Article

Action of a minimal contractile bactericidal nanomachine

Check for updates

Peng Ge[1,2,8], Dean Scholl[3,8], Nikolai S. Prokhorov[4], Jaycob Avaylon[2,5], Mikhail M. Shneider[6], Christopher Browning[7], Sergey A. Buth[4], Michel Plattner[4], Urmi Chakraborty[3], Ke Ding[1,2], Petr G. Leiman[4], Jeff F. Miller[1,2 ✉] & Z. Hong Zhou[1,2 ✉]

R-type bacteriocins are minimal contractile nanomachines that hold promise as precision antibiotics[1–4]. Each bactericidal complex uses a collar to bridge a hollow tube with a contractile sheath loaded in a metastable state by a baseplate scaffold[1,2]. Fine-tuning of such nucleic acid-free protein machines for precision medicine calls for an atomic description of the entire complex and contraction mechanism, which is not available from baseplate structures of the (DNA-containing) T4 bacteriophage[5]. Here we report the atomic model of the complete R2 pyocin in its pre-contraction and post-contraction states, each containing 384 subunits of 11 unique atomic models of 10 gene products. Comparison of these structures suggests the following sequence of events during pyocin contraction: tail fibres trigger lateral dissociation of baseplate triplexes; the dissociation then initiates a cascade of events leading to sheath contraction; and this contraction converts chemical energy into mechanical force to drive the iron-tipped tube across the bacterial cell surface, killing the bacterium.

Contractile nanotube-based machines are widespread in the bacterial domain, functioning to penetrate cell membranes to deliver payloads of proteins or DNA, or to create channels through which ions translocate[6]. Contractile type VI secretion systems (T6SSs) inject proteins into eukaryotic or bacterial cells to facilitate pathogenesis or to kill competing organisms, respectively[7–13]. Phage tail-like bacteriocins, exemplified by the R-type pyocins produced by *Pseudomonas aeruginosa*, employ the same contractility to kill competing bacteria by dissipating their membrane potential[1,2,14]. Myovirus bacteriophages, such as P2 and T4, use a similar contractile mechanism to translocate DNA into bacterial cells[5,9,15–22]. These contractile nanomachines employ a spring-loaded sheath–tube assembly to penetrate target cell surfaces[11]. This process is accompanied by massive structural transformations involving contraction of the sheath triggered by the baseplate, which has been visualized by both single-particle cryo-electron microscopy (cryo-EM)[1,7,10,23,24] and cellular cryo-electron tomography[25]. Although energy storage is mostly similar in phage, pyocin and the T6SS, the triggering mechanism may be different, owing to the presence of additional cell membranes in the case of the T6SS. By contrast, some other delivery systems, such as the T3SS and T4SS[26,27], carry no stored energy for penetration or baseplate-like structure for triggering.

An engineered T4 tube–baseplate complex reported previously[5] shows the static structures of baseplate proteins but lacks the sheath to inform how it receives the contraction signal. More recently, cryo-EM structures of insecticidal contractile toxin-delivery systems from *Photorhabdus asymbiotica* (*Photorhabdus* virulence cassette[23]) and *Serratia entomophila* (antifeeding prophage[24]) showcase the widespread existence of such phage tail-like contractile systems in nature. The relative simplicity and ease of engineering[3,4] of R-type pyocins make them ideal model systems for studying contractile structures. Our previous helical reconstructions of the pyocin R2 sheath and tube in pre-contraction and post-contraction states[1] revealed how energy for contraction is stored and released by shape and charge complementarity. However, the lack of atomic detail on the baseplate and collar precluded understanding of the molecular trigger that initiates sheath contraction, or how the resulting structure is stabilized to facilitate killing.

Here we report the atomic models of the pyocin collar and baseplate derived from single-particle cryo-EM and X-ray crystallography in their pre-contraction and post-contraction states. By comparing these structures, we derived and tested a model for the action of a bactericidal nanomachine that couples specific recognition of target cells with deployment of a generic mechanism of killing, providing insights that are crucial for exploiting these structures as precision antimicrobials[2,28,29].

Overall structures of pyocin R2

We imaged pyocin R2 by cryo-EM (Fig. 1a, b). Under the conditions used for sample isolation, both pre-contracted and post-contracted particles are present in the same preparation[1]. Each pyocin R2 complex consists of three structurally and functionally distinct components: the collar, the trunk and the baseplate with tail fibres. In the post-contraction state, the sheath layer of the trunk contracts by 70%, leaving the central tube exposed and readily visible in the cryo-EM images (Fig. 1a, b). We determined the cryo-EM structures of the pre-contracted pyocin R2 collar, trunk and baseplate regions

[1]Department of Microbiology, Immunology and Molecular Genetics, University of California, Los Angeles (UCLA), Los Angeles, CA, USA. [2]The California NanoSystems Institute (CNSI), University of California, Los Angeles (UCLA), Los Angeles, CA, USA. [3]Pylum Biosciences, South San Francisco, CA, USA. [4]University of Texas Medical Branch, Department of Biochemistry and Molecular Biology, Sealy Center for Structural Biology and Molecular Biophysics, Galveston, TX, USA. [5]Department of Chemistry and Biochemistry, University of California, Los Angeles (UCLA), Los Angeles, CA, USA. [6]Shemyakin-Ovchinnikov Institute of Bioorganic Chemistry, Laboratory of Molecular Bioengineering, Moscow, Russia. [7]Vertex Pharmaceuticals (Europe) Ltd, Abingdon, UK. [8]These authors contributed equally: Peng Ge, Dean Scholl. ✉e-mail: jfmiller@UCLA.edu; Hong.Zhou@UCLA.edu

Fig. 1 | Cryo-EM and overall structure of pyocin in pre-contracted and post-contracted states. a, Shaded surface representation of the cryo-EM reconstructions, coloured according to cylindrical radii as shown in the colour bar. **b**, A representative cryo-EM micrograph. Scale bar, 300 Å. **c**, Regions of the cryo-EM density map (mesh) superimposed with atomic models (sticks) demonstrating the agreement between the observed and modelled amino acid side chains. **d**, Atomic models for pyocin in both the pre-contracted (right) and the post-contracted (left) states. **e**, **f**, Ribbon diagrams of individual subunits of pyocin in the pre-contracted state (**e**) shown along with their corresponding gene loci (**f**). See the 3D rendition in Supplementary Videos 1–3.

separately and made a montage model by computationally stitching the three parts together (Fig. 1a, b and Supplementary Video 1). Substructures were determined at resolutions of 3.8 Å for the collar, 2.9 Å for the trunk and 3.5 Å for the baseplate (Extended Data Figs. 1, 2 and Extended Data Table 1). Although the average resolution of the baseplate reaches 3.5 Å (Fig. 1c and Supplementary Video 2), the resolution of the peripheral regions is lower, possibly owing to blurring from Brownian motion of the connected tail fibres (these regions are modelled according to both their cryo-EM density and our crystal structure of an engineered protein that represents them (Protein Data Bank (PDB) ID: 5CES) (Extended Data Table 2)). Using the same strategy, we determined the structure of the post-contracted pyocin at intermediate resolutions (Fig. 1a, right, and Supplementary Video 3). By matching amino acid side chains that are visible in our cryo-EM structures, we assigned a single gene product, PA0615, to the collar; two gene products, PA0622 and PA0623, to the trunk (PA0622 sheath and PA0623 tube); and seven gene products, PA0616–PA0619 and PA0626–PA0628, to the various components of the baseplate. We then built an atomic model of the
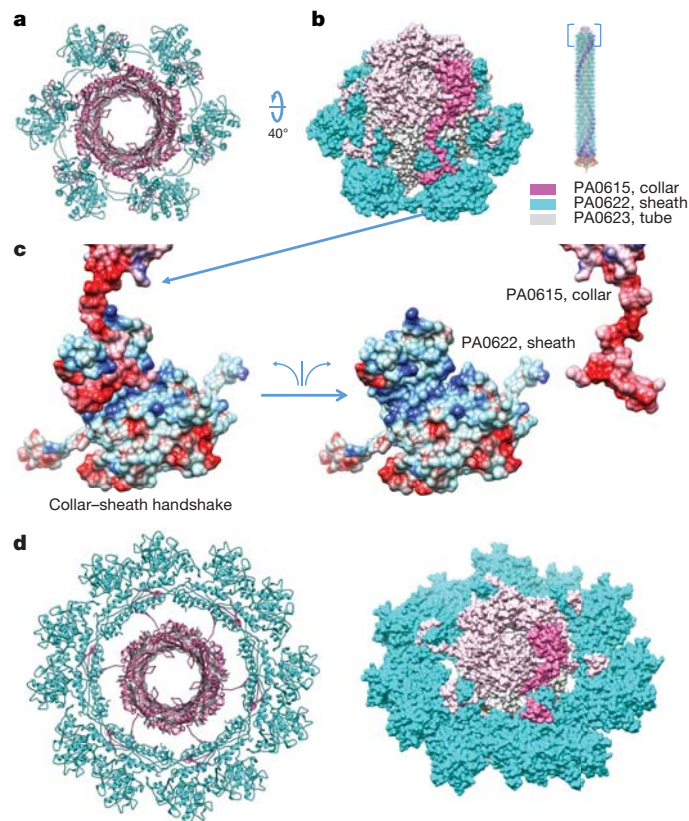
Fig. 2 | Architecture of the collar. a, Top view ribbon diagram of the collar (pink), the outer sheath (cyan) and the inner tube (grey). **b**, Space filling model of the collar–sheath–tube region. **c**, Electrostatic surface model of the collar–sheath handshake with views of the interacting charged surfaces (positive (blue), negative (red) and neutral (same as **b**)). **d**, Ribbon (left) and space filling (right) diagrams of the post-contraction collar–sheath–tube region similar to **a** and **b**, respectively.

complete pyocin, excluding the tail fibre and tape measure proteins (PA0620 and PA0625, respectively) (Fig. 1d–f, Extended Data Table 3 and Supplementary Video 4).

Collar tethers tube to the contracted sheath

The collar is a hexamer formed by the gene product PA0615 (Fig. 2a, b). Each collar monomer has a simple structure with two domains, one globular domain and one β-hairpin domain, joined by an extended loop (Figs. 1e and 2b). The structure of the globular domain is similar to that of the inner tube protein. The collar extends the tube and tethers it to the sheath, thus preventing the tube from dissociating from the sheath after contraction as demonstrated in the *Photorhabdus* virulence cassette[23]. By capping the top of the tube and augmenting the handshake β-sheet of the sheath subunit below with a β-hairpin (similar to the sheath–sheath handshake[1,7,10]), it provides mechanical stability to the junction after the downward pull of the sheath against the tube (Fig. 2a, b).

Specifically, the handshake domain of each sheath subunit is augmented by two β-strands that are donated by subunits from a disc that is closer to the collar. Thus, each subunit of the last disc has a β-sheet with two vacant β-strands. The β-hairpin domain of the collar protein completes this β-sheet. In addition, the hydrogen bond interactions in this augmentation are reinforced by charge–charge interactions: the loop and the β-hairpin domain of the collar are mainly negative, binding to a groove in the sheath that is mainly positive (Fig. 2c), unlike the very

Article

hydrophobic nature of the sheath–sheath interaction. In this configuration, the collar hexamer joins the tube and the sheath at the very top.

We also determined the structure of the post-contraction collar region at an average resolution of 3.5 Å (Extended Data Fig. 1). The most dramatic structural change is that the diameter of the outer sheath increases from 180 Å to 240 Å, which results in the dislocation of the sheath from the tube (Fig. 2d). This dislocation appears to reduce the structural rigidity of the collar protein linker and to increase the local mobility of the tube, as the averaged resolution of the tube portion in the reconstruction is only 7 Å and the linker has become invisible. Nonetheless, the rest of the structure, including the sheath protein and the hairpin domain of the collar, is of sufficient resolution for building atomic models. The hairpin domain of the collar undergoes a slight conformational change during contraction, but the globular domain remains structurally unchanged.

The trunk and the sheath initiator

Since we determined the structures of the bottommost part of the trunk and the baseplate without imposing helical symmetry, we can now resolve the interface regions of both components. Six subunits of the PA0623 tube protein form a hexameric ring with a 24-stranded β-barrel, 28 of which stack into the central tube of each pyocin. Handshake interactions with long N-terminal and C-terminal extension arms interweave PA0622 sheath protein subunits, maintaining the connectivity of the sheath during contraction (Extended Data Fig. 3). Thus, our new non-helical structure confirms our previous observations based on helical reconstructions of the trunk portion of pyocin from images recorded on film[1].

Sheath discs that approach the baseplate break from the shared helical symmetry of the trunk with a slightly but gradually increasing rotation per disc, reaching an extra 4.4° in the bottommost disc (Extended Data Fig. 4). The handshake β-sheet is the same across the entire sheath, except for the bottommost disc where the sheath subunit arms are incorporated into the PA0617 sheath initiator protein (Extended Data Fig. 3). The arm of the sheath protein augments a β-sheet of the sheath initiator protein in the same manner as in the rest of the sheath. Thus, the β-sheet augmentation mechanism allows the sheath to be linked via multiple polypeptide chains to both the collar and the baseplate (see the three different types of handshaking augmentation in Extended Data Fig. 3).

Baseplate: triggering triplex and ripcord

Our reconstruction shows that the pyocin baseplate is composed of eight different protein subunits: ripcord (PA0626), triplex (two copies of PA0618 (Tri1a and Tri1b) and one copy of PA0619 (Tri2)), sheath initiator (PA0617), glue (PA0627), hub (PA0628) and spike (PA016) (Fig. 1e and Extended Data Fig. 5). PA0626 forms the centrepiece of the baseplate (Extended Data Fig. 5a): the central spike complex is 'below', the tube–sheath and the trunk that they form is 'above', and the rest of the baseplate (triplex, sheath initiator and glue) surrounds it. The central spike complex is composed of the trimeric or pseudo-hexameric PA0628 hub (Extended Data Fig. 5b) and the trimeric PA0616 central spike inserted into it (Extended Data Fig. 5c). Our crystal structure of the PA0616 spike, a homologue of the P2 phage gpV[30], shows that it carries a ferric ion at the very tip. The ferric ion is coordinated by a triplet of H×H double histidine motifs, a structure that stabilizes the tip for membrane penetration (Extended Data Fig. 5c).

Each PA0626 monomer has a C-terminal β-sheet domain (amino acids 136–286) and an N-terminal helix-rich domain (amino acids 1–112) joined by an extended linker (amino acids 113–135), which is an unusual domain organization reminiscent of a ripcord (Extended Data Fig. 5a). We reason that this setup may comprise a form of an activation energy barrier for triggering and may also be crucial for assembly:
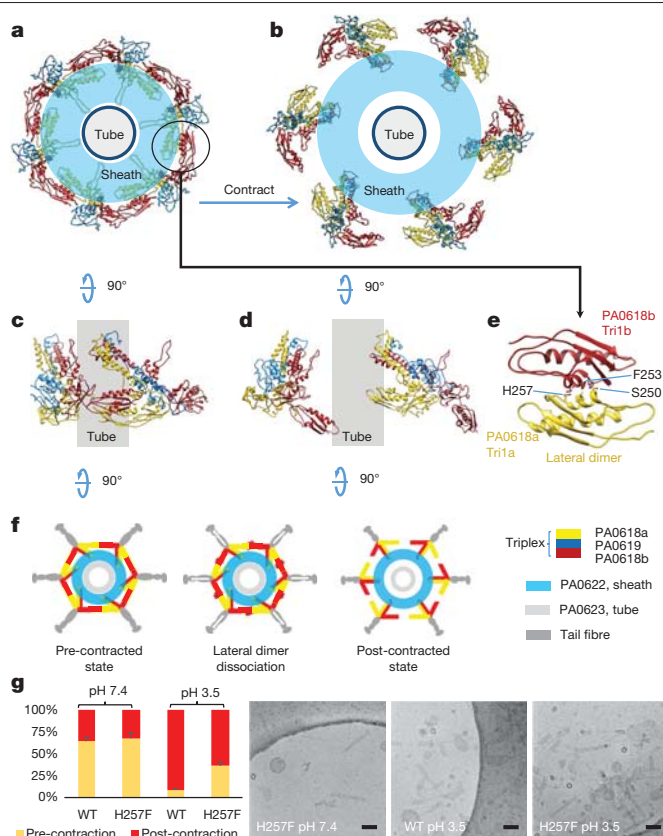
Fig. 3 | Triplex expansion and lateral dimer dissociation. a–e, Ribbon diagrams of triplexes forming an iris ring in the pre-contracted state (**a**), an expanded iris in the post-contracted state (**b**), side view of two adjacent triplexes in the pre-contracted state (**c**) and in the post-contracted state (**d**), and a lateral dimer of PA0618 (**e**). **f**, Schematic of the iris ring expansion as a result of the tail fibre actuation. **g**, The PA0618 H257F mutant. Left, percentages of pre-contraction pyocins in the purified WT and the H257F mutant under cryo-EM at neutral and acidic pH (counts: pre-contraction/all: pH 7.4 WT: 185/289, H257F: 118/175; pH 3.5 WT: 46/530, H257F: 64/178). Error bars represent standard deviations. Right, the representative cryo-EM image for each relevant condition. Scale bars, 300 Å.

the C-terminal domains of six PA0626 subunits form a hexameric disc, which resembles that of the tube proteins and extends the tube (Extended Data Fig. 5a). This disc is integral to the tube–PA0626–hub–spike assembly that penetrates the target as it connects the tube to the hub. The N-terminal domains of PA0626, a four-helical bundle structure, are localized at a higher cylindrical radius inside the baseplate (Extended Data Fig. 5d). Indeed, when we introduce a tobacco etch virus (TEV) protease cleavage site in the linker of the ripcord, the mutant pyocin (626TEV) shows wild-type (WT)-like assembly without TEV protease co-expressed and no assembly with TEV protease co-expressed (Extended Data Fig. 6). We also found that this mutant (626TEV) and another deletion mutant (626ΔWL) have lower activation energy for triggering compared to the WT (Extended Data Fig. 6).

Two copies of PA0618 and one copy of PA0619 proteins form a (PA0618)$_2$–PA0619 heterotrimeric triplex by attaching two copies of PA0618 (Tri1a/Tri1b) to either side of PA0619 (Tri2), resulting in two distinct conformations of PA0618 (Extended Data Fig. 7), which accept the PA0626 N-terminal four-helical bundle domain (Extended Data Fig. 8). At its top, this triplex also binds to the sheath initiator PA0617 protein. A small protein with a LysM fold—PA0627—binds to a side of the triplex and glues PA0617, PA0618b and PA0619 together (Extended Data Fig. 5e). The (PA0618)$_2$–PA0619 triplexes are joined into an iris-like

structure by lateral dimers of the C-terminal domains of PA0618 (Fig. 3a and Extended Data Fig. 7e, f). PA0619 attaches to the tail fibre, allowing it to receive a triggering signal from it (Extended Data Fig. 5f).

By comparing the pre-contraction and post-contraction baseplate models, we found that the iris ring[5], joined by lateral dimerization of PA0618, breaks apart after contraction of the pyocin (Figs. 1d and 3). In this way, the baseplate in the post-contraction state splits into a hexagram shape. The entire complex of the sheath initiator, glue and triplex travels in a rigid-body movement as each sheath subunit does, widening the baseplate to 320 Å (Fig. 3b). The tail fibre still binds to the triplex upon contraction, as its densities are still visibly connected to the triplex, although at low resolution.

Compared to the bacteriophage T4 baseplate[5], pyocin baseplate proteins are minimalistic and bear interesting differences. PA0618 and PA0619 both lack sizeable insertion domains that are required for building the considerably larger T4 baseplate. Instead, these insertions are replaced by loops (Extended Data Fig. 9), although their function as struts that connect the hub and the baseplate is preserved. PA0617 and PA0627 are also minimalistic compared to T4 gp25 and gp53, respectively, simply retaining the core motifs[5] (Extended Data Fig. 9). Furthermore, the pyocin sheath protein lacks domains 3 and 4 of the T4 sheath protein, which add to the energy release of contraction and interact with long tail fibres, respectively[16,17].

Bactericidal action and application

On the basis of our atomic models of R2 pyocin in pre-contracted and post-contracted states, we constructed a minimal contractile machine that includes just 12 stacks of the sheath and rendered a morph movie between the two states to illustrate a possible pathway of action of such a contractile nanomachine (Supplementary Video 5). Notably, because the sheath changes helicity during contraction, the tube undergoes a rotational movement during its power stroke, which may facilitate the spike's penetration of the target cell surface.

The biological function of R-type pyocins is to kill competing bacteria and they do so with an extraordinary efficiency that approaches single-shot killing[4,31]. To achieve this, it is necessary to actuate only in the right place and time. Upon recognition by tail fibre receptor-binding proteins of specific ligands on a target cell, the increased free energy would cause displacement of interacting surfaces within lateral dimers of PA0618, and we suggest that shearing forces that are transduced through tail fibres trigger the baseplate to initiate contraction (Fig. 4). Shearing force encountered by particles during purification may bear similar characteristics. This force drives the baseplate to transit to a larger diameter, breaking the lateral dimers of PA0618 Tri1a/Tri1b (Fig. 3a–f). This process may be reversible to a threshold level, ensuring that a sufficient number of tail fibre-binding interactions occurs to disrupt enough lateral dimers to break the triplex iris ring and begin the irreversible process of sheath contraction[32]. This would provide a 'checksum' mechanism for the baseplate and avoid premature triggering. It may also be used to ensure that the pyocin is positioned upright to prevent sideward, non-productive firing, as the arrangement of triggering tail fibres may also play a role in setting off the baseplate[32] (Fig. 4a). As the iris breaks, the triplexes may pull the four-helical bundles of attached ripcord proteins, lowering the final barrier to contraction (Fig. 4).

Indeed, changing the lateral dimer interface between PA0618 Tri1a and Tri1b subunits alters the firing characteristics of pyocin. Owing to the presence of a histidine residue in the interface, pyocins are sensitive to an acidic environment and their contraction is triggered by pH 3.4 (Fig. 3g), at which the histidine is protonated to disturb the dimer interface (Fig. 3f). We engineered a mutant, H257F, in which the histidine is replaced by a phenylalanine. We found that the mutant is more tolerant to acid than the WT pyocin, whereas cryo-EM images show that purified WT and H257F both contain a similar percentage
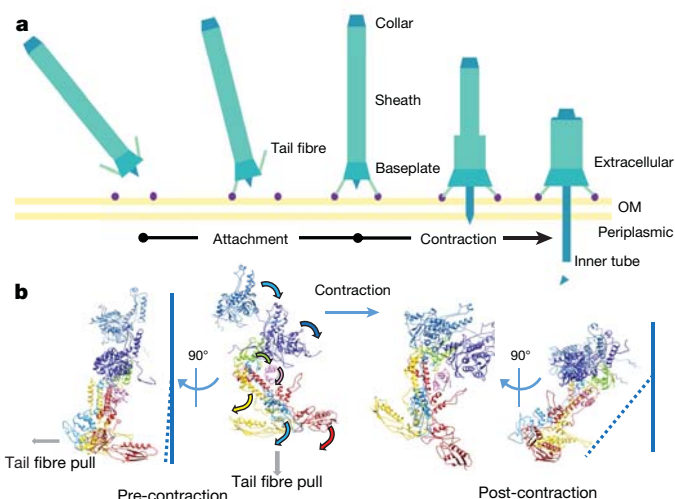
Fig. 4 | Baseplate transition from the pre-contracted to the post-contracted state. a, Illustration of a pyocin landing on a bacterial cell and firing. Release of the spike and hub following injection is postulated on the basis of the lack of these structures on contracted particles that we observed in vitro. OM, outer membrane. **b**, Ribbon diagram of the conserved baseplate components and sheath proteins in their pre-contracted and post-contracted states. Ripcord is believed to travel with the inner tube during the power stroke and therefore is not a conserved component of the baseplate after contraction. Arrows denote potential movements for subunits in the same colours, respectively.

of pre-contraction pyocins at neutral pH (64% and 67%, respectively); however, at pH 3.4, cryo-EM images show that far more WT than H257F pyocins have been triggered into the post-contraction state (9% and 36% of the remaining pre-contraction pyocins, respectively) (Fig. 3g). Two other PA0618 mutations at the lateral dimer interface, S250A and A254C (Fig. 3f), resulted in either defective assembly (S250A) or premature firing (A254C) of the particles.

Among the known structures of similar contractile systems, the sheath and sheath initiator are both conserved, which indicates that key aspects of the contraction mechanism are conserved[6,8]. Whether these similarities also extend to T6SS baseplate proteins is not known in the absence of atomic structures. Thus far, however, the four-helical bundle motif of the ripcord protein described here seems to be unique to pyocin R2 and related phages (for example, PS17 phage) based on genetic information. Indeed, a search in other contractile tail-like systems did not yield any four-helical bundle orthologues, although all contain orthologues for the triplex proteins. This unique triggering system puts pyocin R2 in a special position among contractile structures, perhaps due to its minimal nature that precludes sophisticated and sizeable insertion domains like those of T4.

R-type pyocins and related R-type bacteriocins are being developed as a new class of antimicrobials[2–4,28,29]. A unique feature of these structures is that highly specific target recognition conferred by receptor-binding proteins is directly coupled to the mechanism of action. This exquisite specificity allows selective killing of pathogens without the unintended consequences of off-target effects such as dysbiosis[33], and without selecting for transmissible antibiotic resistance in off-target bacterial species or strains. Engineering receptor-binding proteins to alter R-type pyocin-binding specificity has already been demonstrated[3,4,34]. We now have the information required to fine-tune the triggering mechanism through structure-guided alterations at key interaction points between baseplate components. For applications that require precise ablation of pathobionts from complex bacterial ecosystems, a less sensitive trigger would minimize off-target effects and would only set off the killing mechanism when tightly bound to

Article

the correct bacterial cell. Conversely, in full-blown infections such as septicaemia, in which a single pathogen has grown to high density, a more sensitive trigger would allow for more efficient killing on collision with the target. The adaptability of contractile injection systems and the modularity of receptor-binding proteins, both honed over eons of evolution, provide an opportunity to engineer precision antibiotics for human and animal health.

Online content

Any methods, additional references, Nature Research reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at https://doi.org/10.1038/s41586-020-2186-z.

1. Ge, P. et al. Atomic structures of a bactericidal contractile nanotube in its pre- and postcontraction states. *Nat. Struct. Mol. Biol.* **22**, 377–382 (2015).
2. Scholl, D. Phage tail-like bacteriocins. *Annu. Rev. Virol.* **4**, 453–467 (2017).
3. Scholl, D. et al. An engineered R-type pyocin is a highly specific and sensitive bactericidal agent for the food-borne pathogen *Escherichia coli* O157:H7. *Antimicrob. Agents Chemother.* **53**, 3074–3080 (2009).
4. Williams, S. R., Gebhart, D., Martin, D. W. & Scholl, D. Retargeting R-type pyocins to generate novel bactericidal protein complexes. *Appl. Environ. Microbiol.* **74**, 3868–3876 (2008).
5. Taylor, N. M. et al. Structure of the T4 baseplate and its function in triggering sheath contraction. *Nature* **533**, 346–352 (2016).
6. Leiman, P. G. et al. Type VI secretion apparatus and phage tail-associated protein complexes share a common evolutionary origin. *Proc. Natl Acad. Sci. USA* **106**, 4154–4159 (2009).
7. Kudryashev, M. et al. Structure of the type VI secretion system contractile sheath. *Cell* **160**, 952–962 (2015).
8. Basler, M., Pilhofer, M., Henderson, G. P., Jensen, G. J. & Mekalanos, J. J. Type VI secretion requires a dynamic contractile phage tail-like structure. *Nature* **483**, 182–186 (2012).
9. Leiman, P. G. & Shneider, M. M. Contractile tail machines of bacteriophages. *Adv. Exp. Med. Biol.* **726**, 93–114 (2012).
10. Clemens, D. L., Ge, P., Lee, B. Y., Horwitz, M. A. & Zhou, Z. H. Atomic structure of T6SS reveals interlaced array essential to function. *Cell* **160**, 940–951 (2015).
11. Böck, D. et al. In situ architecture, function, and evolution of a contractile injection system. *Science* **357**, 713–717 (2017).
12. Ho, B. T., Dong, T. G. & Mekalanos, J. J. A view to a kill: the bacterial type VI secretion system. *Cell Host Microbe* **15**, 9–21 (2014).
13. Mougous, J. D. et al. A virulence locus of *Pseudomonas aeruginosa* encodes a protein secretion apparatus. *Science* **312**, 1526–1530 (2006).
14. Stover, C. K. et al. Complete genome sequence of *Pseudomonas aeruginosa* PAO1, an opportunistic pathogen. *Nature* **406**, 959–964 (2000).
15. Aksyuk, A. A. et al. The tail sheath structure of bacteriophage T4: a molecular machine for infecting bacteria. *EMBO J.* **28**, 821–829 (2009).
16. Kostyuchenko, V. A. et al. The tail structure of bacteriophage T4 and its mechanism of contraction. *Nat. Struct. Mol. Biol.* **12**, 810–813 (2005).
17. Leiman, P. G., Chipman, P. R., Kostyuchenko, V. A., Mesyanzhinov, V. V. & Rossmann, M. G. Three-dimensional rearrangement of proteins in the tail of bacteriophage T4 on infection of its host. *Cell* **118**, 419–429 (2004).
18. Hu, B., Margolin, W., Molineux, I. J. & Liu, J. Structural remodeling of bacteriophage T4 and host membranes during infection initiation. *Proc. Natl Acad. Sci. USA* **112**, E4919–E4928 (2015).
19. Hu, B., Margolin, W., Molineux, I. J. & Liu, J. The bacteriophage t7 virion undergoes extensive structural remodeling during infection. *Science* **339**, 576–579 (2013).
20. Hatfull, G. F. Bacteriophage genomics. *Curr. Opin. Microbiol.* **11**, 447–453 (2008).
21. Hendrix, R. W., Hatfull, G. F. & Smith, M. C. Bacteriophages with tails: chasing their origins and evolution. *Res. Microbiol.* **154**, 253–257 (2003).
22. Chen, Z. et al. Cryo-EM structure of the bacteriophage T4 isometric head at 3.3-Å resolution and its relevance to the assembly of icosahedral viruses. *Proc. Natl Acad. Sci. USA* **114**, E8184–E8193 (2017).
23. Jiang, F. et al. Cryo-EM structure and assembly of an extracellular contractile injection system. *Cell* **177**, 370–383.e15 (2019).
24. Desfosses, A. et al. Atomic structures of an entire contractile injection system in both the extended and contracted states. *Nat. Microbiol.* **4**, 1885–1894 (2019).
25. Chang, Y. W., Rettberg, L. A., Ortega, D. R. & Jensen, G. J. In vivo structures of an intact type VI secretion system revealed by electron cryotomography. *EMBO Rep.* **18**, 1090–1099 (2017).
26. Hu, B. et al. Visualization of the type III secretion sorting platform of *Shigella flexneri*. *Proc. Natl Acad. Sci. USA* **112**, 1047–1052 (2015).
27. Low, H. H. et al. Structure of a type IV secretion system. *Nature* **508**, 550–553 (2014).
28. Kirk, J. A. et al. New class of precision antimicrobials redefines role of *Clostridium difficile* S-layer in virulence and viability. *Sci. Transl. Med.* **9**, eaah6813 (2017).
29. Ritchie, J. M. et al. An *Escherichia coli* O157-specific engineered pyocin prevents and ameliorates infection by *E. coli* O157:H7 in an animal model of diarrheal disease. *Antimicrob. Agents Chemother.* **55**, 5469–5474 (2011).
30. Browning, C., Shneider, M. M., Bowman, V. D., Schwarzer, D. & Leiman, P. G. Phage pierces the host cell membrane with the iron-loaded spike. *Structure* **20**, 326–339 (2012).
31. Kageyama, M., Ikeda, K. & Egami, F. Studies of a pyocin. Iii. Biological properties of the pyocin. *J. Biochem.* **55**, 59–64 (1964).
32. Crawford, J. T. & Goldberg, E. B. The function of tail fibers in triggering baseplate expansion of bacteriophage T4. *J. Mol. Biol.* **139**, 679–690 (1980).
33. Gebhart, D. et al. A modified R-type bacteriocin specifically targeting *Clostridium difficile* prevents colonization of mice without affecting gut microbiota diversity. *mBio* **6**, e02368-14 (2015).
34. Scholl, D., Gebhart, D., Williams, S. R., Bates, A. & Mandrell, R. Genome sequence of *E. coli* O104:H4 leads to rapid development of a targeted antimicrobial agent against this emerging pathogen. *PLoS One* **7**, e33637 (2012).

Methods

Purification of pyocins for cryo-EM

The pyocin sample was prepared from the *P. aeruginosa* strain PAO1 as previously described[1,4]. Briefly, a crude pyocin sample was first prepared from PAO1 and harvested by high-salt precipitation and differential centrifugation. This crude sample was further purified using a 10–50% sucrose (w/v) gradient at 77,000g for 1.5 h at 4 °C. After centrifugation, one band was visible at about the 25% position and was extracted gently by fractionation with a 100-µL pipette from the top of the centrifuge tube along its side. The extracted sample was then diluted to a final volume of 4 mL with Tris buffer (10 mM Tris and 130 mM NaCl, pH 7.4). The diluted sample was concentrated using a 100-kDa Amicon molecular filter to about 50 µL. This dilution concentration step was repeated three more times in the same filter as a means of dialysis, ending up with a final sample volume of 50 µL for cryo-EM imaging.

Cloning and purification of proteins for crystallography

Bioinformatic analysis with the help of HHpred[35] was used to identify the *P. aeruginosa* PAO1 genes encoding the central spike protein (PA0616) and Tri1 (PA0618) in the R2 pyocin cluster of *P. aeruginosa* PAO1. The proteins were amplified with the primers given in Supplementary Table 1. PA0616 was cloned into the pEEv3 vector (a pET23d derivative with a TEV protease cleavage site downstream from the N-terminal His-tag) using the BamHI and HindIII restriction sites. PA0618 was cloned into the standard pET23 vector using the NdeI and XhoI restriction sites to give rise to a protein with a non-cleavable C-terminal His-tag.

Both proteins were expressed in *Escherichia coli* B834 (DE3) cells grown in 2×TY medium at 37 °C and aerated at 200 rpm to an optical density at 600 nm (OD$_{600}$) of 0.6 and induced with the addition of 1 mM IPTG. The temperature of the culture was lowered to 18 °C and protein expression was carried out overnight. The cells were pelleted by centrifugation for 10 min at 8,000g at 4 °C and then lysed by sonication. The soluble fraction was separated by centrifugation for 15 min at 25,000g at 4 °C and loaded on the IMAC resin (GE Health Science). The IMAC affinity chromatography was performed in 10 mM Tris-Cl pH 8.0 with a 0–300 mM imidazole gradient. Fractions containing the protein were dialysed into 10 mM Tris-Cl pH 8.0 overnight. The TEV protease in a 1:10 mass-to-mass ratio (TEV protease:target protein) was added to the PA0616 sample, and proteolysis of the N-terminal His-tag was performed simultaneously with a dialysis into 10 mM Tris-Cl pH 8.0 overnight at room temperature. The dialysis of PA0618 was performed overnight at 4 °C. The subsequent steps of protein purification were the same for both proteins. The samples were loaded onto a MonoQ 10/100 GL column equilibrated with 10 mM Tris-Cl pH 8.0 and eluted with a 0–1 M NaCl gradient. Fractions containing the proteins of interest were then subject to size-exclusion chromatography on a Superdex 200 10/300 GL column in the buffer containing 10 mM Tris-Cl pH 8.0 and 150 mM NaCl. Fractions containing the proteins were pooled together and concentrated while the buffer was changed to 20 mM Tris-Cl pH 8.0.

Crystallization and structure determination of PA0616 and PA616d

For crystallization, PA0616 is concentrated to 18 mg/mL and subjected to a sparse matrix random screen using crystallization kits produced by Jena BioSciences. Best crystals of PA0616 were obtained in 60% PEG-400, 100 mM Na$_2$SO$_4$, 100 mM Bis-Tris buffer at pH 8.5 in hanging drop. Several data sets were collected at the beamlines PX-I and PX-III of the Swiss Light Source (SLS) and the best data set was used in structure refinement (Extended Data Table 2). The crystals did not require a special cryoprotectant and could be flash frozen directly from the drop.

The structure of PA0616 was solved by molecular replacement using the P2 phage gpV[30] as a search model (the two proteins display 31% sequence identity). The PA0616 unit cell contained six trimers in the asymmetric unit. The gpV search model contained the OB-fold and the first three strands of the β-helix. The initial solution was found with the help of Molrep[36]. It was rigid body refined with Refmac5[37]. Ten cycles of density modification by solvent flattening and non-crystallographic averaging with Parrot[38] improved the density to a point where the OB-fold and the first three β-strands could be interpreted in terms of PA0616 amino acids. A new model of 1 chain was then superimposed onto the other 17 chains comprising the asymmetric unit. A new round of rigid body refinement was performed with Refmac5[37]. The new density was subjected to an additional 15 cycles of solvent flattening and non-crystallographic averaging with Parrot. The new map was interpretable throughout and the model could be built for all but the last five amino acids, comprising the Fe-binding site. This region of the model was disordered in all other data sets of PA0616 (including different space groups).

To reveal the structure of the Fe-binding site, we designed a shortened mutant of PA0616 (called PA0616d) that comprised 90 C-terminal residues (amino acids 96–185) of the full-length protein. PA0616d was PCR-amplified using primers given in Supplementary Table 1 and cloned into pEEv3 so that it carried a cleavable His-tag at its N terminus. Expression and purification were performed in a way similar to the full-length protein. PA0616d was then concentrated to 15 mg/mL and subjected to a sparse matrix random screen using crystallization kits produced by Jena BioSciences. Best crystals of PA0616d were obtained in 34% PEG-400, 200 mM Na$_2$SO$_4$, 100 mM Na-acetate buffer at pH 5.0 in hanging drop. The crystals did not require a special cryoprotectant and could be flash frozen directly from the drop. The crystallographic data were collected at the beamline PX-I of the SLS. The structure was solved by molecular replacement using the middle part of the overlapping fragment (residues 120–179) of the full-length PA0616 with the help of the Phaser program[39]. The N-terminal part of the PA0616d β-helix was substantially different to that of the full length. Instead of forming a compact structure, it was splayed out and opened like a flower. This structural difference made structure solution difficult, but thanks to the availability of very high-resolution data (Extended Data Table 2), after several rounds of model building and refinement coupled to density modification with Parrot, an interpretable density could be obtained for the rest of the protein (including the Fe-binding site).

The structures of PA0616 and PA0616d were refined with Coot[40], Refmac5[37] and Phenix[41]. The final models of PA0616 and PA0616d were deposited to the PDB under the accession numbers 4S37 and 4S36, respectively.

Production of the PA0618 fragment suitable for crystallographic analysis and the solution of the structure

Full-length PA0618 failed to crystallize. Thus, we subjected it to limited proteolysis by trypsin. Full-length PA0618 was digested with trypsin (trypsin:PA0618 ratio of 1:500) in high-salt buffer (20 mM Tris pH 8.0, 400 mM NaCl and 4 mM CaCl$_2$) for 75 min at room temperature. The reaction was quenched by the addition of PMSF to a final concentration of 2 mM. This procedure resulted in two stable fragments. We loaded the mixture onto a His-Trap column (GE Healthcare) and found that one of the fragments bound to the resin, suggesting that this fragment retained the His-tag and therefore constituted a C-terminal part of the protein (PA0618C). The fragment eluted from the column with a 10 mM Tris pH 8.0 buffer containing imidazole at 250 mM. The eluted protein was further purified by size-exclusion chromatography (Superdex 75 HiLoad 16/60), dialysed into 20 mM Tris pH 8.0 and 5 mM DTT overnight and then purified by anion-exchange chromatography (Mono Q 10/100). Fractions comprising the elution peak were pooled and purified by size-exclusion chromatography (Superdex 75 HiLoad 16/60, 10 mM Tris pH 8.0 and 150 mM NaCl) again. The purified PA0618 C terminus was concentrated to 40 mg/mL and screened for crystallization conditions using Jena BioSciences crystallization kits. Crystallization drops of the best crystallization condition carried a crust, which prevented optimal crystal growth. Lowering of the protein

Article

concentration prevented the formation of the crust, but it also stopped crystal formation. The best crystals were obtained by streak seeding of drops containing PA0618C at 17 mg/mL and 17% PEG-5000 MME, 150 mM $NH_4CH_3CO_2$ and 100 mM MES pH 6.5.

Crystallographic data were collected at SLS beamlines PX-I and PX-III using well solution supplemented with 25% ethylene glycol as a cryo-protectant. The initial phases were obtained by a single-wavelength anomalous diffraction technique (SAD) using a $HoCl_2$ heavy-atom derivative. The crystals were soaked in the well solution containing $HoCl_2$ at 5 mM for 24 h. The model comprising residues 202–295 plus a His-tag was initially built by buccaneer[42] and completed manually. The structure of a better native data set was refined using Coot[40], Refmac5[37] and Phenix[41]. The structure is deposited to the PDB under the accession number 5CES. The resultant model of the C terminus fragment was then used in combination with the cryo-EM-derived structure to build the final model of PA0618.

Cryo-EM

An aliquot of 2.5 μL of the above purified pyocin sample was loaded onto a 'baked'[43] Quantifoil 1.2/1.3 m, 200 mesh grid, blotted for 4 s at force 1, then flash frozen with a Vitrobot Mark IV (FEI). Cryo-EM data were collected as movies in an FEI Titan Krios microscope (operated at 300 kV) equipped with a Gatan imaging filter (GIF) (the slit was not inserted) and a K2 Summit direct electron camera in counting mode using Leginon software package[44] for automation.

The target defocus value is set to 2.0 μm under focus. Each movie contains 50 frames with 5 frames per second with a total accumulated dosage of 60 electrons per $Å^2$. The dose rate is measured at 6 electrons per $Å^2$ per second in the Digital Micrograph software package.

Frames within each movie were aligned to correct for drift as previously described[45], except that an iterative alignment scheme, as previously described elsewhere[46], was used in addition to the original software. We output three averages of selected frames: 1st to 50th frames for particle selection, 3rd to 20th frames for refinement, and 3rd to 13th frames for final reconstruction.

The contrast transfer function (CTF) parameters of these movies were determined from the averages with the 1st to 50th frames by CTFFIND3[47]. The range for acceptable defocus values was set to be between 1 and 3 μm, whereas outliers were rejected.

Image processing and 3D reconstruction

As shown in Fig. 1b, particles at both pre-contraction and post-contraction states are present in the cryo-EM images and can be readily distinguished by eye. The two ends of each pyocin were manually selected as individual particles with EMAN[48] boxer and were kept together before 3D classifications (see below). The box size for these particles is 320 and 420 pixels for the pre-contraction and post-contraction states, respectively; their total numbers of particles are 43,934 and 36,116, respectively. The trunk portion of each pre-contracted pyocin was boxed with EMAN[48] helixboxer with a box width of 384 and was segmented according to a 10% overlapping scheme, and a total of 15,684 segments selected. Because both ends of the post-contraction state contained sufficient structural components from the trunk, we did not pursue a helical reconstruction for the trunk in the post-contraction state. Throughout these particle boxing processes, we only boxed particles that were not overlapping, not broken and not ice-contaminated.

Further image processing steps were performed with Relion v1.2[49]. The boxed particles were first subjected to 2D classification to eliminate poor particles and then to a 3D classification to separate the collar and baseplate for each state (6 classes for the pre-contraction state and 12 classes for the post-contraction state). After separation of the collar and the baseplate, for each end of a pyocin particle and for each state, a further 3D classification was done to again eliminate poor particles (five classes for the pre-contraction baseplate and eight classes for the rest).

The 3D models for the classes after this classification appeared to be mutually shifted along the particle axis. Therefore, the 3D model for all good classes is aligned to the one in the middle. The resulting shifts were projected to the 2D space of the original particles and were applied to the particle centres. The particles of the good classes were then extracted again using the translated coordinates.

The finally selected and re-extracted particles for each of the four structures were subjected to an auto-refinement in Relion v1.2. A sixfold rotational symmetry was applied to the reconstructions. The number of finally included particles is as follows: the pre-contraction collar: 4,109; the pre-contraction baseplate: 26,104; the post-contraction collar: 9,934; and the post-contraction baseplate: 15,582. The overall averaged resolutions of these final structures are estimated by ResMap[50] to be: the pre-contraction collar: 3.9 Å; the pre-contraction baseplate: 3.4 Å; and the post-contraction collar and baseplate: 3.5 Å (Extended Data Fig. 1).

Protein subunit identification and atomic modelling

A purified pyocin preparation was run 5 mm into a 10% SDS gel and stained with Coomassie blue. Each band containing a pyocin protein was excised, digested in-gel with trypsin and subjected to liquid chromatography–tandem mass spectrometry for sequencing analysis (conducted at the University of California Davis Proteomics Core). Proteins were then identified by comparing the mass spectrometry sequence fragments to the *P. aeruginosa* PAO1 sequence (Fig. 1e and Extended Data Table 3).

Atomic models were built ab initio with Coot[40]. We had to assign each of the above-identified protein candidates to a specific region in our cryo-EM density map. At 3.4 Å resolution, this is possible because our map has sufficient resolving power for chain tracing and identification of side chains (Fig. 1c and Extended Data Fig. 2). For each peptide chain, we meticulously compared its secondary structure to the secondary structure predictions of all candidate proteins and found the best match. Once the match was found, the sequence of the candidate protein was compared to the amino acid side chain features in the density map to register the sequence. Then the α-carbon positions were manually traced with the above identified sequence in mind for each of the 11 unique atomic models. The α-carbon trace was converted to a poly-alanine strand, and finally mutated to the correct amino acid sequence. Each of the side chains were manually inspected and fitted into the density along with additional attention to secondary structure conformations. Protein subunits were assembled into an asymmetric unit. Only protein subunits in one asymmetric unit were built. Symmetry-related copies were generated using the 'sym' command in UCSF Chimera[51] with C6 rotational symmetry.

Real-space model refinement

We carried out model refinement with the phenix.real_space_refine command of the Phenix package[41] using default settings in three steps. First, we refined each monomer model with the corresponding density cryo-EM map individually. Second, we combined all of the monomers in an asymmetric unit and refined it with the whole cryo-EM map to separate clashing atoms between adjacent monomers within the asymmetric unit. Third, the refined model of the asymmetric unit was used to generate the full model of either the baseplate or the collar using the 'sym' command in UCSF Chimera with C6 rotational symmetry. This full model was refined globally with the NCS restraints enabled to separate clashing atoms among asymmetric units. At each of the refinement steps, we manually inspected the models to assess the quality of the refinement, made manual adjustments and repeated the refinement steps until a final structure was reached. The refined models of the individual monomers and the full baseplate and collar were validated exhaustively with EMRinger score[52], Ramachandran plot, C-beta, map CC and MolProbity[53] and the results are tabulated in Extended Data Fig. 1.

Structure-based mutagenesis for PA0618

Mutations in the pyocin gene cluster were made in the *P. aeruginosa* strain PAO1 by two-step allelic exchange following the method previously described[4]. Briefly, ~1,000-bp regions of DNA containing mutations were amplified from PAO1 genomic DNA using overlapping PCR and cloned into pEX18Gm (gift from H. Schweizer) that was digested with KpnI and EcoRI. The constructs were transformed by electroporation into PAO1, which was plated on 50 µg/mL gentamicin to select for single crossover integrants. Gentamicin-resistant isolates were then picked, grown for 3 h in LB and plated on LB-sucrose plates to select for second crossover events. Colonies were picked and screened for gentamicin sensitivity. These were then sequence verified for correct mutations. The primers used for these constructs are shown in Supplementary Table 2.

Crude WT and H257F-mutant pyocin were prepared as described above. Both were diluted 1:10 in either Tris buffer (10 mM Tris and 130 mM NaCl, pH 7.4) or phosphate-citrate buffer (28 mM Na_2HPO_4 and 36 mM citric acid, pH 3.4) to adjust the pH. Cryo-EM imaging was done for each of the four resulting samples as described above. The numbers of pre-contraction and post-contraction pyocins were counted through visual inspection of about 200 images for each condition and their percentages were tallied in the bar graph in Fig. 3f.

PA0626-mutant construction

The PA0626 mutants 626TEV and 626ΔWL were made using the modified allelic exchange approach[54]. In brief, the whole WT R2 pyocin gene cluster resided on a pETcoco-1-based plasmid, pSW192. In each case, two overlapping ~1-kb-long fragments carrying the mutation were amplified using pSW192 as a template by the primers listed in Supplementary Table 2.

The donor vectors for the exchange pAK6 and pAK20 were assembled of pairs of fragments by NEBuilder reaction (New England Biolabs) on the backbone of pWM91 plasmid[55] in which the ampicillin resistance gene was replaced with the kanamycin resistance gene. The recipient plasmid pSW192 was maintained in the RecA+ *E. coli* 4 s strain[56]. The donor vectors were transformed into the MFDpyr *E. coli* strain[57], and conjugation between the donor and the acceptor strains was performed on LB agar plates overnight at 37 °C. Selection for the recombination products pSW192:pAK6 and pSW192:pAK20 was done on LB agar plates supplemented with kanamycin at 50 µg/mL. Counterselection for excision products, pSW192 and a mutation-carrying plasmid, was done on agar plates with 1% tryptone, 0.5% yeast extract and 5% sucrose (MilliporeSigma) overnight at room temperature. Colony screening was done by PCR. Plasmid identities were confirmed by sequencing.

626TEV in vivo digestion

For co-expression of pyocins with TEV protease, pS626TEV and pBAD24-based plasmid pB^hTEV coding TEV protease under the control of the arabinose-inducible promoter were co-transformed into the *E. coli* BL21ΔAraΔFhuA strain alongside with all necessary control combinations (WT pyocin coding plasmid pSW192 and pB^hTEV, pSW192 and pBAD24, and pS626TEV and pBAD24). Clones were selected on agar plates for ampicillin and chloramphenicol resistance, grown in liquid LB medium supplemented with ampicillin at 100 µg/mL and chloramphenicol with 10 µg/mL at 37 °C, induced with arabinose (0.01% and 0.03%) and incubated overnight at 30 °C. Fresh lysates were cleared from debris for 5 min at 15,000g in a microcentrifuge, tested for killing activity on *P. aeruginosa* 13 s strain lawns[58] in a spot assay and visualized on a JEM-2100 electron microscope (JEOL) after staining with 2% uranyl acetate.

Recombinant pyocin production

One litre of *E. coli* BL21ΔAraΔFhuA freshly transformed with each plasmid carrying WT or mutant pyocin gene clusters was grown in Lennox LB medium (Invitrogen) supplemented with chloramphenicol (11 µg/mL) in a 4 L Erlenmeyer flask at 37 °C and 240 rpm to an OD_{600} of 1.0. To induce pyocin expression, 0.5 mL of 20% arabinose were added, the temperature was decreased down to 30 °C, and cells were incubated overnight. To remove debris and residual bacteria, the lysate was centrifuged at 15,000g for 30 min in a F9-6x-1000 rotor (Thermo Fisher Scientific). The lysate was supplemented with 4 mg of DNase I and 4 mg of RNase A (MilliporeSigma). 30 g of solid NaCl and 100 g of PEG-8000 (Fisher Scientific) were added and dissolved and the lysate was incubated at 4 °C overnight to ensure the full precipitation of pyocins. Pyocins were pelleted at 15,000g for 15 min in a F9-6x-1000 rotor and pellets were resuspended in 10 mL of SM buffer (8 mM $MgCl_2$, 100 mM NaCl and 50 mM Tris-HCl pH 7.5) with DNase I and RNase A at 1 µg/mL each. The sample was extracted with 10 mL of chloroform, centrifuged in 50-mL falcon tubes at 15,000g for 15 min. The pyocin-containing aqueous phase was collected and pyocins were pelleted at 100,000g for 1 h in a Type 70 Ti rotor (Beckman Coulter). The pellet was dissolved in 0.5 mL of SM buffer with shaking at 100 rpm overnight at 4 °C. The sample was cleared from non-dissolvable material via centrifugation at 15,000g for 5 min in a microcentrifuge at room temperature and loaded on a step gradient of 10%, 20%, 30%, 40% and 50% sucrose (0.9 mL each) and centrifuged at 100,000g for 1 h in a SW 55 Ti rotor (Beckman Coulter). The upper pyocin-containing band was isolated and dialysed against two changes of 0.1× SM buffer. The concentration of the dialysed sample was determined on the basis of the adsorption at 280 nm and brought to 0.1 mg/mL. To get a control sample of fully contracted pyocins, 20 µL of 3 M glycine-HCl pH 2.5 were added to 4 mL of normalized pyocin sample, and contracted and non-contracted samples were dialysed against 10 mM NaCl and 10 mM phosphate buffer pH 7.0. The whole procedure was repeated several times to get at least three independently prepared samples for each circular dichroism experiment performed.

Circular dichroism

Pyocin contraction rates were measured on a JASCO J-815 CD spectrometer at 203 nm over a 67–74 °C range with an increment of 1 °C. The pyocin concentration for contracted and non-contracted samples was maintained equal to 0.1 mg/mL in 10 mM NaCl and 10 mM phosphate buffer pH 7.0. Each measurement took 25 min. The change of ellipticity for non-contracted pyocins was temperature dependent; however, it was not temperature dependent for contracted control samples. For data analysis, we averaged curves for each contracted sample and a resulting control curve subtracted from each contraction rate measurement of a cognate non-contracted sample. The first 2 min of every measurement were trimmed as they correspond to a sample being heated up to the desirable temperature and do not show pyocin contraction. Exponents were fitted into subtracted and trimmed data and the rate constants were determined with the help of MATLAB CFTool (MathWorks). Arrhenius modelling of activation energy was also done in MATLAB.

Reporting summary

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

Data availability

Cryo-EM maps and the associated atomic models have been deposited to the Electron Microscopy Data Bank (EMDB) and the PDB under the accession numbers EMD-20526/PDB: 6PYT (pre-contraction helical trunk), EMD-20643/PDB: 6U5B (pre-contraction baseplate), EMD-20646/PDB: 6U5H (pre-contraction hub in C3 symmetry), EMD-20644/PDB: 6U5F (pre-contraction collar), EMD-20647/PDB: 6U5J (post-contraction collar) and EMD-20648/PDB: 6U5K (post-contraction

Article

baseplate), respectively. X-ray crystal structures have been deposited to the PDB under the accession numbers 5CES (PA0618 C-terminal domain), 4S36 (PA0616 C-terminal domain) and 4S37 (full-length PA0616). All other data are available from the corresponding authors on reasonable request.

Code availability

The modified version of MotionCorr 1 is available on GitHub, licensed under GPLv3 (gepeng1983/motioncorr1exp). Relion 1.2 with helical reconstruction patch is available on GitHub, licensed under GPLv2 (gepeng1983/relion12exp). A later version of Relion (1.4) with the same patch is also available on GitHub (gepeng1983/relion14exp).

35. Söding, J., Biegert, A. & Lupas, A. N. The HHpred interactive server for protein homology detection and structure prediction. *Nucleic Acids Res.* **33**, W244–W248 (2005).
36. Vagin, A. & Teplyakov, A. Molecular replacement with MOLREP. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 22–25 (2010).
37. Murshudov, G. N. et al. REFMAC5 for the refinement of macromolecular crystal structures. *Acta Crystallogr. D Biol. Crystallogr.* **67**, 355–367 (2011).
38. Zhang, K. Y., Cowtan, K. & Main, P. Combining constraints for electron-density modification. *Methods Enzymol.* **277**, 53–64 (1997).
39. McCoy, A. J. et al. Phaser crystallographic software. *J. Appl. Crystallogr.* **40**, 658–674 (2007).
40. Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. Features and development of Coot. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 486–501 (2010).
41. Adams, P. D. et al. PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 213–221 (2010).
42. Cowtan, K. The Buccaneer software for automated model building. 1. Tracing protein chains. *Acta Crystallogr. D Biol. Crystallogr.* **62**, 1002–1011 (2006).
43. Miyazawa, A., Fujiyoshi, Y. & Unwin, N. Structure and gating mechanism of the acetylcholine receptor pore. *Nature* **423**, 949–955 (2003).
44. Suloway, C. et al. Automated molecular microscopy: the new Leginon system. *J. Struct. Biol.* **151**, 41–60 (2005).
45. Li, X. et al. Electron counting and beam-induced motion correction enable near-atomic-resolution single-particle cryo-EM. *Nat. Methods* **10**, 584–590 (2013).
46. Banerjee, S. et al. 2.3 Å resolution cryo-EM structure of human p97 and mechanism of allosteric inhibition. *Science* **351**, 871–875 (2016).
47. Mindell, J. A. & Grigorieff, N. Accurate determination of local defocus and specimen tilt in electron microscopy. *J. Struct. Biol.* **142**, 334–347 (2003).
48. Ludtke, S. J., Baldwin, P. R. & Chiu, W. EMAN: semiautomated software for high-resolution single-particle reconstructions. *J. Struct. Biol.* **128**, 82–97 (1999).
49. Scheres, S. H. W. RELION: implementation of a Bayesian approach to cryo-EM structure determination. *J. Struct. Biol.* **180**, 519–530 (2012).
50. Kucukelbir, A., Sigworth, F. J. & Tagare, H. D. Quantifying the local resolution of cryo-EM density maps. *Nat. Methods* **11**, 63–65 (2014).
51. Pettersen, E. F. et al. UCSF Chimera—a visualization system for exploratory research and analysis. *J. Comput. Chem.* **25**, 1605–1612 (2004).
52. Barad, B. A. et al. EMRinger: side chain-directed model and map validation for 3D cryo-electron microscopy. *Nat. Methods* **12**, 943–946 (2015).
53. Chen, V. B. et al. MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 12–21 (2010).
54. Blomfield, I. C., Vaughn, V., Rest, R. F. & Eisenstein, B. I. Allelic exchange in *Escherichia coli* using the *Bacillus subtilis* sacB gene and a temperature-sensitive pSC101 replicon. *Mol. Microbiol.* **5**, 1447–1457 (1991).
55. Metcalf, W. W. et al. Conditionally replicative and conjugative plasmids carrying lacZ alpha for cloning, mutagenesis, and allele replacement in bacteria. *Plasmid* **35**, 1–13 (1996).
56. Prokhorov, N. S. et al. Function of bacteriophage G7C esterase tailspike in host cell adsorption. *Mol. Microbiol.* **105**, 385–398 (2017).
57. Ferrières, L. et al. Silent mischief: bacteriophage Mu insertions contaminate products of *Escherichia coli* random mutagenesis performed using suicidal transposon delivery plasmids mobilized by broad-host-range RP4 conjugative machinery. *J. Bacteriol.* **192**, 6418–6427 (2010).
58. Scholl, D. & Martin, D. W. Jr. Antibacterial efficacy of R-type pyocins towards *Pseudomonas aeruginosa* in a murine peritonitis model. *Antimicrob. Agents Chemother.* **52**, 1647–1652 (2008).

Author contributions Z.H.Z., J.F.M., P.G. and P.G.L. conceived the project. D.S., J.A. and K.D. prepared pyocin R2 samples used for high-resolution cryo-EM. P.G. and J.A. recorded the cryo-EM data. P.G. and Z.H.Z. processed the cryo-EM data. P.G., J.A. and P.G.L. built the atomic models using cryo-EM data. P.G., P.G.L., J.A. and Z.H.Z. analysed and interpreted the models. D.S., N.S.P. and U.C. created the pyocin mutants and examined their assembly properties and phenotypes. N.S.P. designed the functional assays and circular dichroism experiments to measure the activation energy of sheath contraction. M.M.S. created the expression constructs used for crystallography. C.B. and P.G.L. determined the crystal structure of the PA0616 spike protein. S.A.B. and M.P. determined the crystal structure of the C-terminal domain of PA0618. P.G., Z.H.Z., J.A., P.G.L., D.S. and J.F.M. wrote the paper. All authors contributed to the editing of the manuscript.
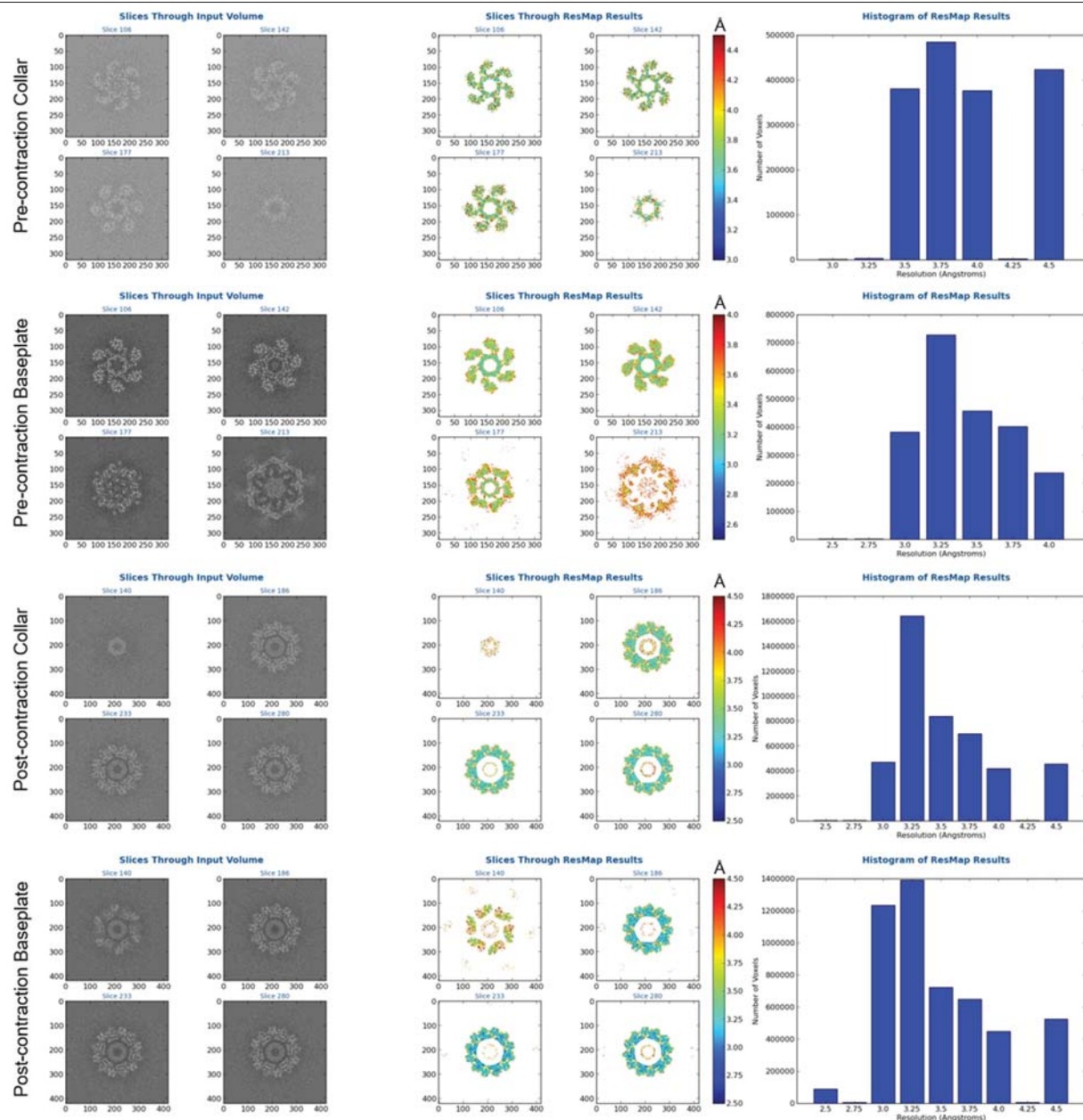
|	Protein	Ramachandran Outliers	Ramachandran Favored	Rotamer outliers	C-beta outliers	Bonds RMSD	Angles RMSD	Map CC	EMRinger Score
PA0615 / Collar		0%	90.9%	0.0%	0.0%	0.007	1.111	0.75	2.35
Baseplate Subunits	PA0617	0%	92.8%	0.0%	0.0%	0.006	0.925	0.82	2.62
	PA0618a	0%	91.7%	0.4%	0.0%	0.005	0.989	0.67	2.39
	PA0618b	0%	92.8%	0.4%	0.0%	0.005	0.920	0.73	2.01
	PA0619	0%	91.3%	0.0%	0.0%	0.005	1.040	0.71	2.54
	PA0626	0%	91.5%	0.0%	0.0%	0.006	0.965	0.82	3.26
	PA0627	0%	94.3%	0.0%	0.0%	0.005	0.957	0.81	4.59
Baseplate Average		0%	92.1%	0.2%	0.0%	0.005	0.967	0.77	2.35

Extended Data Fig. 1 | Resolution assessment and model validation for the cryo-EM structures. ResMap results for the collar and baseplate regions of the pyocin reconstructions in pre-contraction and post-contraction states. Listed in the table are model validation statistics for the collar, baseplate subunits and average. For the 'Slices Through Input Volume' and the 'Slices Through ResMap Results' panels, the units on the axes are indexes for pixels.
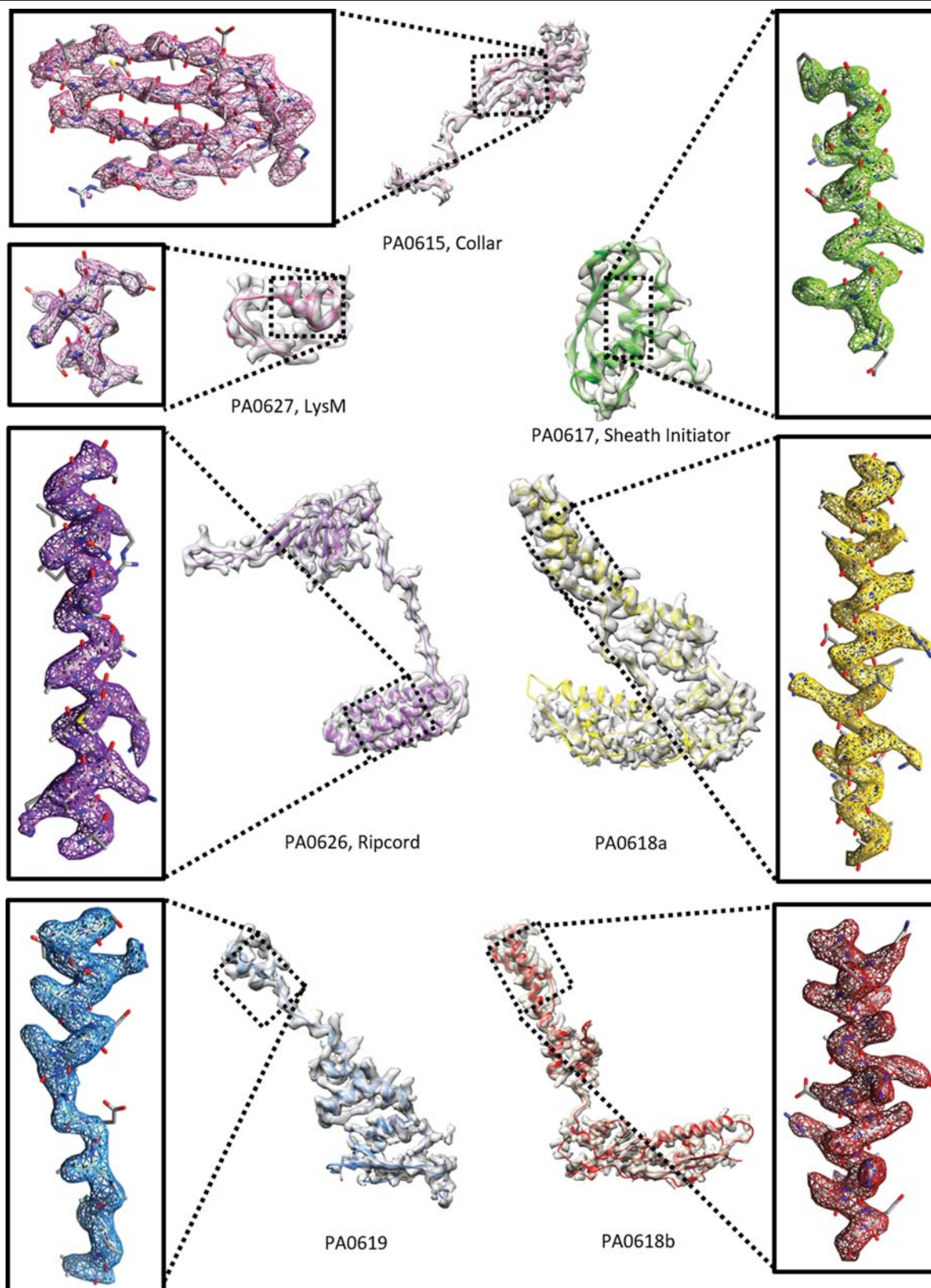
Extended Data Fig. 2 | Model assessment of pre-contraction pyocin subunits. For each of the collar and baseplate proteins, the cryo-EM density map is shown as semi-transparent grey superposed with its atomic model (ribbon). The close-up view of the box region shows the match of the density (wire frames) and the atomic model (sticks).
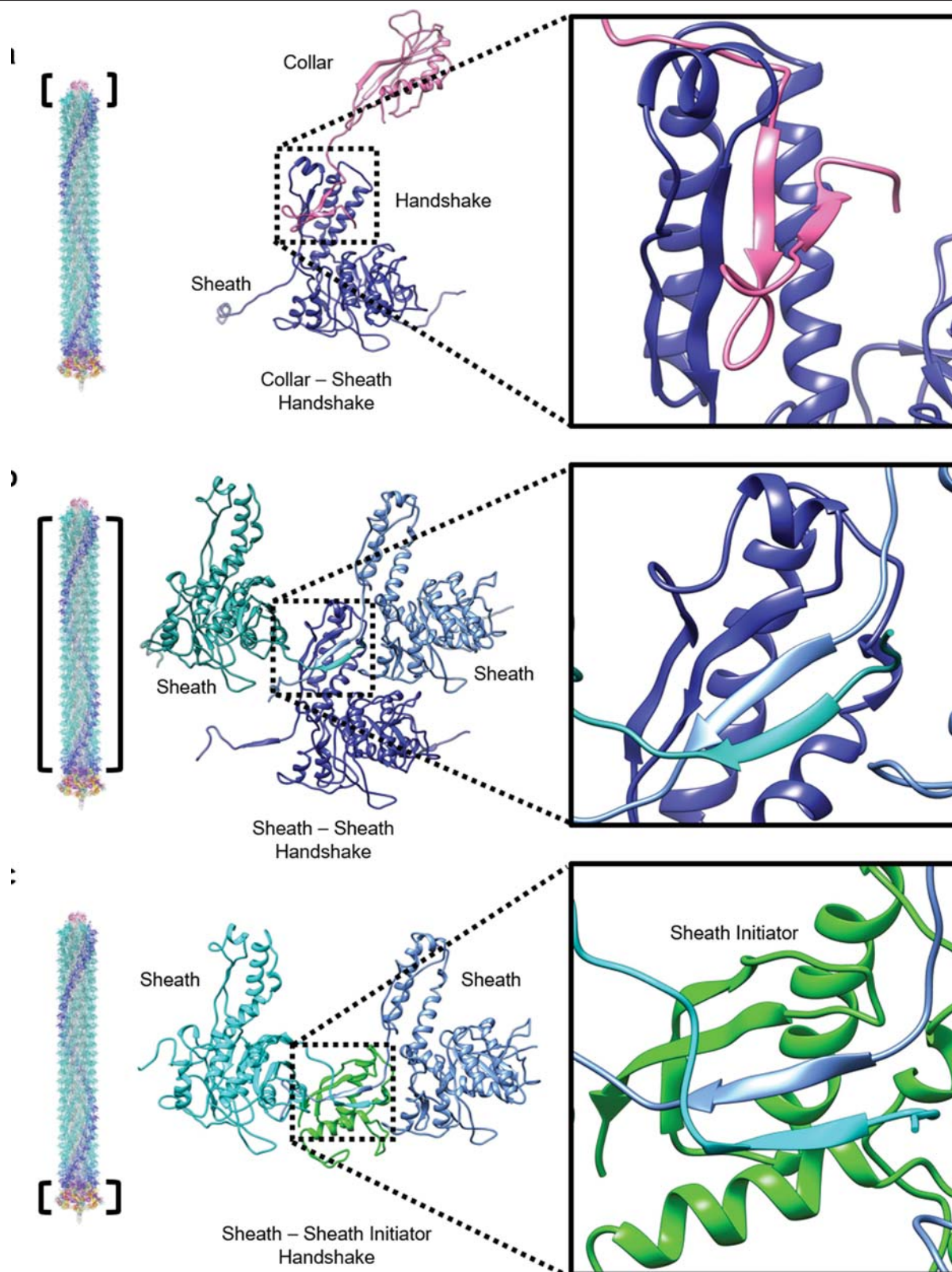
Extended Data Fig. 3 | Types of sheath handshakes in pyocin. Ribbon diagrams depicting the three types of handshake conformations in pyocin. **a**, Collar–sheath handshake. **b**, Sheath–sheath handshake. **c**, Sheath–sheath initiator handshake. All handshakes are composed of a four-stranded β-sheet.

a

b
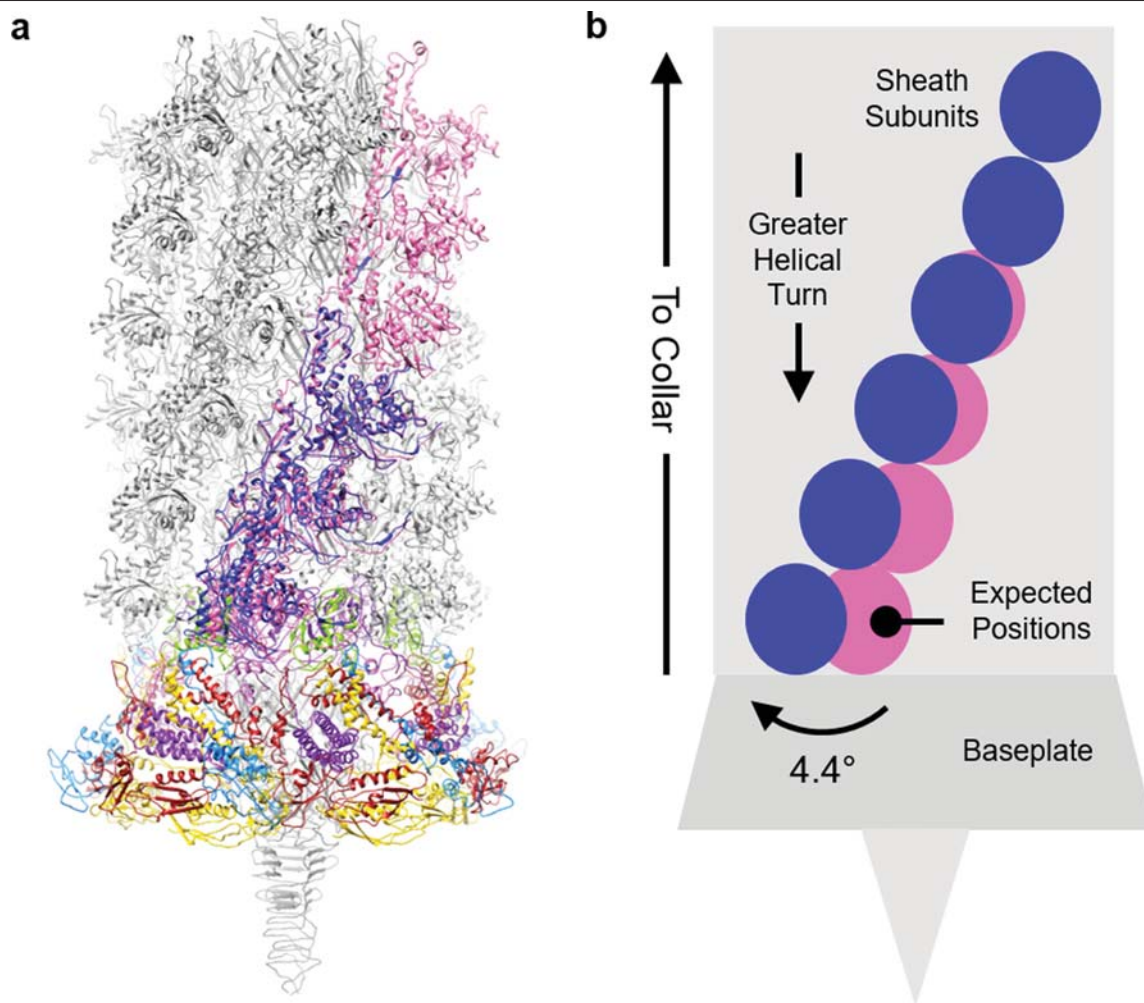
Extended Data Fig. 4 | Trunk transitioning into the baseplate. a, Ribbon diagram depicting the lower portion of pyocin. The pink ribbons depict the expected positions of the sheath subunits according to the helical symmetry of the trunk. **b**, Schematic diagram depicting changes in the quaternary structure of the sheath subunits approaching the baseplate. The pink circles depict the expected positions of the sheath subunits according to helical symmetry of the trunk. The blue circles depict the actual positions with greater sequential helical turn, 4.4° at the last disc of the sheath.
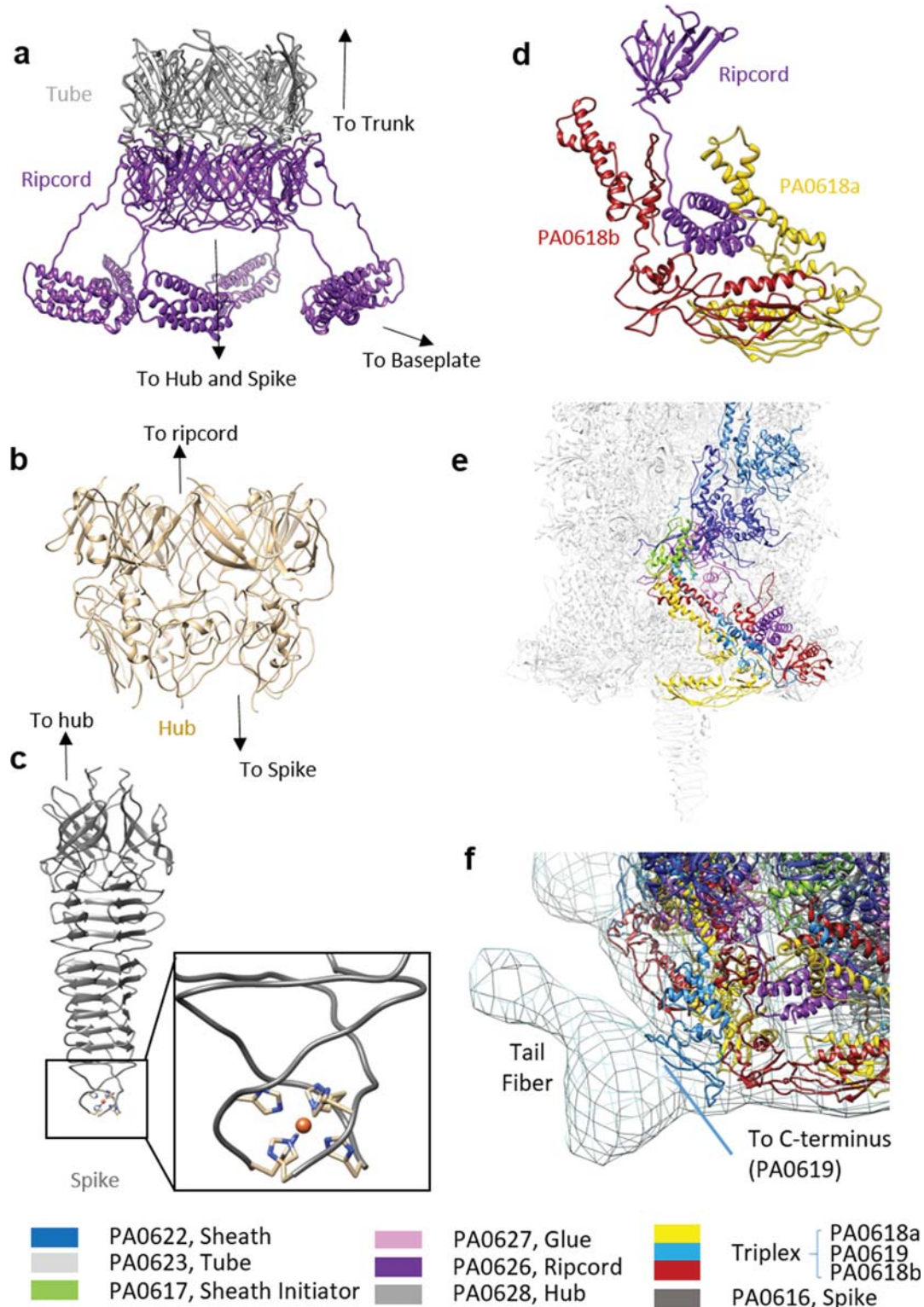
Extended Data Fig. 5 | Inspection of the baseplate. a, Ribbon diagram of the ripcord hexamer with the tube hexamer. **b, c**, Ribbon diagram of the hub (**b**) and the spike (**c**) (with the chelating site of its iron ion highlighted). **d**, Binding of the ripcord into triplexes. **e**, Ribbon diagram of the baseplate with one-sixth of its sixfold symmetric part highlighted in colours as in Fig. 1, showing the relative positions of each subunit. **f**, Baseplate ribbon model superimposed with a blurred cryo-EM density map of the proximal region of the tail fibre.
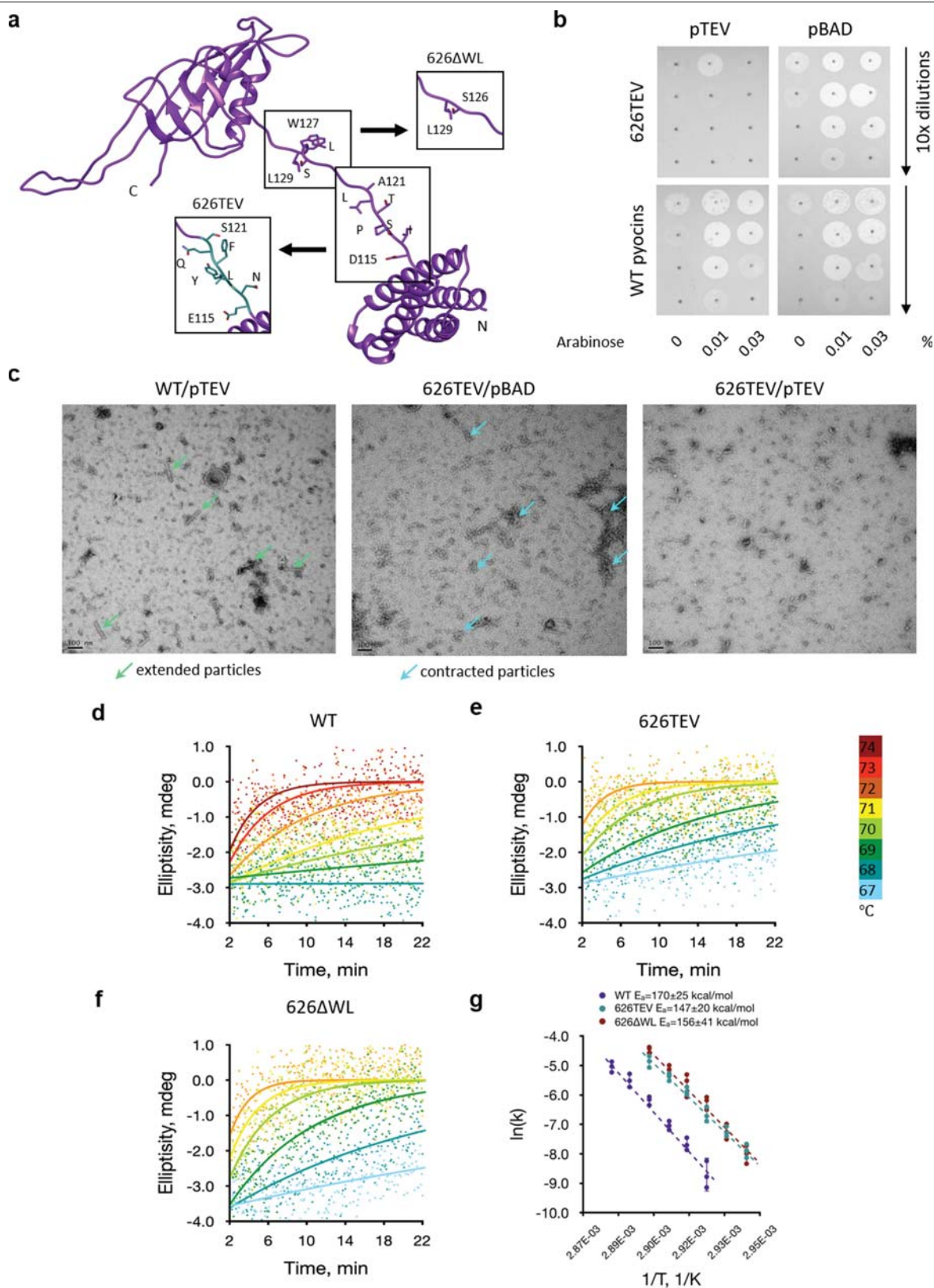
Extended Data Fig. 6 | See next page for caption.

Extended Data Fig. 6 | Functional and morphogenetic implications of ripcord mutagenesis. a, Overview of ripcord mutagenesis. **b**, Co-expression of the WT pyocin and mutant 626TEV with the TEV protease. Pyocin killing activity in the lysates was assessed with the help of a spot assay with the *P. aeruginosa* 13 s strain as prey. Both pyocin and protease expression levels are arabinose dependent, with the rate of protease production being proportional to arabinose concentration and pyocin expression reaching the maximum at the lowest concentrations of arabinose tested (0.01%). Each experiment was repeated biologically three times (also for **c**–**g**). **c**, Representative negative staining EM images of the crude lysates shown in panel **b** induced with 0.01%

arabinose. Despite showing killing activity in the lysates, no extended particles were found in the mutant 626TEV on EM grids. **d**–**f**, Temperature-dependent sheath contraction rates of the WT pyocins (**d**) and mutants (**e** and **f**) measured with the help of circular dichroism. **g**, The rate constants $k(T)$ of WT pyocins, 626ΔWL and 626TEV fitted to the Arrhenius model $k(T) = A \exp(-E_a/RT)$ where T is the absolute temperature, A is a temperature independent constant, E_a is the activation energy and R is the ideal gas constant. The dots on the graph are individual values for three biologically independent measurements of $\ln k(T)$, and the error bars show the 95% confidence interval calculated for them.
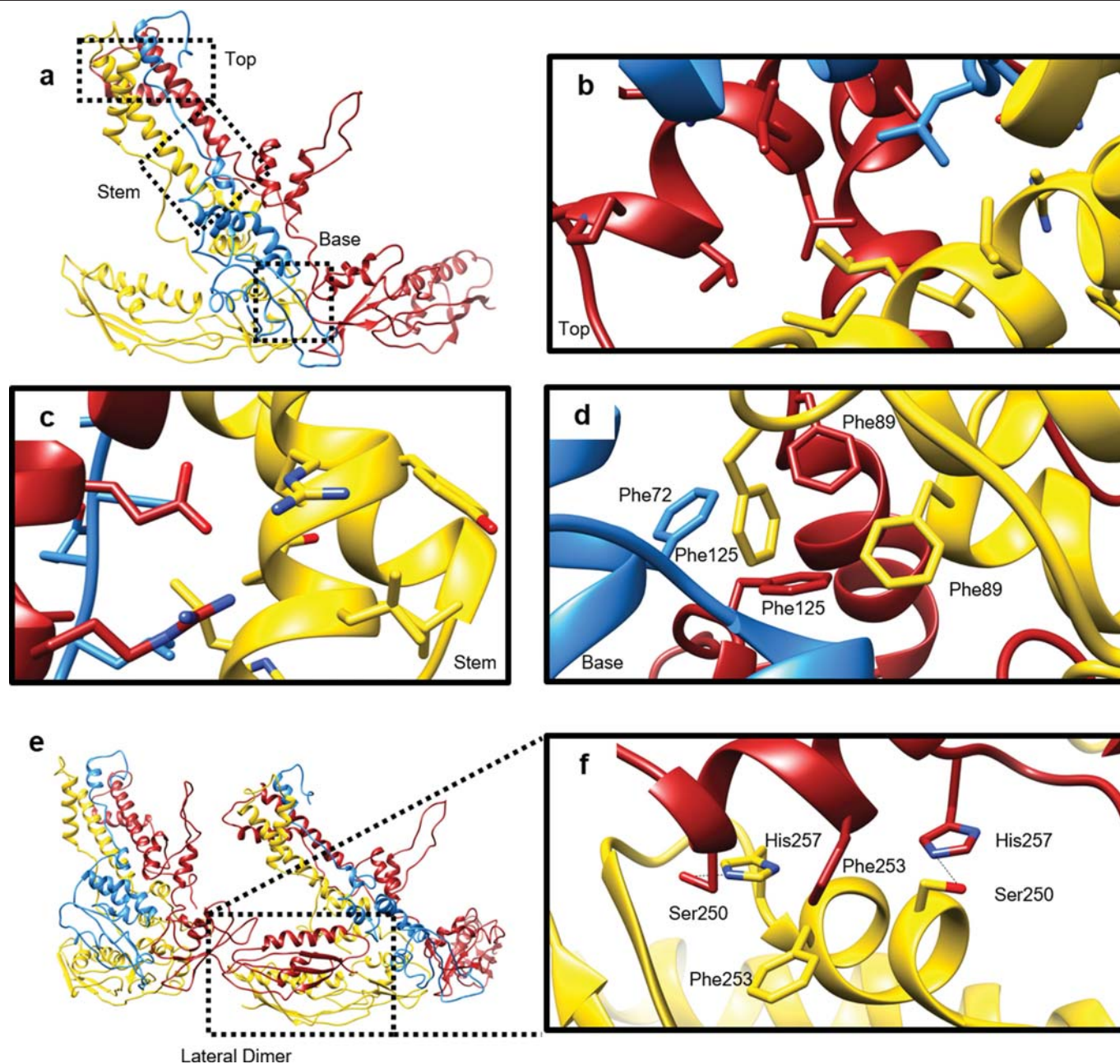
Extended Data Fig. 7 | Interactions important for triplex formation.
a, Ribbon diagram of the atomic model of the pyocin triplex. **b**–**d**, The ribbon model with the depicted side chains in the top (**b**), stem (**c**) and base (**d**) regions of the triplex. In panel **d**, the phenylalanine pi-stacking coordination between PA0618a (yellow, Phe89 and Phe125), PA0618b (red, Phe89 and Phe125) and PA0619 (blue, Phe72) is shown. **e**, Ribbon model diagram of the lateral dimer. **f**, Close-up of the ribbon model diagram of the lateral dimer, highlighting the key interacting residues (Phe253–Phe253, His257–Ser250 and Ser250–His257).
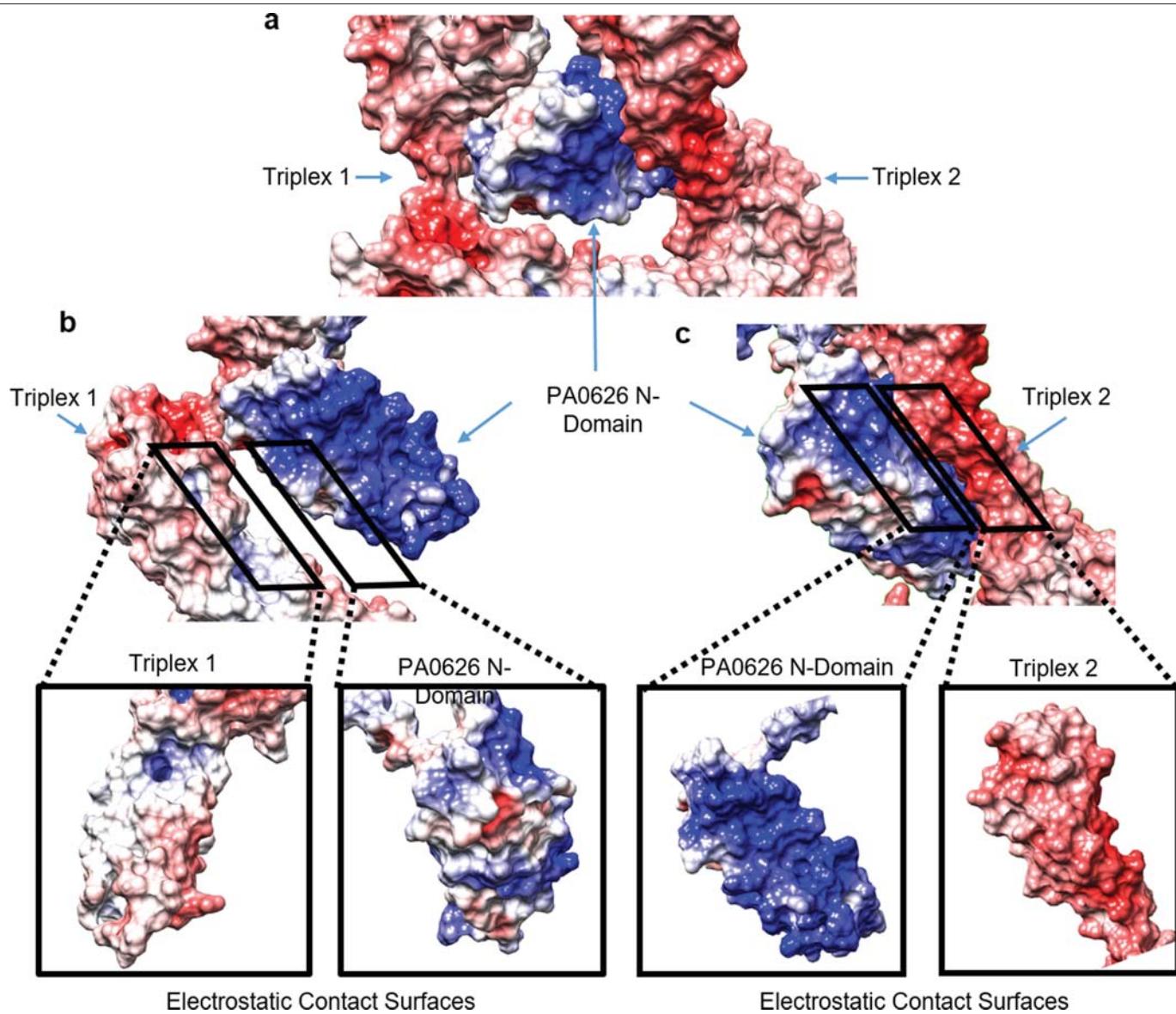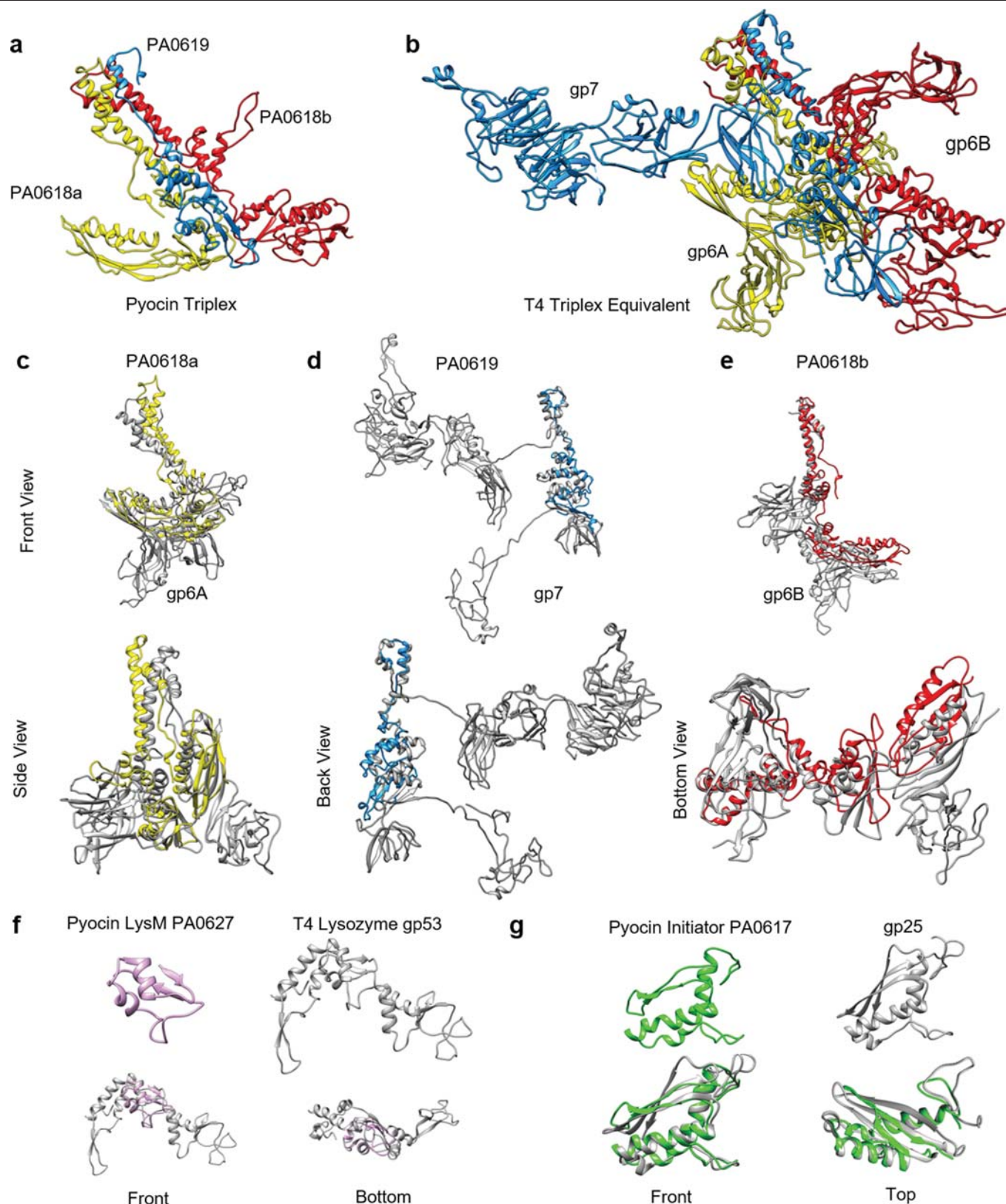
Extended Data Fig. 8 | Electrostatic views of the ripcord handle.
a, Electrostatic surface diagram of the ripcord with adjacent triplexes.
b, **c**, Electrostatic properties of the interfaces between the ripcord and triplexes 1 and 2, respectively. The colour code corresponds to positive (blue), neutral (white) and negative (red).

Extended Data Fig. 9 | Comparison of related protein subunits from pyocin R2 and T4 phage. a, Ribbon diagram of the pyocin triplex. **b**, Ribbon diagram of the T4 triplex equivalent with subunits marked by corresponding colour to panel **a**. **c**, Ribbon diagram of pyocin PA0618a (yellow) and T4 gp6A (grey). **d**, Ribbon diagram of pyocin PA0619 (blue) and T4 gp7 (grey). **e**, Ribbon diagram of pyocin PA0618b (red) and T4 gp6B (grey). **f**, Ribbon diagram of pyocin PA0627 (pink) and T4 gp53 (grey). **g**, Ribbon diagram of pyocin PA0617 (green) and T4 gp25 (grey).

Extended Data Table 1 | Cryo-EM data statistics

	Pre-contraction Trunk (EMDB-20526) (PDB 6PYT)	Pre-contraction Baseplate (EMDB-20643) (PDB 6U5B)	Pre-contraction Hub (C3 sym.) (EMDB-20646) (PDB 6U5H)	Pre-contraction Collar (EMDB-20644) (PDB 6U5F)	Post-contraction Collar (EMDB-20647) (PDB 6U5J)	Post-contraction Baseplate (EMDB-20648) (PDB 6U5K)
Data collection and processing						
Magnification	130,000	130,000	130,000	130,000	130,000	130,000
Voltage (kV)	300	300	300	300	300	300
Electron exposure (e–/Å^2)	30	30	30	30	30	30
Defocus range (μm)	1.2-3.4	1.2-3.4	1.2-3.4	1.2-3.4	1.2-3.4	1.2-3.4
Pixel size (Å)	1.07	1.07	1.07	1.07	1.07	1.07
Symmetry imposed	C6 + helix	C6	C3	C6	C6	C6
Initial particle images (no.)	15,680	43,934	43,934	43,934	36,116	36,116
Final particle images (no.)	15,500	22,001	21,844	4,110	15,582	9,934
Map resolution (Å)	2.9	3.5	3.9	3.8	3.5	3.5
FSC threshold	0.143	0.143	0.143	0.143	0.143	0.143
Map resolution range (Å)	2.9-200	3.4-200	3.7-200	3.8-200	3.5-200	3.5-200
Refinement						
Initial model used (PDB code)	*de novo*	*de novo*	*de novo*	*de novo*	*de novo*	*de novo*
Model resolution (Å)	3.1	3.4	4.0	3.7	3.5	3.5
FSC threshold	0.5	0.5	0.5	0.5	0.5	0.5
Model resolution range (Å)	2.9-200	3.4-200	3.7-200	3.8-200	3.5-200	3.5-200
Map sharpening B factor (Å^2)	-80	-80	-80	-80	-80	-80
Model composition						
Non-hydrogen atoms	4,9920	103,392	7,326	107,550	59,634	110,142
Protein residues	6,624	13,608	954	14,232	7,878	14,520
Ligands	0	0	0	0	0	0
B factors (Å^2)						
Protein	45.39	41.54	55.81	13.62	37.18	76.64
Ligand						
R.m.s. deviations						
Bond lengths (Å)	0.009	0.005	0.005	0.007	0.005	0.005
Bond angles (°)	1.000	0.941	1.054	0.999	0.700	0.770
Validation						
MolProbity score	1.83	1.87	2.13	1.92	2.15	2.13
Clashscore	5.45	6.35	9.36	6.33	5.69	5.85
Poor rotamers (%)	0.94	0.17	0.39	0.38	3.03	3.00
Ramachandran plot						
Favored (%)	90.48	91.13	86.41	89.01	93.32	92.97
Allowed (%)	9.52	8.87	13.59	10.99	7.68	7.03
Disallowed (%)	0.00	0.00	0.00	0.00	0.00	0.00

FSC, Fourier shell correlation; R.m.s., root mean square.

Article

Extended Data Table 2 | Crystallographic data statistics

Dataset name*	PA0618C HoCl$_2$	PA0618C	PA0616d	PA0616
Data collection				
Space group	P4$_3$	P4$_3$	P6$_3$22	P2$_1$
Cell dimensions:				
a, b, c (Å)	102.67, 102.47, 45.76	72.61, 72.61, 46.52	46.68, 46.68, 144.55	94.56, 137.052, 164.44
α, β, γ (°)	90.00, 90.00, 90.00	90.00, 90.00, 90.00	90.00, 90.00, 120.00	90.00, 106.88, 90.00
Wavelength (Å)	1.5352	0.91963	0.9763	1.6984
Resolution (Å)	50.0 – 2.80 (2.97 – 2.80)#	50.0 – 2.10 (2.23 – 2.10)	50.0 – 1.46 (1.55 – 1.46)	50.0 – 2.20 (2.33 – 2.20)
R_{meas} (%)	9.7 (58.3)	3.2 (73.6)	3.4 (53.4)	6.5 (95.2)
CC$_{1/2}$	99.9 (91.5)	100.0 (81.8)	100.0 (98.4)	99.8 (83.2)
I / σ_I	25.35 (4.65)	30.41 (2.61)	56.52 (5.35)	12.70 (1.33)
Completeness (%)	99.2 (96.3)	99.6 (97.8)	99.6 (97.7)	97.1 (92.3)
Redundancy	13.9 (13.5)	6.9 (6.7)	21.8 (16.0)	3.21 (2.95)
Anomalous signal$	1.770 (0.863)			
Refinement				
Software (version)		Phenix_refine (1.9_1692)	Phenix_refine (1.8.2_1309)	Phenix_refine (1.8.2_1309)
Resolution (Å)		50.0 – 2.10	50.0 – 1.46	50.0 – 2.20
No. unique reflections		27,538	16,615	203,160
No. atoms				
Protein		1,403	744	23,767
Ligand/ion		0	1	238
Water		58	118	2,432
Rwork / Rfree		0.190 / 0.236	0.149 / 0.188	0.163 / 0.208
B-factors:				
Protein (Å^2)		74.92	28.64	54.02
Ligand/ion (Å^2)		n/a	24.51	117.84
Water (Å^2)		68.68	46.50	55.16
R.m.s. deviations				
Bond lengths (Å)		0.003	0.008	0.004
Bond angles (°)		0.705	1.274	0.889
Ramachandran plot:				
Favored (%)		97.77	98.86	96.03
Allowed (%)		2.23	0.0	3.47
Outliers (%)		0.0	1.14	0.50
PDB code		**5CES**	**4S36**	**4S37**

*A single crystal was used for each of the datasets.
#The statistics for the highest-resolution shell are given in parenthesis.
$As calculated by the program XDS (http://xds.mpimf-heidelberg.mpg.de/).

Extended Data Table 3 | Pyocin protein identification by mass spectrometry

MW (kDa)	Amino Acids	Gene Number (prf)	Gene Number (PA)	Copies (of monomer)
7.5	68	22	PA0627	Unknown
11.8	108	12	PA0617	6-18
18.1	168	18	PA0623	100-200
18.9	171	10	PA0615	Unknown
19.4	185	11	PA0616	6-18
20.0	177	14	PA0619	Unknown
31.3	290	21	PA0626	6-18
32.0	295	13	PA0618	6-18
35.9	329	23	PA0628	6-18
41.2	386	17	PA0622	100-200
77.7	745	20	PA0625	1

MW, molecular weight.

Corresponding author(s): Jeff F Miller, Z Hong Zhou

Last updated by author(s): Dec 16, 2019

Reporting Summary

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

Statistics

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

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

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

Software and code

Policy information about availability of computer code

Data collection	Leginon 3.3 (leginon.org);
Data analysis	Refmac5; Parrot; Phaser; Coot 0.8.9; Phenix 1.14-3260; CTFFind 3.1; MotionCorr 1 (with custom patch, available upon request); EMAN 1.81; Relion 1.2 (with custom IHRSR patch, available upon request, current version 1.4); Chimera 1.11.2; EMringer (https://fraserlab.com/2015/02/18/EMringer/); MatLab (MathWorks)

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

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:
- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

CryoEM maps and the associated atomic models have been deposited to EMDB, PDB under the accession numbers EMD-20526/PDB:6PYT (pre-contraction helical trunk), EMD-20643/PDB:6U5B (pre-contraction baseplate), EMD-20646/PDB:6U5H (pre-contraction hub in C3 symmetry), EMD-20644/PDB:6U5F (pre-contraction collar), EMD-20647/PDB:6U5J (post-contraction collar) and EMD-20648/PDB:6U5K (post-contraction baseplate), respectively. X-ray crystal structures have been deposited to PDB under the accession numbers 5CES (PA0618 C-terminal domain), 4S36 (PA0616 C-terminal domain) and 4S37 (full length PA0616). All other data are available from the corresponding authors upon reasonable request.

Field-specific reporting

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

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

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

Life sciences study design

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

Sample size	The presented study is of structural biological nature. We averaged huge numbers of particles (See Ext Data Table 1) to reach the stated resolutions, which is then tested and confirmed in the work. For Figure 4f, the phenotypical evaluation of H257F mutant, 200 micrographs were taken for each condition and particles counted. For Ext Data Figure 6, samples were prepared independently three times, 10 micrographs were taken for each of them (b and c), giving us 0 pre-contraction and about 150 post-contraction pyocins for the WT and 0 pre- and about 150 post-contraction pyocins for the mutant analyzed in total. For (d-g), each experiment was repeated three times.
Data exclusions	No data were excluded.
Replication	Ext Data Fig 6: all experiments were repeated three times.
Randomization	The structural determination process is random by nature since the particles are randomly presented to us in solution.
Blinding	The structural determination process is blinded by nature since the particles are presented to us in solution without selection by prior knowledge.

Reporting for specific materials, systems and methods

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

Materials & experimental systems

n/a	Involved in the study
☒	Antibodies
☒	Eukaryotic cell lines
☒	Palaeontology
☒	Animals and other organisms
☒	Human research participants
☒	Clinical data

Methods

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