**h**, Time traces monitoring  $\rho^{\text{wt}}$  (top) and  $\rho^{\text{P167L}}$  (bottom) RNA-stimulated ATPase activities in the absence of Psu or in the presence of 1 or 5  $\mu\text{M}$  of Psu, recorded *via* TLC. Data were recorded as biological replicates, ( $\rho^{\text{wt}}$  or  $\rho^{\text{P167L}}$  alone,  $n = 3$ ;  $\rho^{\text{wt}}$  or  $\rho^{\text{P167L}}$  plus Psu,  $n = 2$ ). Data were fitted to a single-exponential equation;  $A = A_0 + (A_{\text{max}} - A_0) * (1 - \exp(-k * t))$ ; A, fraction of ATP hydrolyzed at time

1173  $t$ ;  $A_0$ , fraction of ATP hydrolyzed at time zero;  $A_{max}$ , fraction of ATP hydrolyzed at infinite time  
1174 (amplitude);  $k$ , rate constant. Quantified amplitudes and rate constants are listed on the right.  
1175  $i$ , Inhibition of  $\rho^{wt}$  (top) and  $\rho^{P167L}$  (bottom) RNA-stimulated ATPase activities by increasing  
1176 concentrations of  $Psu^{wt}$ ,  $Psu^{R43A}$  or  $Psu^{D49R}$ . Data represent biological replicates,  $n = 2$ . Data were  
1177 fitted to an [inhibitor] vs. response function,  $A = A_{min} + (A_{max} - A_{min}) / (1 + ([inhibitor] / IC_{50}))$ ;  $A$ , fraction of  
1178 ATP hydrolyzed at a given  $Psu$  (inhibitor) concentration;  $A_{min}$  and  $A_{max}$ , fitted minimum and  
1179 maximum fractions of ATP hydrolyzed. Quantified  $IC_{50}$  values are listed on the right; n.d., not  
1180 determined.  
1181



**Fig. 4: Effects of Psu on nucleotide binding by  $\rho$ .**

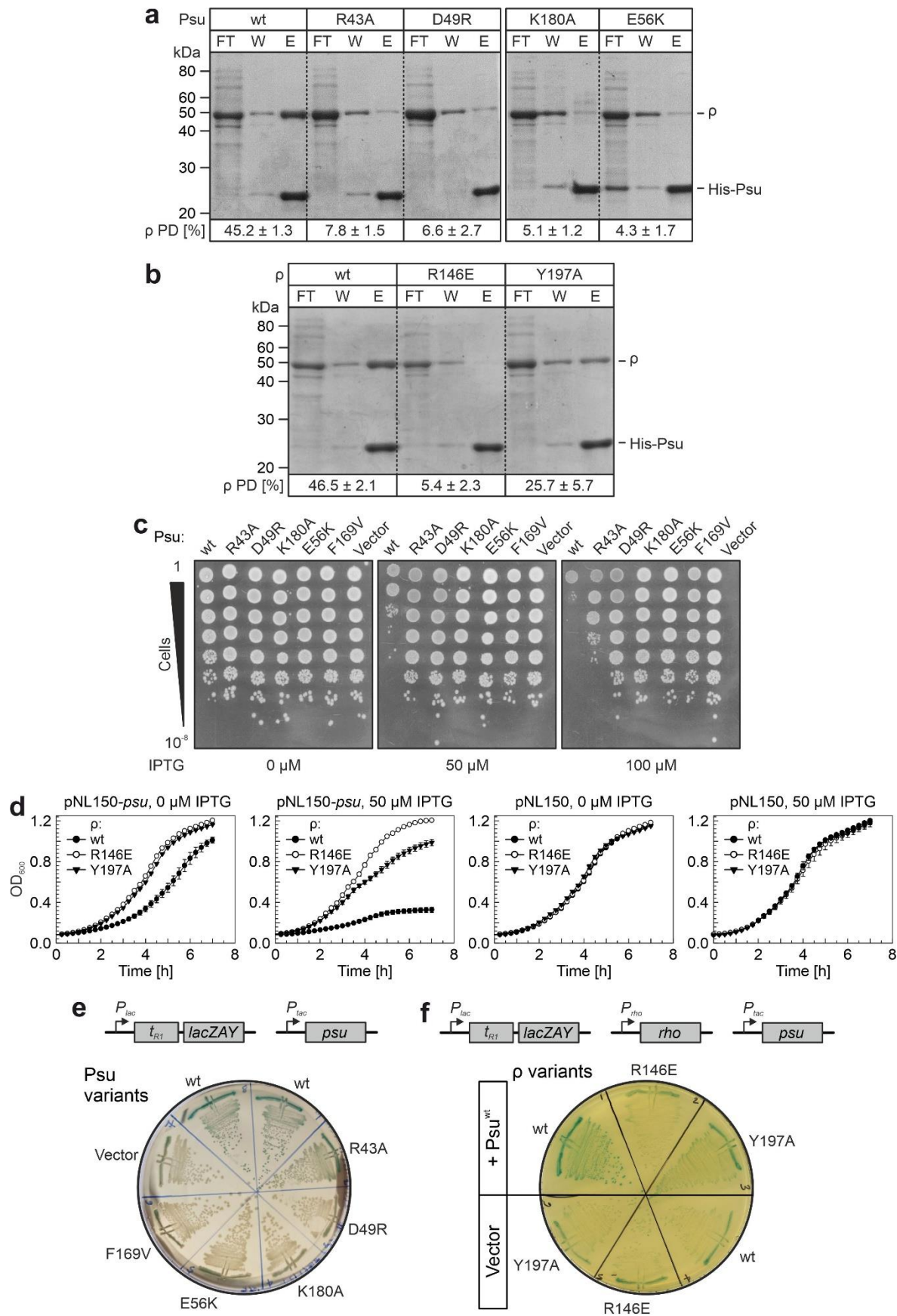
**a**, Time traces monitoring  $\rho^{wt}$  (left),  $\rho^{P167L}$  (middle) or  $\rho^{P167L}$ -Psu (right) binding to increasing

concentrations of *mant*-ATPγS (0.5 to 3 μM; colored from light grey to black), recorded via stopped-flow/FRET measurements. Data represent means, n = 3-6. Data were fit to a single-exponential equation;  $A = A_0 + (A_{max} - A_0) * (1 - \exp(-k_{app} * t))$ ; A, fluorescence at time t;  $A_{max}$ , final signal;  $A_0$ , initial fluorescence signal;  $k_{app}$ , characteristic time constant.

**b**, Linear regression ( $k_{app} = k_1[mant\text{-}ATP\gamma S] + k_{-1}$ ) of the apparent rate constants,  $k_{app}$ , at increasing *mant*-ATPγS concentrations, derived by the single exponential fitting of the time-traces depicted in (a). The association ( $k_{on}$ ) and dissociation ( $k_{off}$ ) rate constants were derived from the slopes and y-intercepts.  $K_d$ 's were calculated as the ratios of the  $k_{off}$  and  $k_{on}$  values.

**c**, Time traces monitoring  $\rho^{wt}$  (left) or  $\rho^{wt}/Psu$  (middle) binding to *mant*-ATPγS, and of  $\rho^{wt}$ -*mant*-ATPγS binding to Psu (right), recorded via stopped-flow/FRET measurements. Data represent means, n = 3-6.  $k_{app}$  values depicted in the graphs were derived by fitting of the data to a single-exponential equation as in (a) ( $\rho^{wt}$  or  $\rho^{wt}/Psu$  binding to *mant*-ATPγS), or by fitting of the data to a double exponential equation ( $\rho^{wt}/Psu$  binding to *mant*-ATPγS);  $A = A_0 + A_1(1 - \exp(-k_{app1} * t)) + A_2(1 - \exp(-k_{app2} * t))$ ; A, fluorescence at time t;  $A_0$ , initial fluorescence signal;  $A_1$  and  $A_2$ , amplitudes of the signal change;  $k_{app1}$  and  $k_{app2}$ , characteristic time constants.

**d**, Time traces monitoring the dissociation (replacement with ATPγS) of *mant*-ATPγS from  $\rho^{wt}$ -*mant*-ATPγS-Psu (left) or  $\rho^{P167L}$ -*mant*-ATPγS-Psu (right), recorded via stopped-flow/FRET measurements. Data represent means, n = 3-6.  $k_{app}$  values depicted in the graphs were derived by fitting of the data to a single-exponential equation;  $A = A_{min} + (A_0 - A_{min})\exp(-k_{app} * t)$ ; A, fluorescence at time t;  $A_{min}$ , final signal;  $A_0$ , initial fluorescence signal;  $k_{app}$ , characteristic time constant.



**Fig. 5: *In vivo* defects of Psu and  $\rho$  variants.**

**a**, *In vivo* binding defects of Psu variants. *E. coli* cell lysates after production of N-terminally His<sub>6</sub>-tagged Psu variants and  $\rho^{wt}$  were passed through Ni<sup>2+</sup>-NTA beads, and the flow-through (FT), wash (W), and eluted (E) fractions were collected. The fraction of pulled-down (PD)  $\rho^{wt}$  was calculated as  $(E/(FT+W+E)) \times 100$ . Psu variants are indicated above the gels. Molecular mass markers are indicated on the left. Protein bands are identified on the right. Quantified data ( $\rho$  PD [%]) represent means  $\pm$  SD; n = 3.

**b**, *In vivo* binding defects of  $\rho$  variants. *E. coli* cell lysates after production of N-terminally His<sub>6</sub>-tagged Psu<sup>wt</sup> and  $\rho$  variants were passed through Ni<sup>2+</sup>-NTA beads, and the flow-through (FT), wash (W), and eluted (E) fractions were collected. The fraction of pulled-down (PD)  $\rho$  variants was calculated as  $(E/(FT+W+E)) \times 100$ .  $\rho$  variants are indicated above the gel. Molecular mass markers are indicated on the left. Protein bands are identified on the right. Quantified data ( $\rho$  PD [%]) represent means  $\pm$  SD; n = 3.

**c**, Serial dilutions of *E. coli* overnight cultures producing Psu variants (indicated above the plate images) were spotted on LB agar plates with different IPTG concentrations (indicated below the plate images). Serial dilutions are indicated on the left. Experiments were repeated independently at least two times with similar results.

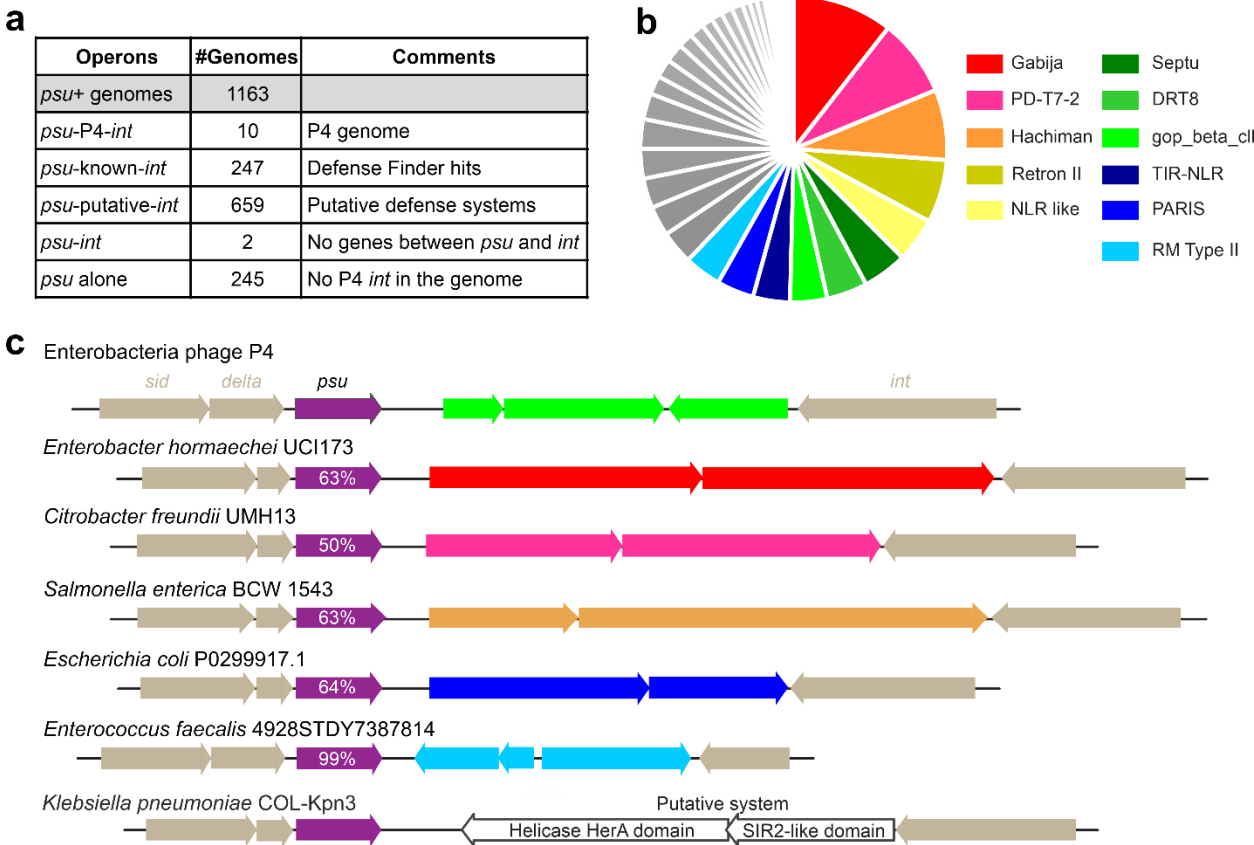
**d**, *E. coli* MG1655 $\Delta\rho\Delta rac$  producing the indicated  $\rho$  variants were transformed with a plasmid guiding production of Psu<sup>wt</sup> (first and second panel). As a control, strains were transformed with an empty vector (third and fourth panel). Growth curves were recorded in the absence of IPTG or in the presence of 50  $\mu$ M IPTG, as indicated above the panels. Data points represent means  $\pm$  SEM; n = 4-5.

**e**, *E. coli* MG1655 $\Delta rac\Delta lac$  with a  $P_{lac}-t_{R1}-lacZYA$  reporter were transformed with plasmids guiding the production of the indicated Psu variants, or with an empty vector. Transformants were streaked on LB-X-gal agar plates supplemented with 15  $\mu$ M IPTG. The reporter system produced

blue-green colonies when the *in vivo* p-dependent termination function was inhibited. Experiments were repeated independently at least two times with similar results.

*f*, *E. coli* MG1655 $\Delta\rho\Delta rac\Delta lac$  with a  $P_{lac}-t_{R1}-lacZYA$  reporter and producing the indicated p variants were further transformed with a plasmid guiding the production of  $Psu^{wt}$ , or with an empty vector. Transformants were streaked on LB-X-gal agar plates supplemented with 15  $\mu$ M IPTG. The reporter system produced blue-green colonies when the *in vivo* p-dependent termination function was inhibited. Experiments were repeated independently at least two times with similar results.





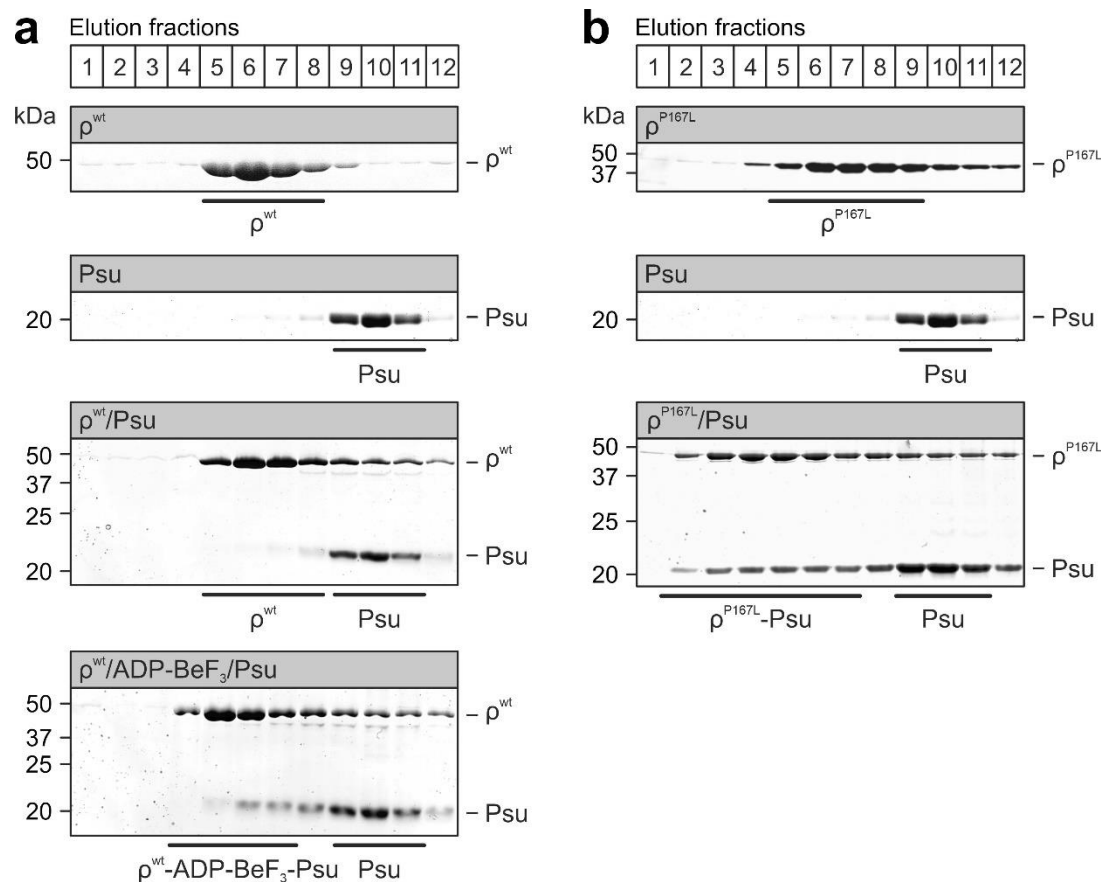
**Fig. 6: Psu as a marker of anti-phage defense systems.**

**a**, Defense system analysis of *psu*<sup>+</sup> genomes. See Supplementary Dataset 1 for genomic context details. We note that the absence of the *int* gene could be due to incomplete genome assembly in some cases.

**b**, The 37 classes of *psu*-associated defense systems identified by DefenseFinder are shown in a pie chart, with the eleven most abundant defense systems indicated in color; see Extended Data Fig. 8 for a complete list.

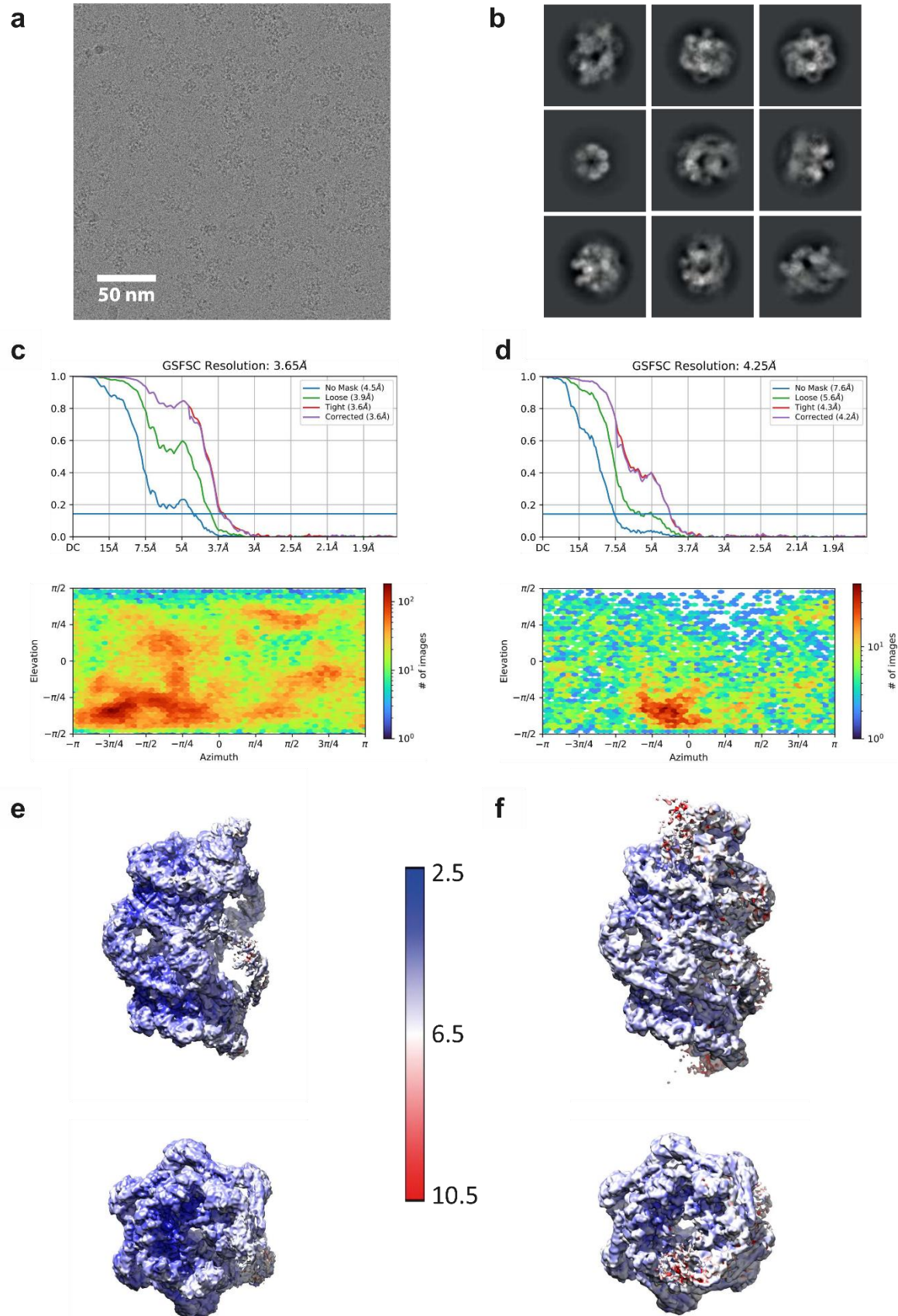
**c**, Examples of *psu*-associated defense system loci; % identity to P4 Psu is shown. The known defense genes are colored as in **b**; a putative defense system with a SIR2-like domain is shown in white.

# Extended data



## Extended Data Fig. 1: $\rho$ -Psu interaction in solution.

**a,b**, SDS-PAGE analyses of analytical SEC elution fractions monitoring the interaction of  $\rho^{wt}$  (a) or  $\rho^{P167L}$  (b) with Psu. Elution fractions are indicated at the top. The same elution fractions were analyzed for each run and were aligned below each other. Proteins and protein mixtures analyzed are indicated above each gel. Molecular mass markers are indicated on the left. Protein bands are identified on the right. Fractions containing isolated proteins or complexes are identified below each gel. Experiments were repeated independently at least two times with similar results.



1269

1270 **Extended Data Fig. 2: CryoEM/SPA analysis of a  $\rho^{\text{wt}}$ -ATPyS-Psu complexes.**

1271 **a**, Representative cryoEM micrograph of the  $\rho^{\text{wt}}$ -ATPyS-Psu complexes. Scale bar, 50 nm. 6,266

1272 micrographs were recorded, particle images were picked from 5,986 high-quality micrographs.

1273 **b**, 2D class averages of  $\rho^{\text{wt}}$ -ATPyS-Psu particle images after reference-free 2D classification.

1274 **c,d**, Upper panels, global resolution estimation for the  $\rho^{\text{wt}}$ -ATPyS-Psu cryoEM reconstructions

1275 (complex I, left; complex II, right) by gold-standard Fourier shell correlation (FSC). Blue line,

1276 FSC<sub>0.143</sub>. Lower panels, viewing direction distribution plots of the particle images used for the final

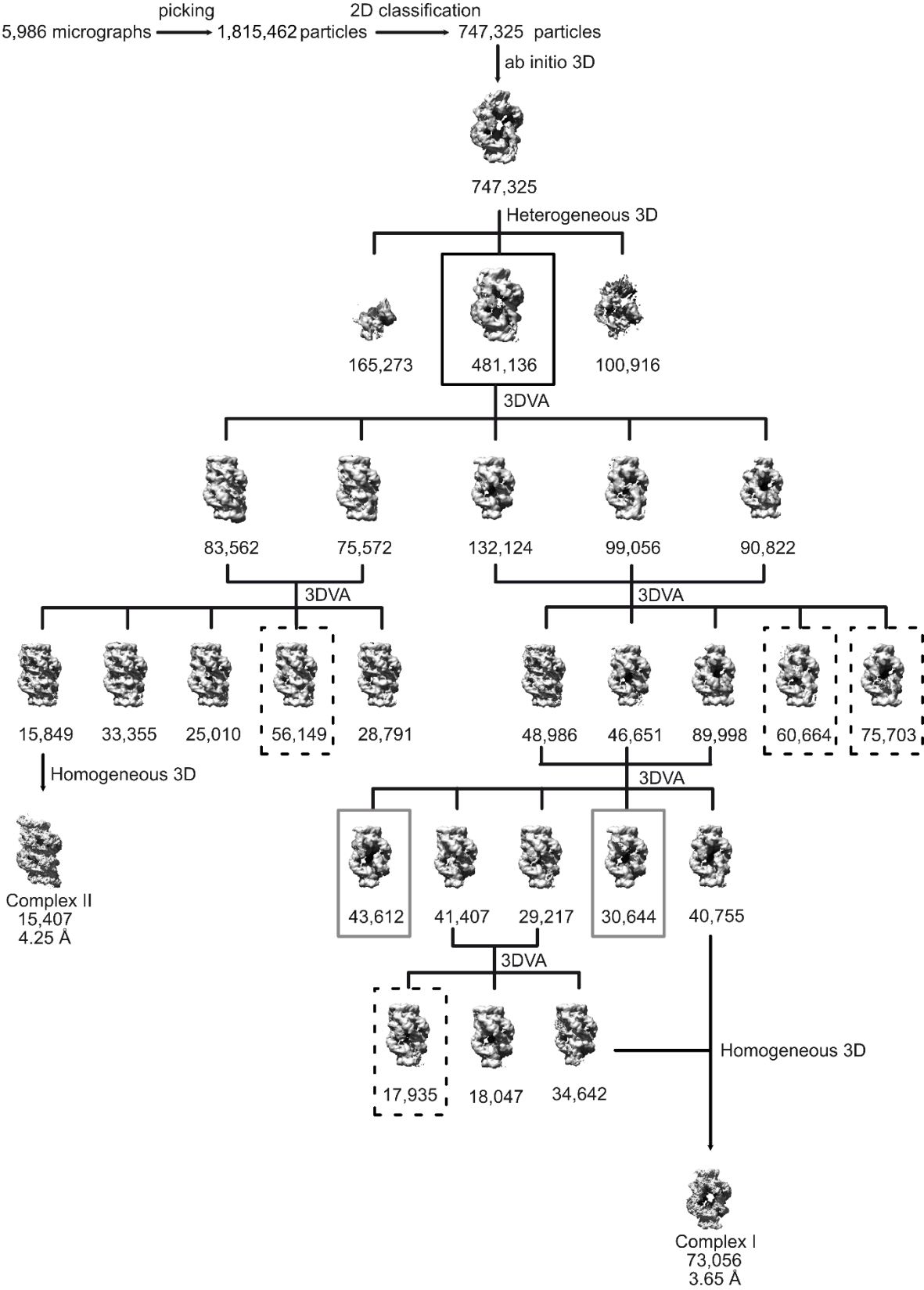
1277  $\rho^{\text{wt}}$ -ATPyS-Psu cryoEM reconstructions (complex I, left; complex II, right) as obtained during NU

1278 refinement with cryoSPARC.

1279 **e,f**, Local resolution estimation as determined with cryoSPARC, ranging from 2.5 Å to 10.5 Å for

1280 the  $\rho^{\text{wt}}$ -ATPyS-Psu cryoEM reconstructions (complex I, left; complex II, right).

1281

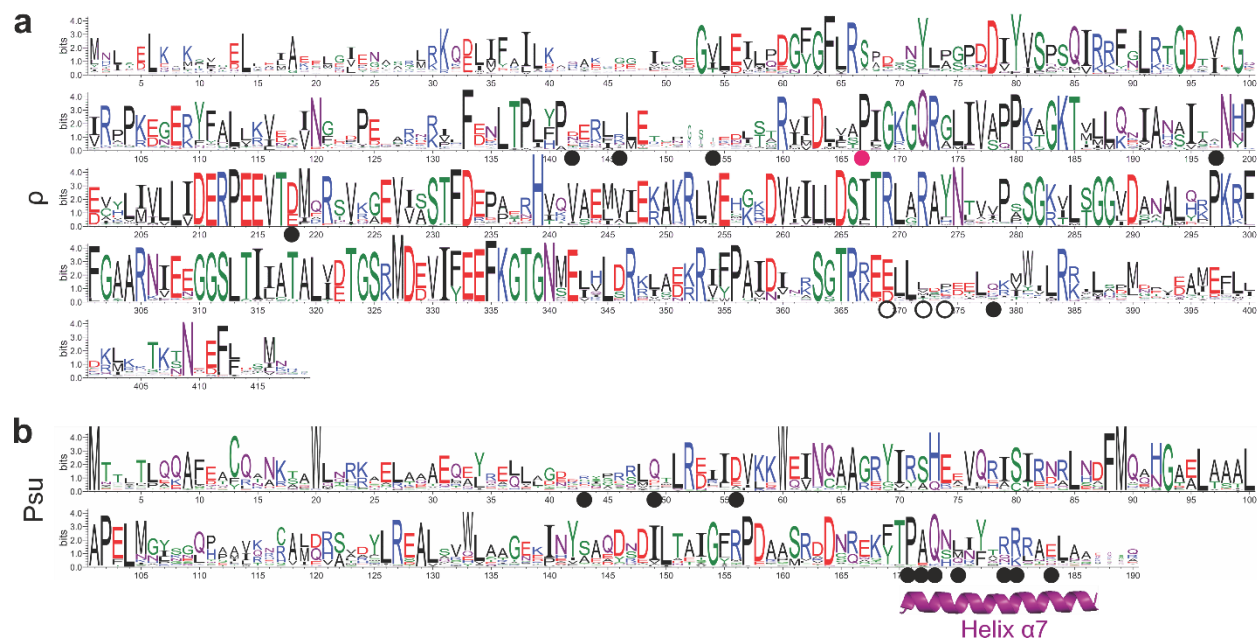


1282

1283

**Extended Data Fig. 3:  $\rho^{\text{wt}}$ -ATPyS-Psu cryoEM data refinement.**

1,815,462 particle images were picked from 5,986 micrographs and subjected to reference-free 2D classification. 747,325 particle images were selected for iterative cycles of heterogeneous 3D refinement into 3 classes. The best-appearing class, consisting of 481,136 particle images, was selected and subjected to 3D variability analysis (3DVA) generating five clusters. Two of these were combined (left branch) and further classified by 3DVA into five classes, one of which, consisting of 15,849 particles was subjected to homogeneous 3D refinement, which yielded a reconstruction at 4.25 Å resolution (complex II). The residual clusters (right branch) were further classified by iterative 3DVA cycles. Two of the resulting classes appeared virtually identical and were combined for homogeneous refinement yielding a reconstruction at 3.65 Å resolution (complex I). The reconstructions boxed in solid gray boxes represent assemblies with altered rotational orientation of Psu-bridged  $\rho^{\text{wt}}$  rings, equivalent to  $\rho^{\text{P167L}}$ -ATPyS-Psu complex I. The reconstructions boxed in dashed boxes represent assemblies with intermediate oligomeric states.

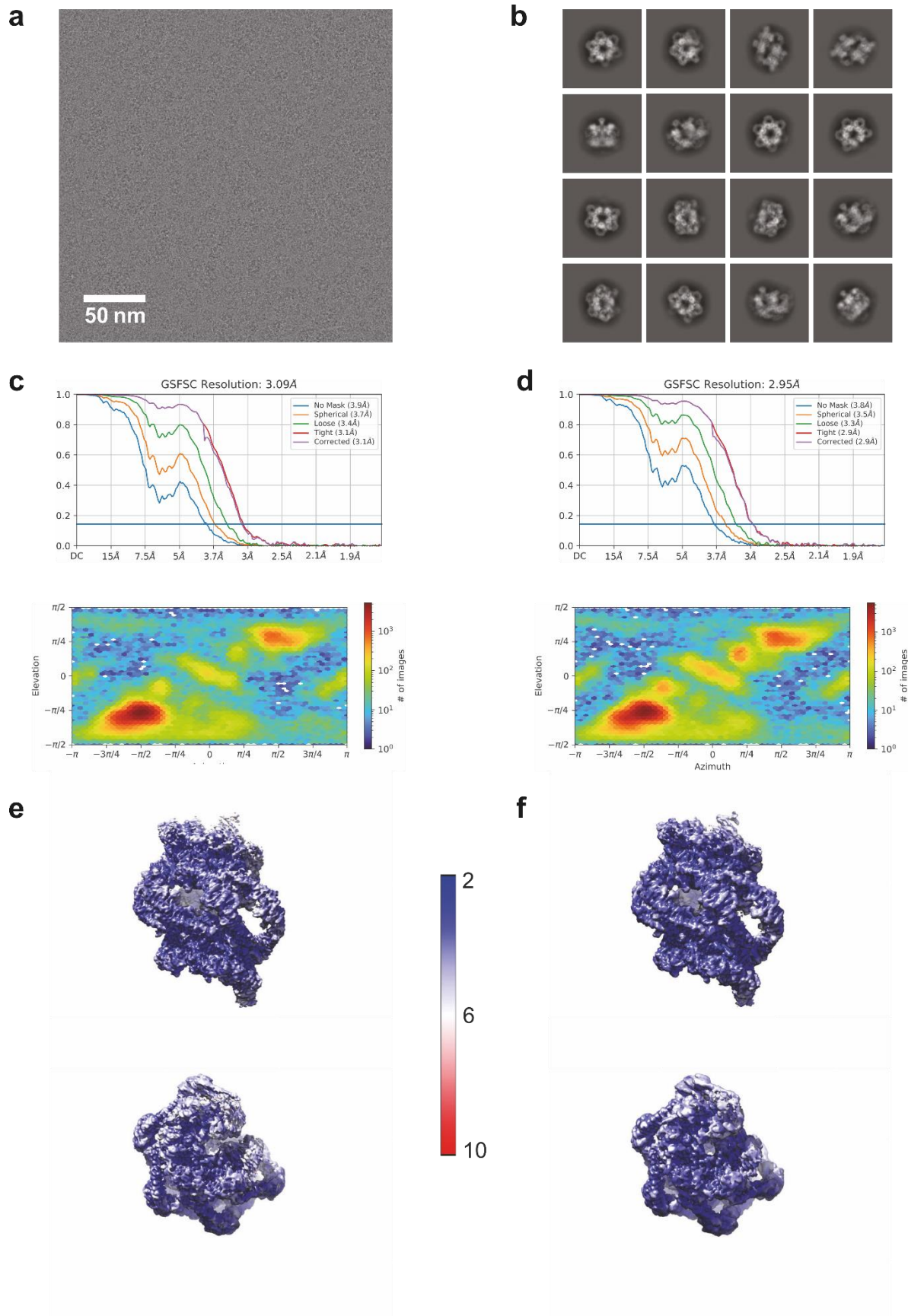


# **Extended Data Fig. 4: $\rho$ and Psu sequence conservation.**

**a**, Conservation pattern of  $\rho$  based on a multiple sequence alignment of  $\rho$  proteins from representatives of the entire Bacteria kingdom (Supplementary Dataset 1A). The universally conserved P167 residue is marked by a magenta sphere. Residues exhibiting side chain interactions between *E. coli*  $\rho$  and phage P4 Psu are marked by black spheres; residues exhibiting backbone interactions are marked by white spheres.

**b**, Conservation pattern of Psu proteins mined in this study (Supplementary Dataset 1B).





1309

1310 **Extended Data Fig. 5: CryoEM/SPA analysis of  $\rho^{P167L}$ -ATPyS-Psu complexes.**

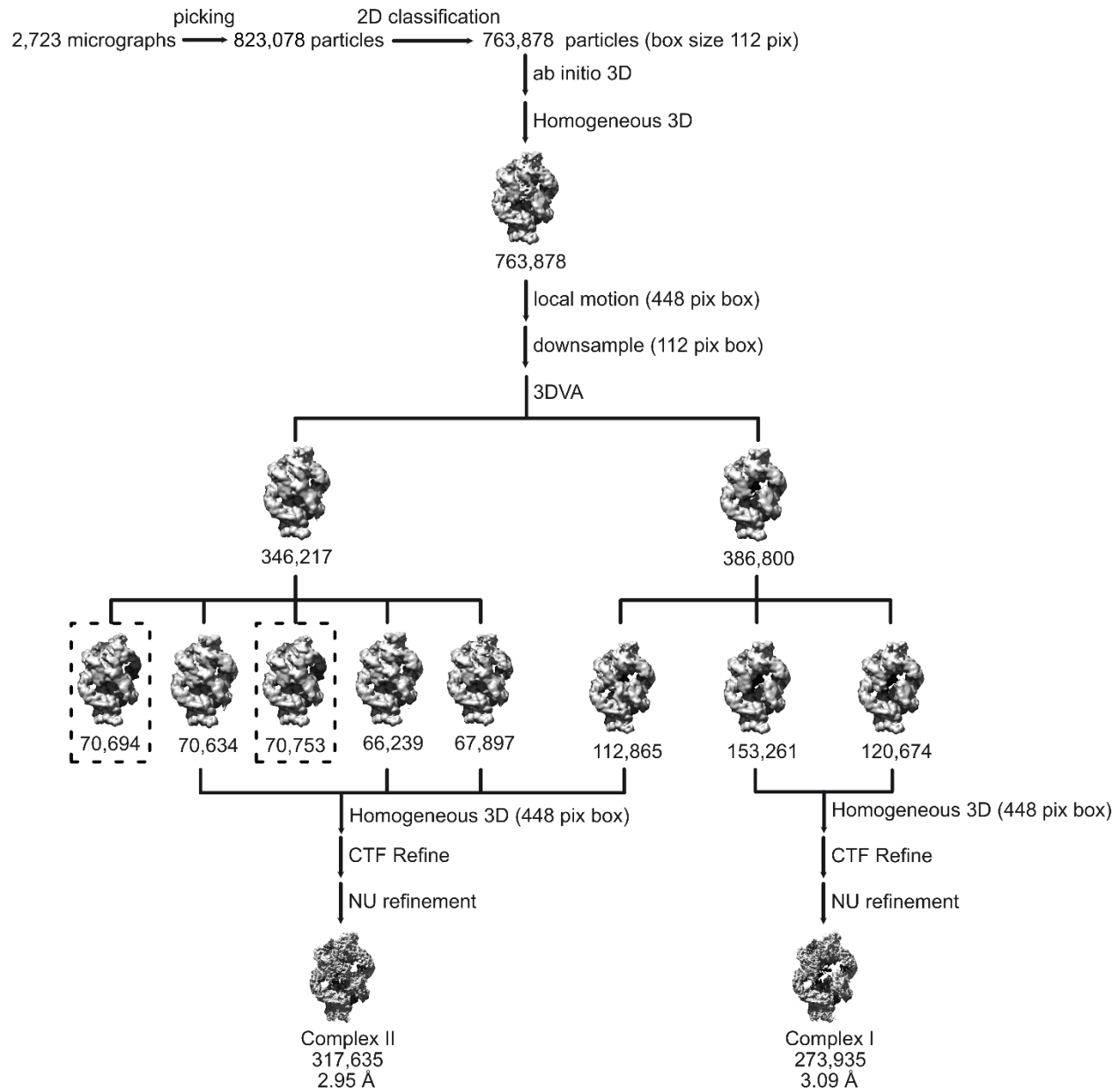
1311 **a**, Representative cryoEM micrograph of the  $\rho^{P167L}$ -ATPyS-Psu complexes. Scale bar, 50 nm.  
1312 2,766 micrographs were recorded, particle images were picked from 2,723 high-quality  
1313 micrographs.

1314 **b**, 2D class averages of  $\rho^{P167L}$ -ATPyS-Psu particle images after reference-free 2D classification.

1315 **c,d**, Upper panels, global resolution estimation for the  $\rho^{P167L}$ -ATPyS-Psu cryoEM reconstructions  
1316 (complex I, left; complex II, right) by gold-standard Fourier shell correlation (FSC). Blue line,  
1317 FSC<sub>0.143</sub>. Lower panels, viewing direction distribution plots of the particle images used for the final  
1318  $\rho^{P167L}$ -ATPyS-Psu cryoEM reconstructions (complex I, left; complex II, right) as obtained during  
1319 NU refinement with cryoSPARC.

1320 **e,f**, Local resolution estimation as determined with cryoSPARC, ranging from 2.5 Å to 10 Å for  
1321 the  $\rho^{P167L}$ -ATPyS-Psu cryoEM reconstructions (complex I, left; complex II, right).

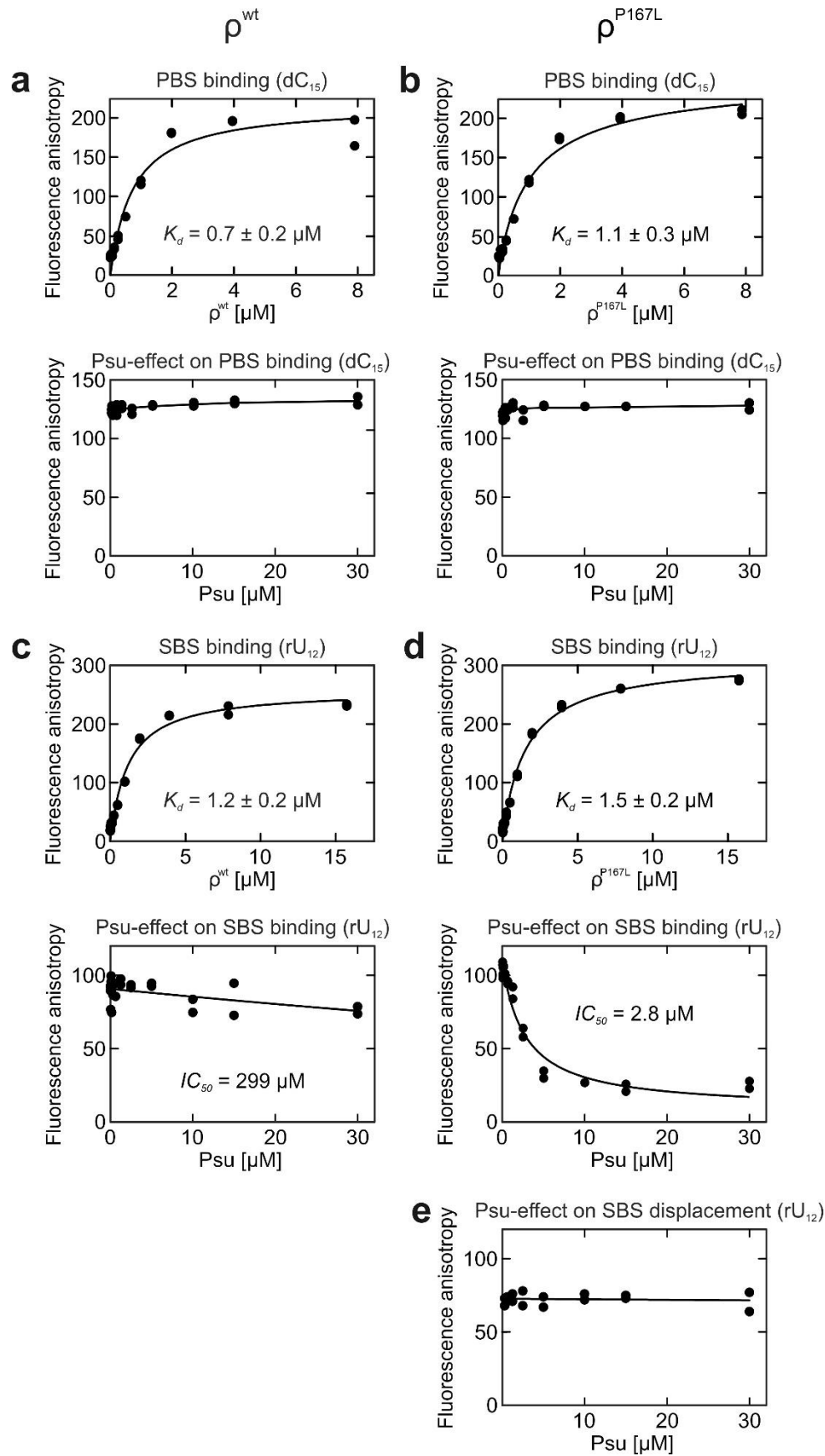
1322



# Extended Data Fig. 6: $p^{P167L}$ -ATPyS-Psu cryoEM data refinement.

823,078 particle images were picked from 2,723 micrographs and subjected to reference-free 2D classification. 763,878 particle images were selected and used for *ab initio* 3D reconstruction, followed by homogeneous refinement. Particles were subjected to local motion correction and re-extracted with a box size of 448 px. After down-sampling to a box size of 112 px, 3DVA was applied to separate the dataset into two classes. Each of these was individually subjected to

1331 3DVA, yielding five and three classes, respectively. Similar appearing classes of both branches  
1332 were combined and homogeneously refined followed by global and local CTF refinement. Final  
1333 NU refinement yielded reconstructions of 317,635 particle images at 2.95 Å resolution (complex  
1334 II) and 273,935 particle images at 3.09 Å resolution (complex I). The remaining two classes  
1335 highlighted by dashed boxes represent additional assemblies.  
1336



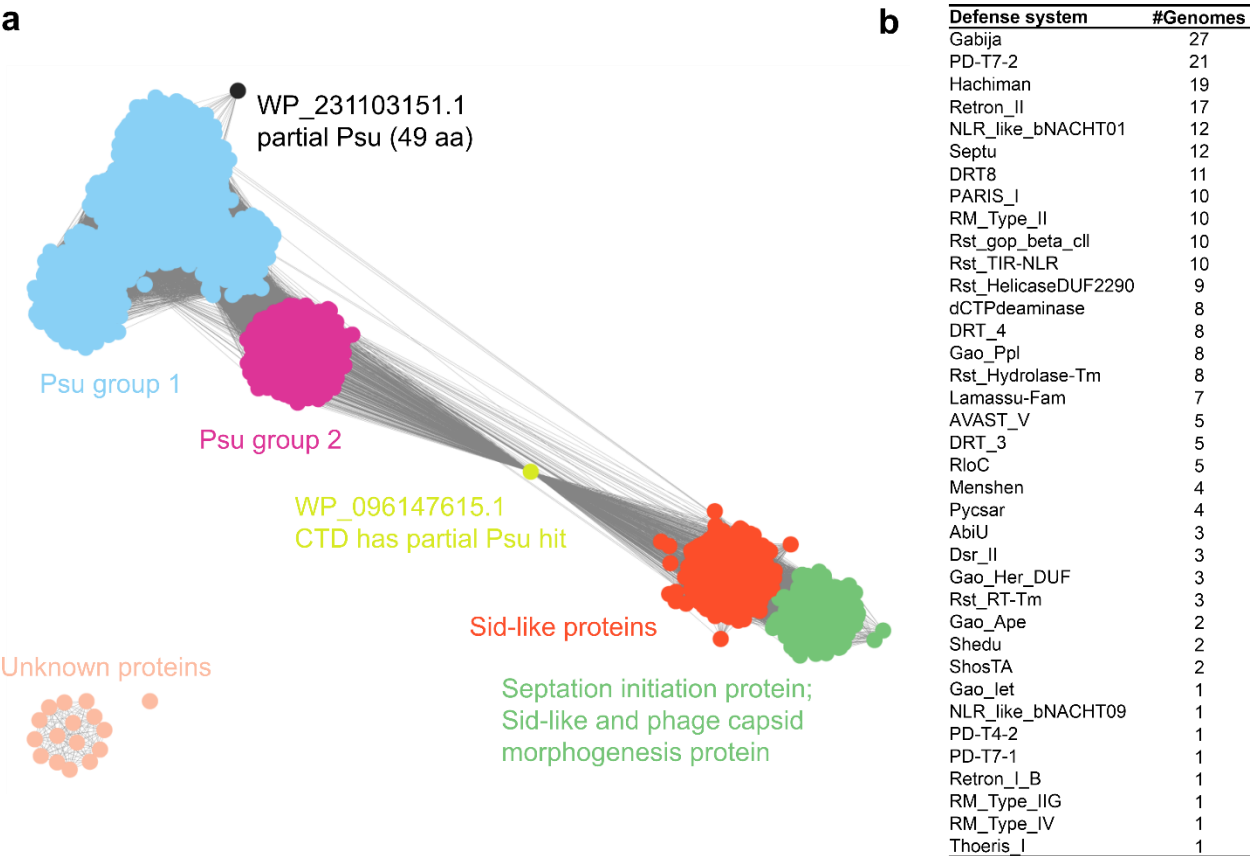
**Extended Data Fig. 7: Effects of Psu on nucleic acid binding at the p PBS and SBS.**

**a,b**, Fluorescence anisotropy measurements monitoring the interaction of increasing concentrations of  $\rho^{\text{wt}}$  (**a**) or  $\rho^{\text{P167L}}$  (**b**) with dC<sub>15</sub>-FAM at the PBSes (top), or the effect of increasing concentrations of Psu on  $\rho^{\text{wt}}$  (**a**) or  $\rho^{\text{P167L}}$  (**b**) dC<sub>15</sub>-FAM binding at the PBSes (bottom).

**c,d**, Fluorescence anisotropy measurements monitoring the interaction of increasing concentrations of  $\rho^{\text{wt}}$  (**c**) or  $\rho^{\text{P167L}}$  (**d**) with rU<sub>12</sub>-FAM substrate at the SBS (top), or the effect of increasing concentrations of Psu on  $\rho^{\text{wt}}$  (**c**) or  $\rho^{\text{P167L}}$  (**d**) rU<sub>12</sub>-FAM binding at the SBS (bottom).

**e**, Lack of Psu-mediated release of SBS-bound RNA.

Data were recorded as technical duplicates,  $n = 2$ . Data in (**a-d**, top) were fit to a single-exponential Hill equation for dC<sub>15</sub>-FAM or rU<sub>12</sub>-FAM association at p PBSes or SBS.  $A = A_{\text{max}}[\text{protein}]^h / (K_d^h + [\text{protein}]^h)$ ; in which  $A_{\text{max}}$  is the fitted maximum of nucleic acid bound;  $K_d$  is the dissociation constant;  $h$  is the Hill coefficient. Data in (**c,d**, bottom) were fit to an [inhibitor] vs. response equation:  $A = A_{\text{min}} + (A_{\text{max}} - A_{\text{min}}) / (1 + ([\text{inhibitor}] / IC_{50}))$ ; in which  $A$  is the anisotropy signal at a given concentration of Psu (inhibitor);  $A_{\text{min}}$  and  $A_{\text{max}}$  are the fitted minimum and maximum of nucleic acid bound.



**Extended Data Fig. 8: Psu homologs and associated defense systems.**

**a**, The two Psu groups from the 'greedy search'. See Methods for details.

**b**, A complete count of the DefenseFinder hits.



# Supplementary information

## Widespread gene regulator Psi inhibits transcription termination factor p by forced hyper-oligomerization

Daniela Gjorgjevikj<sup>1</sup>, Naveen Kumar<sup>2</sup>, Bing Wang<sup>3</sup>, Tarek Hilal<sup>1,4</sup>, Nelly Said<sup>1</sup>, Bernhard Loll<sup>1</sup>, Irina Artsimovitch<sup>3</sup>, Ranjan Sen<sup>2,5</sup>, Markus C. Wahl<sup>1,6,\*</sup>

<sup>1</sup> Freie Universität Berlin, Institute of Chemistry and Biochemistry, Laboratory of Structural Biochemistry, Takustr. 6, D-14195 Berlin, Germany

<sup>2</sup> Laboratory of Transcription, Centre for DNA Fingerprinting and Diagnostics, Inner Ring Road, Uppal, Hyderabad-500039, India.

<sup>3</sup> Department of Microbiology and Center for RNA Biology, The Ohio State University, Columbus, OH, USA

<sup>4</sup> Freie Universität Berlin, Institute of Chemistry and Biochemistry, Research Center of Electron Microscopy and Core Facility BioSupraMol, Fabeckstr. 36a, 14195 Berlin, Germany

<sup>5</sup> School of Biological Sciences, Indian Association for the Cultivation of Science, 2A & B, Raja S. C. Mullick Road, Jadavpur, Kolkata 700032, India

<sup>6</sup> Helmholtz-Zentrum Berlin für Materialien und Energie, Macromolecular Crystallography, Albert-Einstein-Str. 15, D-12489 Berlin, Germany

\* Correspondence to: markus.wahl@fu-berlin.de

# 1383 Supplementary Tables

## 1384 Supplementary Table 1: CryoEM data collection, refinement and validation statistics.

	p <sup>wt</sup> -ATPyS-Psu		p <sup>P167L</sup> -ATPyS-Psu	
	Complex I (PDB 8PEU) (EMDB 17637)	Complex II (PDB 8PEW) (EMDB 17639)	Complex I (PDB 8PEX) (EMDB 17640)	Complex II (PDB 8PEY) (EMDB 17641)
Data collection and processing				
Microscope	FEI Titan Krios G3i			
Voltage [keV]	300			
Camera	Falcon 3EC			
Magnification (nominal / calibrated)	96,000	96,000	96,000	96,000
Pixel size at detector [Å/pixel]	0.832	0.832	0.832	0.832
Total electron exposure [e <sup>-</sup> /Å <sup>2</sup> ]	42	42	42	42
Exposure rate [e <sup>-</sup> /pixel/s]	0.7	0.7	0.7	0.7
Frames collected during exposure	33			
Defocus range [μm]	0.80 - 2			
Automation software	EPU (version 2.8.1)			
Micrographs collected	6,066	6,066	2,766	2,766
Micrographs used	5,986	5,986	2,723	2,723
Total extracted particles	1,815,462	1,815,462	823,078	823,078
Final particles	73,056	15,407	273,935	317,635
Point-group or helical symmetry parameters	C1	C1	C1	C1
Global resolution [Å]; FSC <sup>a</sup> <sub>0.143</sub> (unmasked / masked)	4.5 / 3.6	7.6 / 4.2	3.9 / 3.1	3.8 / 2.9
Local resolution range [Å]	1.8-30.0	3.1-30.0	2.2-30.0	2.1-30.0
Map sharpening <i>B</i> factor [Å <sup>2</sup> ] / ( <i>B</i> factor range)	-82	-57	-81	-83
Map sharpening method	Local <i>B</i> factor			
Refinement package	PHENIX (version 1.20_4459); real.space.refine			
Model composition				
Non-hydrogen atoms	57,632	83,524	54,751	55,957
ρ copies / residues	12 / 39,588	18 / 52,588	12 / 39,588	13 / 41,172
Psu copies / residues	10 / 14,780	16 / 23,468	10 / 14,780	10 / 14,780
ATPyS molecules	12	16	12	12
Mg <sup>2+</sup> ions	12	16	12	12
Model Refinement				
Model-map scores				
CC <sup>b</sup> (mask)	0.71	0.72	0.83	0.85
CC (volume)	0.69	0.68	0.82	0.85
Average <i>B</i> factors [Å <sup>2</sup> ]				
Overall	162	216	167	150
ρ copies	162	226	171	149
Psu copies	162	190	156	153
ATPyS molecules	202	277	130	129
Mg <sup>2+</sup> ions	221	282	113	120
Rmsd <sup>c</sup> from ideal values				
Bond lengths [Å]	0.002	0.003	0.003	0.003
Bond angles [°]	0.621	0.713	0.634	0.606
Validation <sup>d</sup>				
MolProbity score	1.72	1.96	1.15	1.69
CaBLAM outliers [%]	0.9	1.0	1.1	1.2
Clashscore	14.3	18.7	13.8	13.8
Poor rotamers [%]	0.0	0.0	0.0	0.0
Cβ deviations	0	0	0	0

EMRinger score	0.78	-0.26	1.56	1.50
<b>Ramachandran plot</b>				
Favored [%]	97.7	96.9	98.1	97.8
Allowed [%]	2.3	3.1	1.9	2.2
Outliers [%]	0.0	0.0	0.0	0.0
<b>Ramachandran Z-score (rmsd)</b>				
Whole	0.78 (0.10)	-0.10 (0.08)	0.82 (0.10)	0.75 (0.10)
Helix	1.18 (0.07)	0.48 (0.07)	1.28 (0.09)	1.32 (0.09)
Sheet	0.34 (0.19)	0.00 (0.16)	0.25 (0.19)	-0.12 (0.18)
Loop	-0.48 (0.13)	-0.84 (0.10)	-0.40 (0.12)	-0.58 (0.13)

1385

1386

1387

1388

1389

1390

- a FSC, Fourier shell correlation
- b CC, correlation coefficient
- c Rmsd, root-mean-square deviation
- d Using MolProbity<sup>91</sup>

**Supplementary Table 2: Defective Psu and p variants.**

<b>Psu variants defective in p binding/inhibition</b>		
<b>Variant</b>	<b>Likely effect</b>	<b>Reference</b>
L21P	Destabilization of the $\alpha 1/\alpha 2$ coiled-coil	42
V45F	Local mis-folding, disruption of R43 and D49 interactions with Y197 and R146 of p	16,19
S72L	Destabilization of the $\alpha 1/\alpha 2$ coiled-coil	42
E56K	Disruption of the interaction with the backbone amide of p Q374	42
$\Delta 95-98$	Disruption of Psu dimerization	37
$\Delta 97-100$	Disruption of Psu dimerization	
P157L	Destabilization of Psu dimer	42
P157S	Destabilization of Psu dimer	42
R166C	Effect on helix $\alpha 7$ positioning	42
R166P	Effect on helix $\alpha 7$ positioning	42
F169V	Effect on helix $\alpha 7$ positioning through stacking with W60	37
$\Delta$ CTD10	Partial deletion of helix $\alpha 7$ interaction interface	38
$\Delta$ CTD20	Deletion of helix $\alpha 7$ interaction interface	38
<b>p variants defective in Psu binding/resistant to Psu</b>		
<b>Variant</b>	<b>Likely effect</b>	<b>Reference</b>
R144E	Local mis-folding, effect on interaction of R146 with D49 of Psu	33
$\Delta 144-148$	Local mis-folding, disruption of interaction with D49 of Psu	33
R146E	Disruption of interaction with D49 of Psu	33
E148R	Local mis-folding	33
R149E	Disruption of stacking with Y197 of p to position it for interaction with R43 of Psu	33
$\Delta 149-153$	Local mis-folding, effect on the positioning of Y197 of p for interaction with R43 of Psu	33
N151D	Effect on the positioning of Y197 of p for interaction with R43 of Psu	33

1394 **Supplementary Table 3: PCR primers (5' to 3').**

Primer	Sequence	Usage
DG30	TCAGGAGAAGGCGCAAACGTCAGTCGC	FP <sup>a</sup> to generate <i>psu</i> -R43A in pET21b
DG31	GCGACTGACGTTTGCGCCTTCTCCTGA	RP <sup>b</sup> to generate <i>psu</i> -R43A in pET21b
DG32	GTCAGTCGCCTGCGCGAATTACGCAAT	FP to generate <i>psu</i> -D49R in pET21b
DG31	ATTGCGTAATTCGCGCAGGCGACTGAC	RP to generate <i>psu</i> -D49R in pET21b
RS84new	GCG CGCGCCATATGAATCTTACCGAATTAAAG	FP upstream of <i>rho</i> with <i>Nde</i> I site to generate <i>rho</i> -Y197A in pET21b
RS85	GCGCGCCTCGAGTTATGAGCGTTTCATCATTTTC	RP downstream of <i>rho</i> with <i>Xho</i> I site and with stop codon to generate <i>rho</i> -Y197A in pET21b
RS123	GCGCGCGCCATATGGAAAGCACAGCCTTACAGCAGGCC	FP upstream of <i>psu</i> with <i>Nde</i> I site to generate <i>psu</i> -D49R, <i>psu</i> -R43A and <i>psu</i> -K180A in pET28b
RS124	GCGCGCCTCGAGCACTGACTGACGTGATGCCAGTTGC	RP downstream of <i>psu</i> with <i>Xho</i> I and without stop codon to generate <i>psu</i> -D49R, <i>psu</i> -R43A and <i>psu</i> -K180A in pET28b
RS2178	CAGAAACGTCAGTCGCCTGCGCGAATTACGCAATATTATCG	FP to generate <i>psu</i> -D49R by SDM <sup>c</sup>
RS2179	CGATAATATTGCGTAATTCGCGCAGGCGACTGACGTTTCTG	RP to generate <i>psu</i> -D49R by SDM
RS2180	TCTGTCAGGAGAAGGCGCAAACGTCAGTCGCCTG	FP to generate <i>psu</i> -R43A by SDM
RS2181	CAGGCGACTGACGTTTGCGCCTTCTCCTGACAGA	RP to generate <i>psu</i> -R43A by SDM
RS2182	AGAACATGATTTTTTCGCGTGCAAGTGC GCAACTGGCATCAC	FP to generate <i>psu</i> -K180A by SDM
RS2183	GTGATGCCAGTTGCGCACTTGACGCGGAAAAAATCATGTTCT	RP to generate <i>psu</i> -K180A by SDM
RS2224	GAGCATTGCTGCCAACCACCCGG	FP to generate <i>rho</i> -Y197A by SDM
RS2225	TGAGCAATGTTCTGCAGC	RP to generate <i>rho</i> -Y197A by SDM y

1395  
1396 <sup>a</sup> FP, forward primer  
1397 <sup>b</sup> RP, reverse primer  
1398 <sup>c</sup> SDM, site-directed mutagenesis  
1399

1400 **Supplementary Table 4: Plasmids and strains.**

Plasmid	Description	Source
pETM11- <i>rho</i>	pETM11 with wt <i>rho</i> ; no encoded tag; <i>Kan</i> <sup>R</sup>	9
pUC18- <i>λt<sub>R1</sub>-rut</i>	pUC19 with the <i>λt<sub>R1</sub> rut</i> region cloned at <i>Xba</i> I/ <i>Kpn</i> I sites; <i>Amp</i> <sup>R</sup>	27
pRS455	pET21b with wt <i>psu</i> cloned at <i>Nde</i> I/ <i>Xho</i> I sites; no encoded tag; <i>Amp</i> <sup>R</sup>	This study
pRS964	pET21b with <i>rho</i> -P167L cloned at <i>Nde</i> I/ <i>Xho</i> I sites; no encoded tag; <i>Amp</i> <sup>R</sup>	33
pRS100	pET21b with wt <i>rho</i> cloned at <i>Nde</i> I/ <i>Xho</i> I sites; no encoded tag; <i>Amp</i> <sup>R</sup>	38
pRS259	pNL150 with <i>psu</i> -F169V; <i>Cam</i> <sup>R</sup>	37
pRS458	pET28b with wt <i>psu</i> cloned at <i>Nde</i> I/ <i>Xho</i> I sites; encoded N-terminal His-tag; <i>Kan</i> <sup>R</sup>	38
pRS555	pNL150 with <i>psu</i> -E56K; <i>Cam</i> <sup>R</sup>	38
pRS624	pET28b with <i>psu</i> -E56K cloned at <i>Nde</i> I/ <i>Xho</i> I sites; encoded N-terminal His-tag; <i>Kan</i> <sup>R</sup>	42
pRS649	pCL1920 with wt <i>rho</i> with its promoter cloned at <i>Hind</i> III/ <i>Sac</i> I sites; <i>Spec</i> <sup>R</sup> , <i>Strep</i> <sup>R</sup>	93
pRS1117	pNL150 with wt <i>psu</i>	37
pRS1199	RS649 with <i>rho</i> -R146E; <i>Spec</i> <sup>R</sup> , <i>Strep</i> <sup>R</sup>	33
pRS1219	pET21b with <i>rho</i> -R146E cloned at <i>Nde</i> I/ <i>Xho</i> I site; no encoded tag; <i>Amp</i> <sup>R</sup>	33
pRS2158	pNL150 with <i>psu</i> -D49R; <i>Cam</i> <sup>R</sup>	This study
pRS2159	pNL150 with <i>psu</i> -K180A; <i>Cam</i> <sup>R</sup>	This study
pRS2160	pNL150 with <i>psu</i> -R43A; <i>Cam</i> <sup>R</sup>	This study
pRS2194	RS649 with <i>rho</i> -Y197A; <i>Spec</i> <sup>R</sup> , <i>Strep</i> <sup>R</sup>	This study
pRS2236	pET21b with Y197A <i>rho</i> cloned at <i>Nde</i> I/ <i>Xho</i> I site, Non-His-tag; <i>Amp</i> <sup>R</sup>	This study
pRS2237	pET28b with R43A <i>psu</i> cloned at <i>Nde</i> I/ <i>Xho</i> I site, His-tag at N-terminal; <i>Kan</i> <sup>R</sup>	This study
pRS2238	pET28b with D49R <i>psu</i> cloned at <i>Nde</i> I/ <i>Xho</i> I site, His-tag at N-terminal; <i>Kan</i> <sup>R</sup>	This study
pRS2239	pET28b with K180A <i>psu</i> cloned at <i>Nde</i> I/ <i>Xho</i> I site, His-tag at N-terminal; <i>Kan</i> <sup>R</sup>	This study
Strain	Genotype	Source
RS1263	<i>E. coli</i> MG1655; K-12	Lab stock
RS1309	<i>E. coli</i> MG1655Δ <i>rho</i> Δ <i>rac</i> , with shelter plasmid pHYD1201 ( <i>Amp</i> <sup>R</sup> )	107



RS2047	MG1655 $\Delta rac\Delta lac$ , $\lambda$ RS45 lysogen carrying $P_{lac}$ - $\lambda t_{R1}$ - $lacZYA$	75
--------	---	----

1401