






## Neck and capsid architecture of the robust *Agrobacterium* phage Milano

Ravi R. Sonani <sup>1</sup>, Nathaniel C. Esteves<sup>2</sup>, Abigail A. Horton<sup>2</sup>, Rebecca J. Kelly<sup>2</sup>, Amanda L. Sebastian<sup>2</sup>, Fengbin Wang <sup>1,4</sup>, Mark A. B. Kreutzberger<sup>1</sup>, Petr G. Leiman <sup>3✉</sup>, Birgit E. Scharf <sup>2✉</sup> & Edward H. Egelman <sup>1✉</sup>

Large gaps exist in our understanding of how bacteriophages, the most abundant biological entities on Earth, assemble and function. The structure of the “neck” region, where the DNA-filled capsid is connected to the host-recognizing tail remains poorly understood. We describe cryo-EM structures of the neck, the neck-capsid and neck-tail junctions, and capsid of the *Agrobacterium* phage Milano. The Milano neck 1 protein connects the 12-fold symmetrical neck to a 5-fold vertex of the icosahedral capsid. Comparison of Milano neck 1 homologs leads to four proposed classes, likely evolved from the simplest one in siphophages to more complex ones in myo- and podophages. Milano neck is surrounded by the atypical collar, which covalently crosslinks the tail sheath to neck 1. The Milano capsid is decorated with three types of proteins, a minor capsid protein (mCP) and two linking proteins cross-linking the mCP to the major capsid protein. The extensive network of disulfide bonds within and between neck, collar, capsid and tail provides an exceptional structural stability to Milano.

<sup>1</sup>Department of Biochemistry and Molecular Genetics, University of Virginia School of Medicine, Charlottesville, VA 22903, USA. <sup>2</sup>Department of Biological Sciences, Virginia Tech, Blacksburg, VA 24061, USA. <sup>3</sup>Department of Biochemistry and Molecular Biology, University of Texas Medical Branch, Galveston, TX 77555, USA. <sup>4</sup>Present address: Department of Biochemistry and Molecular Genetics, University of Alabama at Birmingham, Birmingham, AL 35233, USA. ✉email: [pgleiman@utmb.edu](mailto:pgleiman@utmb.edu); [bscharf@vt.edu](mailto:bscharf@vt.edu); [ehe2n@virginia.edu](mailto:ehe2n@virginia.edu)

**B**acteriophages have regained attention in the last few years<sup>1–3</sup> due to their potential in combating bacterial infections and overcoming antibiotic resistance. *Caudovirales*, the taxonomic order for the tailed double-stranded DNA bacteriophages, is characterized by the presence of a tail—a host cell attachment organelle that is connected to the genome-containing spherical or elongated capsid via a neck<sup>4</sup>. *Caudovirales* is sub-classified into three families: Myoviridae, which carry a long contractile tail; Siphoviridae that possess a long non-contractile tail; and Podoviridae with a very short non-contractile tail<sup>5</sup>. The tubular part of the tail is formed by helical polymerization of a tube protein in siphophages and an additional sheath protein in myophages<sup>6</sup>.

The neck connects the tail to the unique vertex of the capsid that is occupied by a dodecameric portal protein<sup>7</sup>. The neck and capsid contain a varying number of proteins that display a substantial structural diversity. For example, the neck may or may not be surrounded by collar-like structures<sup>8–15</sup>. Besides serving as a symmetry adaptor between a capsid and tail, the neck controls the release of phage DNA into the host cell during infection<sup>8,16,17</sup>. The tail to capsid junction involves a symmetry mismatch that is bridged by the neck, and is therefore crucial for the assembly of tailed phage. The mismatch is overcome by a head-proximal component of the neck establishing unique contacts to the icosahedral capsid. The structure of the neck alone has been studied for several phages<sup>8–18</sup>, and two studies have examined the symmetry mismatch at the neck-capsid interface<sup>8,18</sup>, which were distinctly different between the podophage Sf6 and myophage T4.

The fundamental skeleton of the capsid shell in all *Caudovirales* is formed by the icosahedral arrangement of hexamers and pentamers of the major capsid protein (MCP) having a conserved HK97-fold<sup>19–21</sup>. This capsid shell can be “decorated” by additional domains of the MCP or minor capsid proteins (mCPs) that play an important role in capsid stability and may participate in host recognition during infection.

Bacteriophage Milano<sup>22</sup> is a myophage that infects *Agrobacterium tumefaciens*, a plant pathogen that is being widely used in the genetic manipulation of plants. Milano requires an actively rotating flagellum on the surface of the bacterial host to which it binds to initiate an infection. Milano’s genome (68 kb) has only recently been sequenced<sup>22</sup>. Its predicted proteome contains 1.4% cysteine residues, which is significantly higher than the cysteine content in other bacteriophages (e.g., 0.9% in T4 and 0.7% in A511 phage). Remarkably, these cysteines are even more prevalent (2.4%) in predicted structural proteins.

Here, we report the cryo-EM structures of various regions of the neck, including the neck-capsid and neck-tail junctions, as well as the capsid of Milano. We describe how the neck 1 protomers undergo structural polymorphism to connect the 12-fold neck with the 5-fold capsid vertex. Our results reveal the exceptional structure of the Milano collar, departing from the typical triple-helical coiled-coil fold observed in other phages, and how it crosslinks the tail sheath to the neck. The Milano capsid is decorated with a mCP and two linking proteins (LP1 and LP2), crosslinking the MCP to the mCP. Our structures further reveal an extensive network of disulfide bonds within and between neck, collar, capsid and tail, and we show how these bonds provide structural robustness to the phage.

## Results

**Cryo-EM reconstructions and structures of Milano neck region.** Six cryo-EM volumes have been reconstructed for different parts of the Milano neck by imposing different symmetries (Fig. 1, Table 1). The 15-fold, 12-fold and 3-fold symmetrized maps of the neck were used to build atomic models of the collar,

portal, and neck devoid of capsid (portal + neck 1 + neck 2 + tail terminator + collar), respectively. The 5-fold symmetrized map was used to build the model of the neck-capsid junction — the neck 1 and collar interacting with the 5-fold capsid vertex. A 3-fold symmetrized map focused at the neck-tail junction was used to understand the interactions between the collar and the tail. Finally, a no-symmetry-imposed map was used to generate an atomic model of the complete neck.

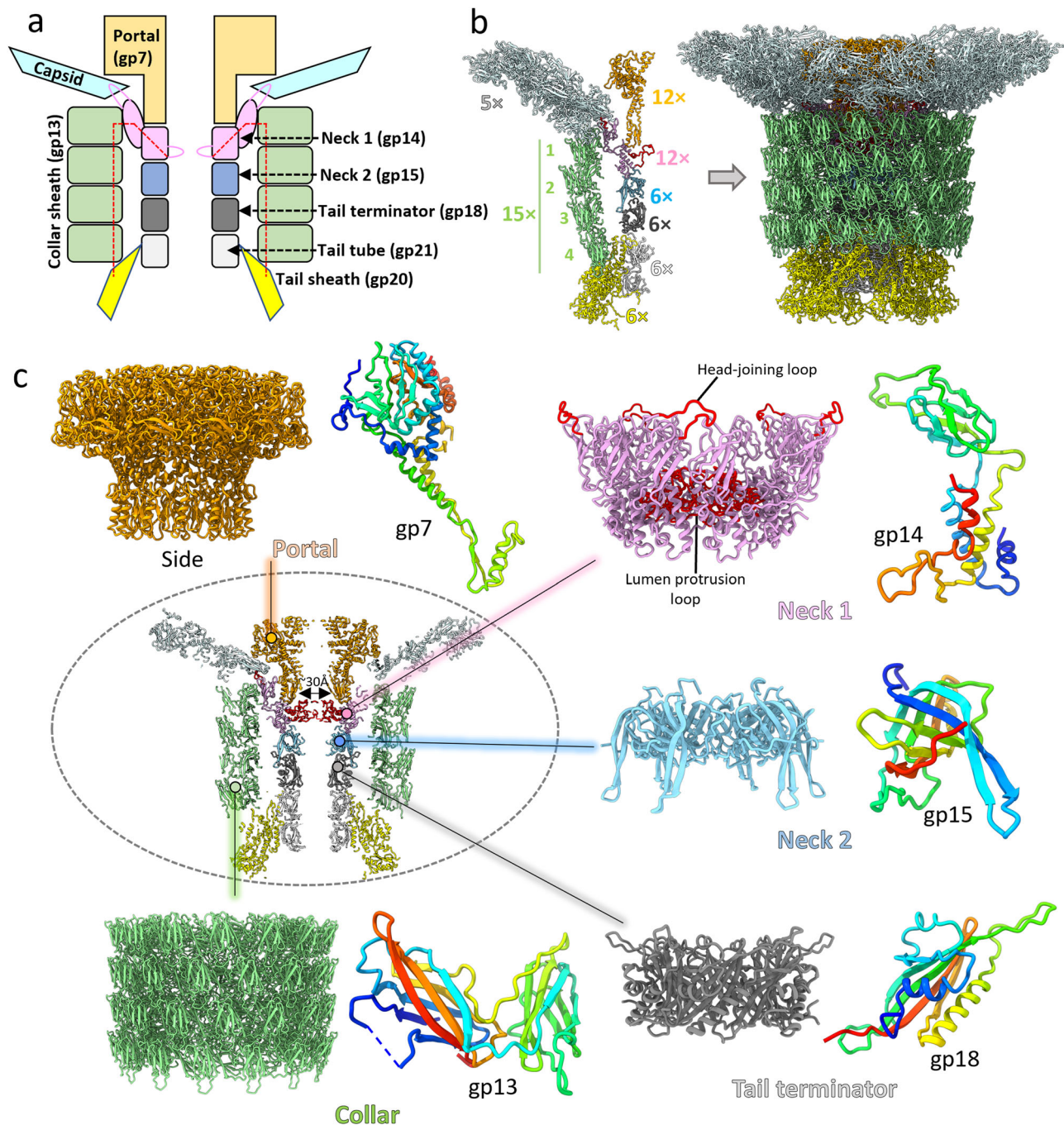
The inner part of the Milano neck complex is similar to that found in T4 and *Vibrio* XM1 bacteriophages<sup>10,11</sup> (Fig. 1, Supplementary Figs. S1, S2). Namely, the dodecameric portal ring (12 × gp7), which is embedded into the capsid shell, connects sequentially to the dodecameric neck 1 ring (12 × gp14), hexameric neck 2 ring (6 × gp15) and hexameric tail terminator ring (6 × gp18). The latter caps the tail tube and interacts with the sheath (see below). Besides these conserved common elements, the Milano neck complex possesses a unique, external, sheath-like structure surrounding the neck and consisting of four pentadecameric rings of gp13 (15 × gp13) that we name the collar.

As observed in the structure of R-type pyocin<sup>23</sup>, the tail terminator, while connecting the neck to the tail tube, buries its C-terminal arm in the handshake domain of tail sheath protomers belonging to the top ring<sup>23</sup> (Supplementary Fig. S3). The arm forms a  $\beta$ -strand that augments the  $\beta$ -sheet of the topmost sheath subunit in a manner resembling interactions between sheath subunits elsewhere in the sheath. The structures of portal, neck 2, and tail terminator rings of Milano show considerable similarity to their counterparts in other bacteriophages at both protomer and oligomeric levels (Supplementary Fig. S2)<sup>10,11,16,18,24–28</sup>. We thus focus on the unique structural features of the Milano neck, the neck 1 and collar.

**Milano neck 1 proteins are polymorphic to bridge the neck to the capsid.** Milano neck 1 protomer (gp14) consists of two main domains (Fig. 2): an  $\alpha$ -helical domain (residues 1–50, 141–159 and 189–199) containing four helices, a  $\beta$ -sandwich domain (residues 51–109 and 126–140) containing six anti-parallel  $\beta$ -strands, and two loops: a head joining loop (residues 110–125), and a lumen protrusion loop (residues 160–188) (Fig. 2a). The  $\beta$ -sandwich domain of Milano neck 1 is absent in its homologs in XM1, HK97, SPP1, P22, Sf6 and Mu, but present in T4 and T7 phages (Fig. 2b, Supplementary Fig. S2A)<sup>11,16,26,29,30</sup>.

In the dodecameric neck 1 assembly, the  $\alpha$ -helical and  $\beta$ -sandwich domains of gp14 form the core and peripheral ring, respectively (Figs. 1c, 2c). The core ring connects the neck 2 ring to the portal. The lumen protrusion loops form a diaphragm-like structure that constricts the channel spanning all proteins of the neck and the portal protein. These loops decrease the diameter of the neck lumen from ~30 Å to ~18 Å. The loop is formed by mostly polar residues, including a positively charged Arg167, which gives this diaphragm an overall positive charge (Supplementary Figs. S1C, S4). There are two alternating conformations of the lumen protrusion loop, with six tilted up and six tilted down. The RMSD between these two conformations of the loop is 9.7 Å (Fig. 2b).

The peripheral ring of neck 1 breaks the C12 symmetry of the core and connects the neck to the capsid. The 12 ring-forming gp14 protomers exhibit structural polymorphism. Their  $\beta$ -sandwich domains adopt different orientations in each protomer and can be described as 12 steps in a continuous tilting relative to the  $\alpha$ -helical domain (Fig. 2b). When the  $\alpha$ -helical domains of the 12 gp14 subunits are superimposed, they display an RMSD of 2.6 Å, while the  $\beta$ -sandwich domains have a varying orientation with the maximal distance between their distal parts reaching 24 Å. The  $\beta$ -sandwich domains are clustered in three groups with



**Fig. 1 Cryo-EM structure of Milano neck.** **a** Schematic of structural organization of Milano neck. The covalent connection between the neck 1 lumen protrusion loop and the tail sheath through the collar is indicated by the red dashed line. **b** Milano consists of stacked protein rings having different symmetries. The portal, neck 1, neck 2/tail terminator and collar has C12, pseudo-C12, C6 and C15 symmetry, respectively. Tail terminator connects the neck to the C6 tail tube and tail sheath. The C12 portal ring is embedded into the capsid at the 5-fold portal vertex. **c** Enlarged view of portal, neck 1, neck 2, collar and tail terminator rings with their component subunits colored blue to red, N- to C-terminal.

3, 4, and 5 members each (Fig. 2c, d) creating a 12-to-5-fold symmetry adaptor assembly to bridge neck-capsid symmetry mismatch.

Five out of 12 gp14  $\beta$ -sandwich domains ( $n$ ,  $n + 2$ ,  $n + 5$ ,  $n + 7$ ,  $n + 10$ , CCW direction looking from the capsid side) display pseudo-5-fold symmetry (Fig. 2c, d). These domains are covalently linked to the MCP (gp9) by two disulfide bonds, gp<sup>14</sup>Cys117–gp<sup>9</sup>Cys252 and gp<sup>14</sup>Cys120–gp<sup>9</sup>Cys440, which emanate from a head joining loop of the  $\beta$ -sandwich domains (Fig. 2e). The head joining loops of the remaining seven gp14

protomers are not resolved in our maps and thus likely to be disordered. The  $\beta$ -sandwich domains form a positively charged “bowl”, which houses a negatively charged portal clip domain region<sup>31</sup> thus strengthening the connection between the portal and the neck tube (Supplementary Fig. S1C).

In summary, in the gp14 neck 1 dodecamer, the  $\alpha$ -helical domain connects the portal to neck 1 ring smoothly while maintaining C12 symmetry, while its  $\beta$ -sandwich domain binds to the five-fold symmetric capsid vertex by adopting 12 different conformations. Electrostatics appears to play a role in the



**Table 1 Cryo-EM data collection and refinement statistics of Milano Neck reconstructions and models.**

	Portal	Collar	Neck C3	Neck-Capsid junction	Neck-tail junction	Neck C1	Capsid
PDB entry	8FWB	8FWC	8FWE	8FWG	8FWM	8FXR	8FXP
EMBD entry	29500	29501	29503	29504	29512	29541	29540
Voltage (kV)	300	300	300	300	300	300	300
Magnification (x)	81,000	81,000	81,000	81,000	81,000	81,000	81,000
Electron exposure (e-/ Å <sup>2</sup> )	50	50	50	50	50	50	50
Pixel size (Å/pixel)	1.08	1.08	1.08	1.08	1.08	1.08	1.44
Particle images (no.)	10,216	10,216	10,216	10,083	10,216	10,083	15,740
Symmetry imposed							
Point group	C12	C15	C3	C5	C3	C1	C5
Map global resolution (Å)							
Map:map FSC (0.143)	3.1 Å	3.0 Å	3.5 Å	3.5 Å	3.5 Å	4.5 Å	4.0 Å
Model:map FSC (0.5)	3.4 Å	3.1 Å	3.6 Å	3.6 Å	3.6 Å	4.7 Å	4.5 Å
Refinement and model validation							
Clash score	11.58	7.24	12.62	10.48	13.63	12.14	17.83
R.M.S. deviations							
Bond Length (Å)	0.005	0.009	0.011	0.005	0.004	0.003	0.011
Bond Angle (°)	0.819	0.993	0.865	0.733	0.844	0.680	0.781
Ramachandran plot (%)							
Outliers	0.00	0.00	0.18	0.10	0.35	0.15	0.41
Allowed	12.82	10.33	11.50	12.06	13.63	10.09	13.71
Favoured	87.18	89.67	88.32	87.84	86.01	89.76	85.88

interaction of the  $\beta$ -sandwich domain with the portal clip region. Finally, lumen protrusion loops form a two layered diaphragm-like structure in the lumen of the neck by adopting two alternating conformations (Fig. 2f).

**The collar crosslinks the tail sheath to neck 1.** The Milano neck is covered by a sheath-like structure in place of the collar/whiskers (gp *wac*) of T4<sup>10</sup>, collar spikes (gp40) of XM1<sup>11</sup>, and pre-neck appendage (equivalent to collar, gp12) of phi29<sup>32</sup> phages. No protruding fibrous elements such as whiskers/spikes/appendages emanate from the Milano collar.

The Milano collar is made up of four rings of gp13 pentadecamers (15 × gp13). Since this pentadecamer has both the C5 symmetry of the capsid and a C3 symmetry (present in the C6 symmetry of the tail), it is resolved in both the C3 and C5 symmetric reconstructions of the neck (Fig. 1). Interestingly, Milano gp13 is devoid of  $\alpha$ -helices, which are found in collar proteins of T4, XM1 and phi29 phages with many of these helices forming triple helical coiled coils<sup>10,11,32</sup> (Supplementary Fig. S5). The gp13 subunit contains two  $\beta$ -strand-rich jellyroll-like domains. Domain 1 (residues 1–69, 167–230) forms the wall of the collar, and we thus named it the wall domain (W domain), whereas domain 2 (residues 70–166) decorates the collar wall, and we named it the decorating domain (D domain) (Fig. 3a, b). DALI search results suggest that the D domain of gp13 is structurally similar to glycan-interacting and cell surface binding proteins (Supplementary Table S1).

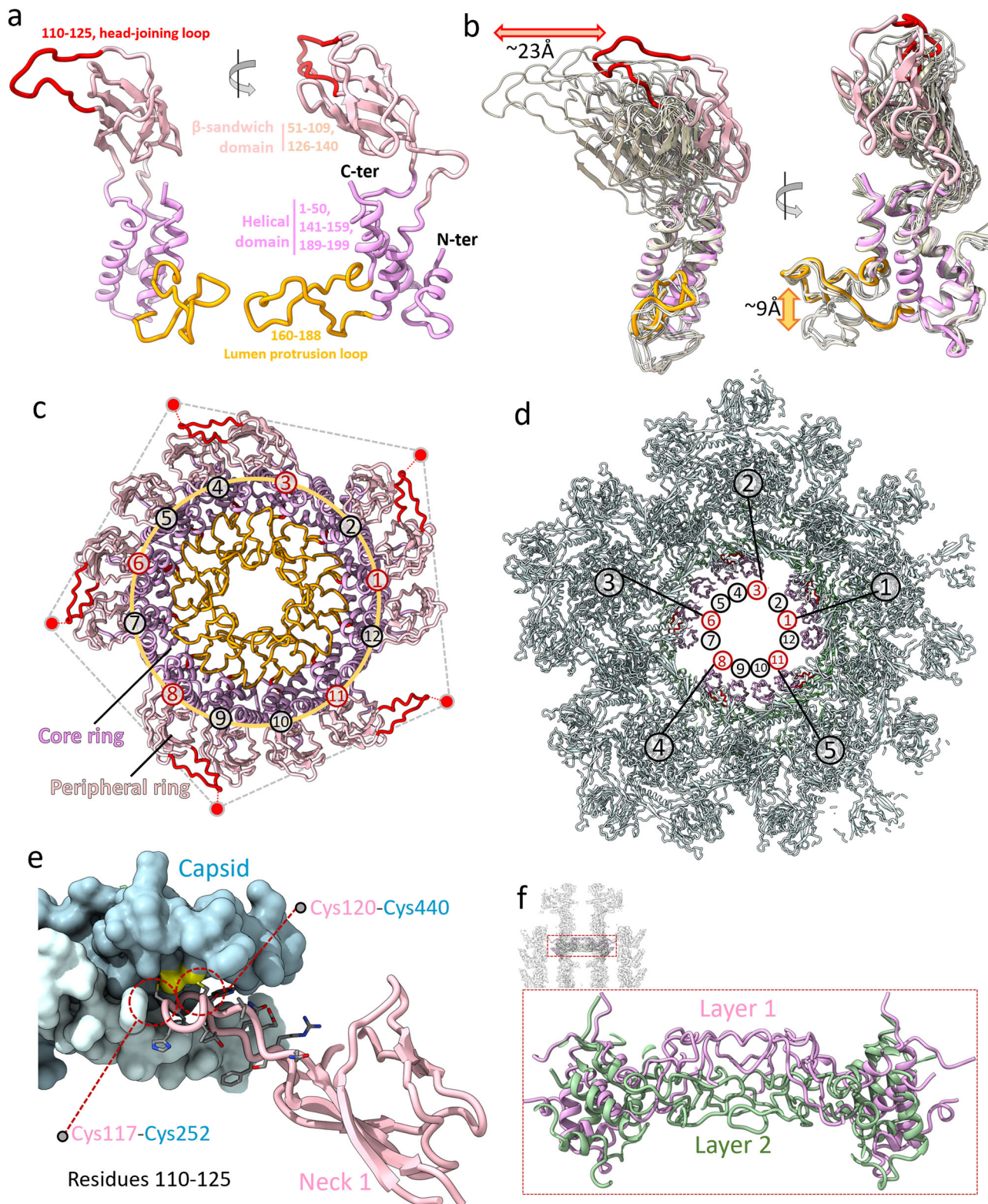
Four collar rings are stacked and stabilized by complementary electrostatic charges on the interacting surfaces of their W domains (Fig. 3e). The W domain of each gp13 protomer is covalently linked to its neighboring subunits by four disulfide bonds, one with each neighbor within the same ring (Cys23–Cys229) and one with each neighbor along the left-handed 12-start helices (Cys22–Cys211) (Fig. 3c). The collar connects the neck 1 with the uppermost layer of the sheath without interacting with the neck components in its midsection. Namely, the head-proximal end of the collar interacts with the  $\beta$ -sandwich domain of neck 1 (gp14), while the tail-proximal end of the collar interacts with the tail sheath (gp20) (Figs. 1a and 3d) (Table 2). Every third (n, n + 3, n + 6, n + 9, n + 12) collar subunit is attached to the neck 1 subunits (n, n + 2, n + 5, n + 7, n + 10,

CCW direction looking from the capsid side) through the disulfide bond gp<sup>14</sup>Cys136–gp<sup>13</sup>Cys19, linking the collar to neck 1 by five disulfide bonds. It is possible that other collar subunits might be making additional disulfide bonds not resolved in our map.

The interaction of the C15-symmetric collar with the C6-symmetric tail sheath is more complex. Five collar subunits (e.g., n, n + 1, n + 2, n + 3, n + 4) comprising the asymmetric unit interact with four tail sheath subunits belonging to the two topmost sheath rings (or layers) that we named Fn, Fn + 1, Sn, Sn + 1, where F and S denote first and second ring, respectively. The collar subunits n (Cys19), n + 3 (Cys19), and n + 4 (Cys23) form disulfide bonds with tail sheath subunits Sn (Cys467), Sn + 1 (Cys467) and Fn + 1 (Cys244), respectively, thus linking the collar to the tail sheath by nine (three in each of the three asymmetric units) disulfide bonds (Table 2). Intra- and inter-region disulfide bonds of the collar form a covalent connection between the tail sheath and the neck 1 lumen protrusion loop through the collar, running from the tail sheath as follows: [sheath gp<sup>20</sup>Cys467/Cys244] → [collar gp<sup>13</sup>Cys19–Cys24<sub>ring4</sub> → gp<sup>13</sup>Cys221–Cys211<sub>ring4</sub> → gp<sup>13</sup>Cys22–Cys24<sub>ring3</sub> → gp<sup>13</sup>Cys221–Cys211<sub>ring3</sub> → gp<sup>13</sup>Cys22–Cys24<sub>ring2</sub> → gp<sup>13</sup>Cys221–Cys211<sub>ring2</sub> → gp<sup>13</sup>Cys22–Cys19<sub>ring1</sub>] → [neck 1 gp<sup>14</sup>Cys136–lumen protrusion loop, 160–188] (Fig. 3d).

**An extensive network of disulfide bond provides structural stability.** Our atomic models reveal that most of the cysteine residues of Milano structural proteins are involved in disulfide bonds, covalently crosslinking the neck, collar, capsid, and tail. Such an extensive network of disulfide bonds would be expected to strengthen the overall structure of the Milano virion. To test this, we compared the ability of Milano to withstand disintegration by ultrasonic disruption with that of phage T4 and Chi, which contain an average number of cysteines and disulfide bonds.

We found that T4 and Chi fragmented within 10 s under sonication-induced mechanical stress, while Milano particles stayed intact for more than 30 s (Supplementary Fig. S6). In contrast, upon reduction of the disulfide bonds with dithiothreitol (DTT), Milano particles disintegrated within 10 s of sonication similar to T4 and Chi in the absence of DTT (Supplementary Fig. S6). Furthermore, DTT-treated Milano particles disintegrated



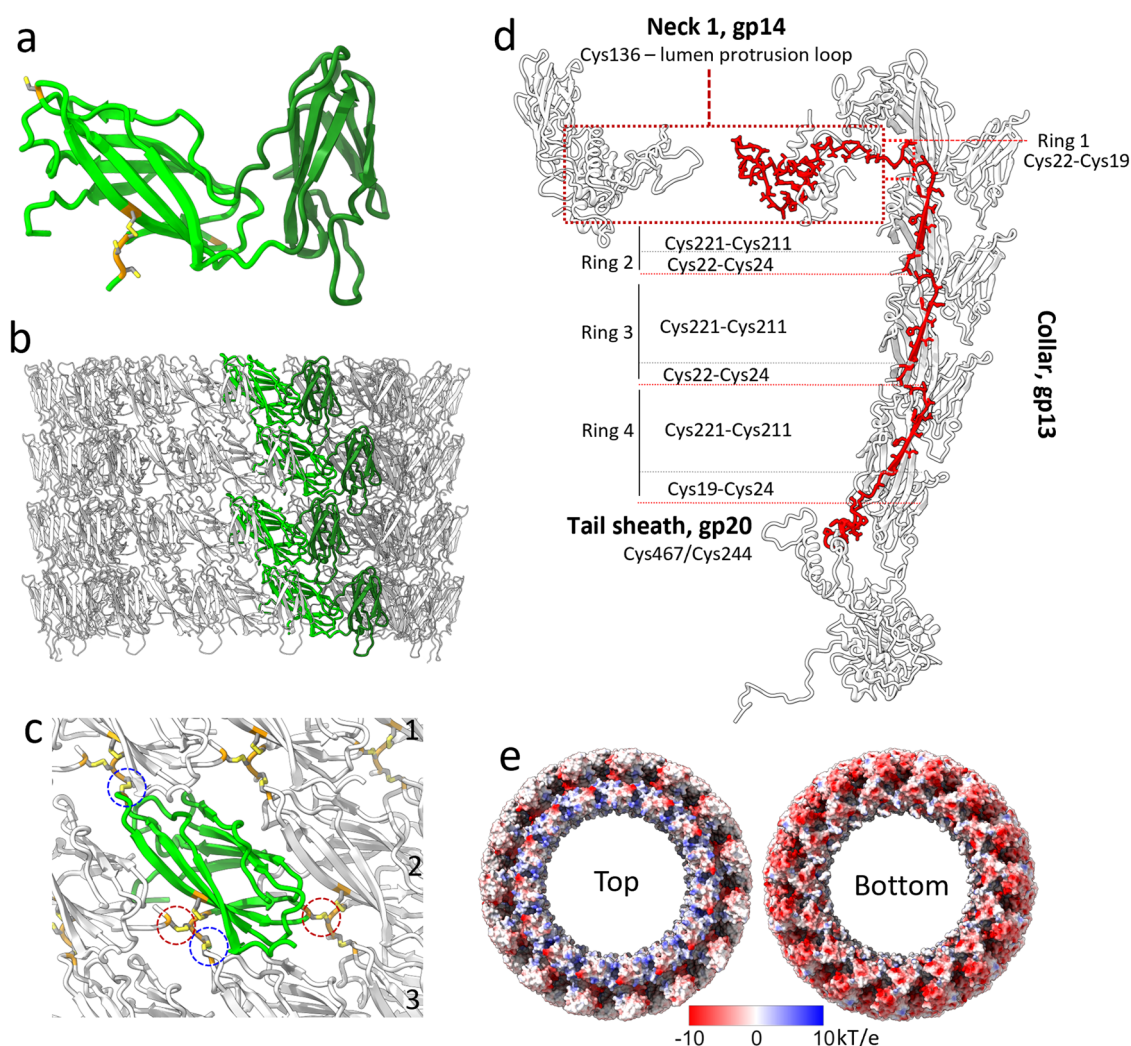
within 90 s when heated to 60 °C, while untreated Milano was resistant to 60 °C heating for more than 5 min (Supplementary Fig. S6). The loss of structure in the presence of a reducing agent suggests that the disulfide bonds provide extended stability to Milano particles.

Other myophages do not possess such a high percentage of cysteines and resultant disulfide bonds. However, the closely

related *Agrobacterium* phages, 7-7-1<sup>33</sup> and OLIVR4 (Taxon ID: 2723772), exhibit high sequence conservation including cysteine residues in their structural proteins (Supplementary Fig. S7). Therefore, the disulfide bonds and robustness of Milano appear to be correlated with the *Agrobacterium* host. The unique collar and the abundance of disulfide bridges in Milano may be an adaptation to the specific environment associated with *Agrobacterium*.



**Fig. 2 Structural organization of neck 1.** **a** Structure of neck 1 protein, gp14. The helical domain,  $\beta$ -sandwich domain, head joining (residues, 110–125) and lumen protrusion loops (residues 160–188) are colored dark pink, light pink, red and orange, respectively. **b** Superimposed view of all 12 protomers of gp14 in the Milano neck 1 ring, aligned at their helical-domain. The different conformations of  $\beta$ -sandwich domain and dual conformations of lumen protrusion loop are indicated by red and orange double-headed arrows, respectively. **c** Top view of neck 1 assembly in Milano. Color scheme is same as Fig. 2a. The  $\beta$ -sandwich domains are clustered asymmetrically in groups of three, four and five. The head joining loop of five protomers ( $n$ ,  $n + 2$ ,  $n + 5$ ,  $n + 7$ ,  $n + 10$ , CCW direction) are connected to MCPs at the 5-fold portal vertex, indicated by red circles. **d** Same view of Fig. 2c with capsid (grey), showing five out of twelve gp14 protomers ( $n$ ,  $n + 2$ ,  $n + 5$ ,  $n + 7$ ,  $n + 10$ , CCW direction) deviating from C12 symmetry to connect with the 5-fold capsid. **e** Interaction of the head joining loop of gp14 with MCP (gp9). The gp14 and gp9 subunits are shown as pink ribbon and cyan surface models, respectively. Cysteines involved in disulfide bonds are colored yellow and indicated by red dashed circles. **f** The lumen protrusion loop of gp14, by adopting a dual conformation, creates a two layered diaphragm-like structure in the lumen of the Milano neck. Two different layers created by different conformations are shown in pink and green.



**Fig. 3 Structural organization of the collar.** The structure of (a) collar protomer gp13 and (b) assembly. The W (wall)- and D (decorating) domains of gp13 are colored light and dark green, respectively. Cysteines involved in disulfide bonds are colored yellow. **c** Disulfide bond network between collar subunits. Each gp13 protomer makes four disulfide bonds, one with each neighbor within the same ring (Red circles, Cys23–Cys229), and one with each neighbor related by left-handed 15-start helices in an adjacent ring (Blue circles, Cys22–Cys211). **d** The collar forms a covalent structural link between the tail sheath and the neck 1 lumen protrusion loop: [sheath gp20Cys467/Cys244] → [collar gp13Cys19–Cys24<sub>ring4</sub> → gp13Cys221–Cys211<sub>ring4</sub> → gp13Cys22–Cys24<sub>ring3</sub> → gp13Cys221–Cys211<sub>ring3</sub> → gp13Cys22–Cys24<sub>ring2</sub> → gp13Cys221–Cys211<sub>ring2</sub> → gp13Cys22–Cys19<sub>ring1</sub>] → [neck 1 gp14Cys136–lumen protrusion loop, 160–188]. **e** Electrostatic potential of top and bottom surfaces of gp13 pentadecameric rings.

**The capsid is decorated with a minor capsid protein and linking proteins.** The Milano capsid has a  $T = 9$  icosahedral symmetry. It consists of three structural elements that are formed by the MCP (gp9): a pentamer at the icosahedral 5-fold axis, hexamer 1 at the icosahedral 3-fold, and hexamer 2 positioned on a local 2-fold (Fig. 4a). The icosahedral asymmetric unit is formed by a single subunit of the pentamer, two subunits of hexamer 1,

and a complete hexamer 2. Hexamer 1 and hexamer 2 differ in their decorating proteins, as described below.

The Milano MCP has a HK97-like fold found in all MCPs of dsDNA tailed phages (Supplementary Fig. S2)<sup>18,27,28</sup>. It consists of an A domain that forms the core, and a P domain, E loop, and N arm that form the peripheral parts of hexamers and pentamers (Fig. 4b). The conformational plasticity in the P domains, E loops,

**Table 2 Inter-chain disulfide bonds within and between sub-structures of Milano neck.**

Within sub-structure		
	Residue	Bonded To
Collar	gp13, Cys22	gp13, Cys211
	gp13, Cys23	gp13, Cys229
Capsid	MCP, gp9, Cys302	MCP, gp9, Cys455
	MCP, gp9, Cys186	MCP, gp9, Cys456 (intrachain)
	MCP, gp9, Cys190	MCP, gp9, Cys385 (intrachain)
	MCP, gp9, Cys223	MCP, gp10, Cys105
	MCP, gp9, Cys252	MCP, gp10, Cys74
	mCP, gp10, Cys99	LP1, gp16, Cys3
	mCP, gp10, Cys60	LP1, gp16, Cys7
	mCP, gp10, Cys99	LP2, gp128, Cys6
	mCP, gp10, Cys60	LP2, gp16, Cys10
	LP2, gp16, Cys23	LP2, gp128, Cys26
	LP1, gp16, Cys20	LP2, gp128, Cys26
Between sub-structures		
Collar—Neck 1	Collar gp13, Cys19	Neck 1, gp14, Cys136
Tube terminator—Tail tube	Tube terminator, gp18, Cys156	Tail tube, gp21, Cys3
Collar—Tail sheath	Collar, gp13, Cys19	Tail sheath, gp20, Cys467
	Collar, gp13, Cys23	Tail sheath, gp20, Cys244
Neck 1—Capsid	Neck 1, gp14, Cys117	MCP, gp9, Cys252
	Neck 1, gp14, Cys120	MCP, gp9, Cys440

MCP major capsid protein, mCP minor capsid protein, LP1 linking protein 1, LP2 linking protein 2.

and N arms of the MCP enable the stable construction of quasi-equivalent pentamers and hexamers (Fig. 4f). Hexamers and pentamers are further decorated by additional proteins: mCP (gp10) and two linking proteins (LP)—LP1 (gp16) and LP2 (gp128). The mCP contains a  $\beta$ -strand-rich jellyroll-like domain (Fig. 4b). LP1 is predicted by AlphaFold to have a  $\beta$ -sheet Ig-type domain with an N-terminal unstructured strand. LP2 is a small 38 residue protein with a high degree of sequence similarity to the N-terminal strand of LP1 (Fig. 4b; Supplementary Fig. S8).

A dimer of mCP decorates the external surface of the capsid by cementing the pentamer-hexamer 1 and hexamer 1-hexamer 2 interfaces, where it forms disulfide bonds with the MCPs (gp<sup>9</sup>Cys223—gp<sup>10</sup>Cys105 and gp<sup>9</sup>Cys252—gp<sup>10</sup>Cys74) (Fig. 4c). The resolved parts of LPs—residues 1–13 of LP1 and 1–10 of LP2—occupy quasi-equivalent positions with an mCP at the periphery of pentamer and hexamers and run towards the centers of the pentamers and hexamers, where they interact with the MCPs (Fig. 4c; Supplementary Fig. S9; Table 2). LP1 and LP2 form all such linkages in the pentamer and hexamer 1, respectively. In hexamer 2, all but one linkage is formed by LP2, while the unique pentamer-proximal link is formed by LP1 (Fig. 4c; Supplementary Fig. S8). Notably, the LP1–mCP and LP2–mCP interaction interfaces also include a disulfide bonds. LP1<sup>Cys3/LP2Cys6</sup> and LP1<sup>Cys7/LP2Cys10</sup> form disulfide bond with mCP<sup>Cys99</sup> and mCP<sup>Cys60</sup>, respectively (Table 2; Supplementary Fig. S9).

The C-terminal Ig-type domain of LP1 appears to extend away from the capsid shell without forming much contact with it (Fig. 4d, e). Consequently, the density of the LP1 Ig-type domain is not resolved due to positional disorder. The capsid-decorating Hoc protein of T4, which consists of several Ig-type domains, extends away from the T4 capsid in a similar fashion and the density for these domains is also not resolved<sup>34</sup>. The tethering of

Milano’s LP1 to the capsid via an N-terminal arm of LP1 is similar to the way the capsid auxiliary protein gp16 of phage YSD1 binds to its capsid<sup>28</sup>.

While LP2 seems to have only a structural role in strengthening the capsid, the mCP and LP1 may have an additional function. The jellyroll-like domain of mCP and Ig-type domain of LP1 are widespread in phage capsid proteins. The likely function of such proteins and domains is to participate in low-affinity binding to cell surface components (e.g. surface polysaccharides and receptors), which keeps the phage in the vicinity of potential host cells<sup>34–36</sup>. DALI analysis<sup>37</sup> shows that these proteins are structurally similar to many glycan and receptor binding proteins (Supplementary Table S1), for example, the GlcNAc-binding protein A that mediates attachment of *Vibrio* bacteria to the mammalian intestinal mucin (PDB ID: 2XWX)<sup>38</sup>, the variable domain of T-cell receptor (PDB ID: 1BWM) that binds the MHC I complex antigenic peptide<sup>39</sup>, and the C-terminal domain of a sensor kinase of a *Bacteroides* hybrid two-component system<sup>40</sup>.

Discussion

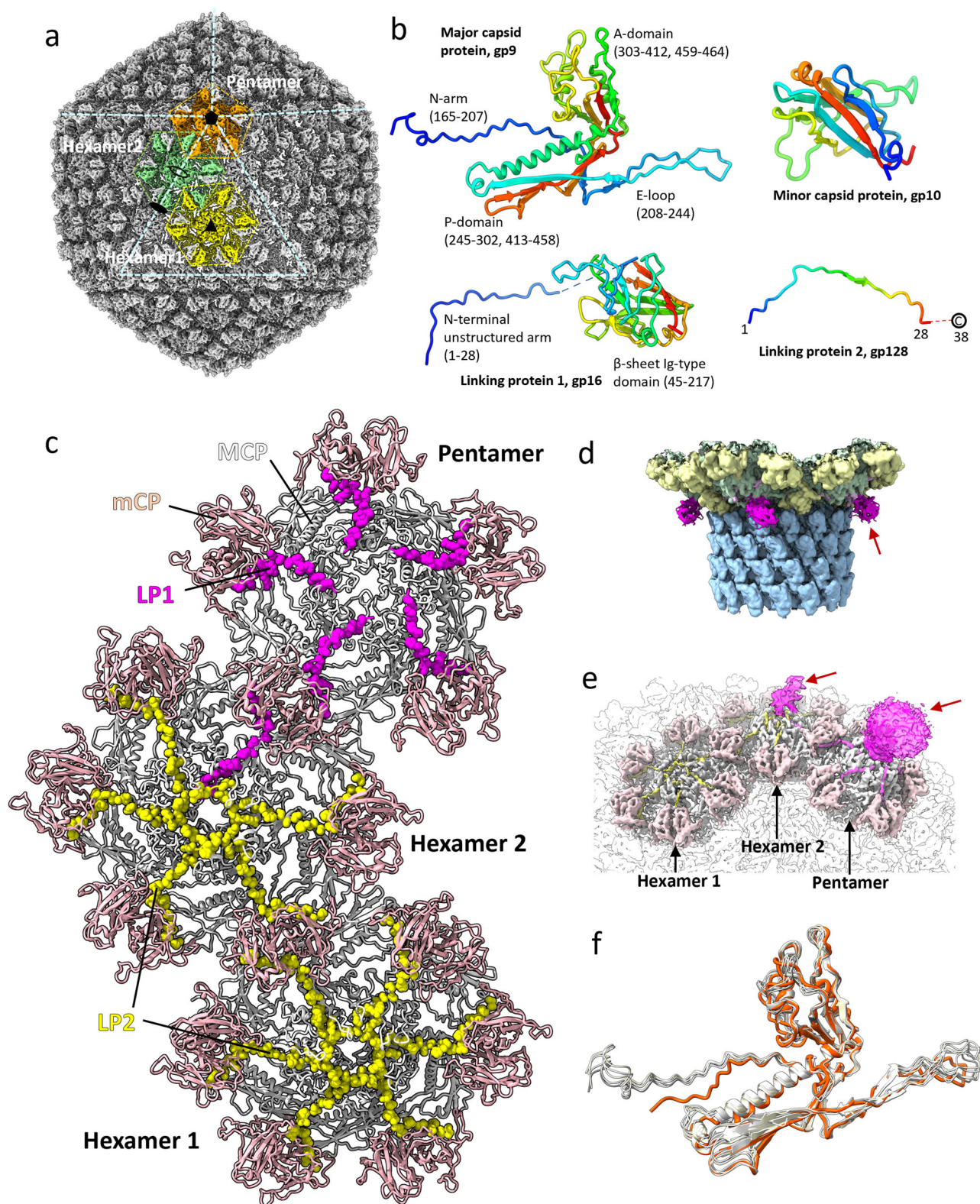
Our structural results show that Milano’s neck 1 ring plays a dual structural role: (1) it joins the neck to the capsid and resolves the symmetry mismatch between the two, and (2) it forms a two-layered diaphragm-like structure in the lumen of the neck. The head-to-neck symmetry mismatch has been described for T4<sup>18</sup> and Sf6<sup>8</sup> bacteriophages. The N-terminal whisker of the T4 portal protein (gp20) and the C-terminal whisker of the *Shigella* Sf6 head-to-tail adaptor protein (gp7) bridge the neck and head symmetries by conformational differences among its protomers. Our work characterizes how the  $\beta$ -sandwich domains of Milano neck 1 protomers (gp14) adopt different conformations to compensate for the head-to-neck symmetry mismatch and bridge them covalently. While this domain is largely absent in neck 1 homologs of some other bacteriophages, the equivalent proteins of T7 (gp11, PDB id: 6R21)<sup>17</sup> and T4 (gp13, AlphaFold-predicted, Uniprot# P11110) possess equivalent  $\beta$ -sandwich domains. However, the role of these  $\beta$ -sandwich domains in symmetry adjustment remains to be determined in T4 and T7 phages.

Structural comparisons of the Milano neck 1 homologs suggest the existence of at least four structural classes of neck 1-like proteins (Fig. 5). Class 1, containing proteins with ~100 residues, represents the simplest form composed of a single helical domain (e.g. gp15 in SPP1 and gp6 in HK97). In Class 2, ~140–160 residue-long proteins possess an additional extended C-terminal strand (e.g. gp4 in P22, gp7 in Sf6, and gp36 in Mu). In Class 3, ~190–210 residue-long proteins possess an additional  $\beta$ -sandwich domain and/or lumen protrusion loop (gp14 in Milano, gp11 in T7, and gp1 in XM1). In Class 4, ~310 residue-long proteins possess two additional  $\alpha$ -helical domains protruding out from the  $\beta$ -sandwich domain (gp13 in T4).

These observations suggest an evolutionary pattern of neck 1 protein developing multiple functions. The latter three classes (mostly found in myo- and podophages) appear to have evolved divergently from the same origin as class 1, the simplest structure mostly present in siphophages. Notably, the lumen protrusion loop exists only in myophages, whereas the presence of the  $\beta$ -sandwich domain is spread across myo- and podophages. Therefore, the structural joint between the neck and the portal appears to be the fundamental function of neck 1 in siphophages, whereas the adaptation to genome gating and neck-capsid symmetry bridging may have occurred later in myo- and myo-/podophages, respectively.

The diaphragm-like structure in the Milano neck has two layers formed by two alternating conformations of the neck 1





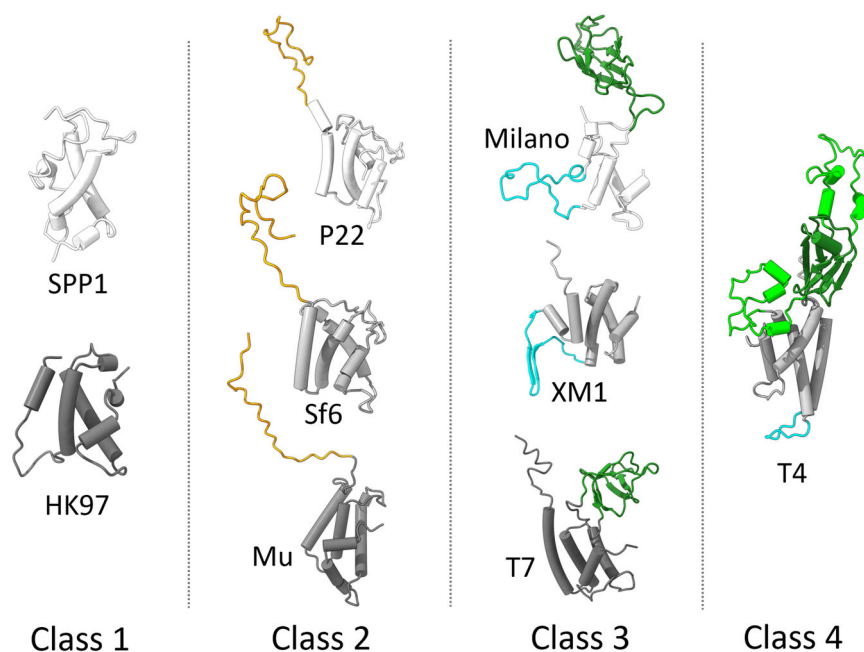
lumen protrusion loop. It is likely to be flexible in the monomeric state as its AlphaFold-predicted structure exhibits a very different conformation from the two observed experimentally in the cryo-EM map. It has been shown that structural homologs of Milano neck 1 exist as monomers in solution and oligomerize only in the presence of the portal and other neck proteins<sup>29,41</sup>. Two conformations of the neck 1 lumen protrusion loops are also likely to

be acquired during neck assembly. Stacks of layer 1 and 2 are stabilized by inter-layer hydrogen bonds between polar residues.

The function of the neck 1 diaphragm structure in the Milano neck remains to be determined. The diaphragm decreases the effective diameter of the neck lumen to  $<20$  Å, which is too narrow for the double-stranded DNA to pass through. Given its polar and positively charged nature, it is likely to interact with the



**Fig. 4 Structure of Milano capsid.** **a** Cryo-EM map of the Milano capsid shows that it consists of three building blocks: pentamer (orange), hexamer 1 (yellow) and hexamer 2 (green), each composed of the MCP (gp9). Icosahedral 2-fold, 3-fold and 5-fold symmetry are shown by filled black oval, triangle and pentagon, respectively. The local 2-fold symmetry is shown by an empty oval. **b** Structures of MCP (gp9), mCP (gp10), linking protein 1 (LP1, gp16) and linking protein 2 (LP2, gp128) are shown with N-to C-terminal colored blue-to-red. **c** Structure of pentamer, hexamer 1 and hexamer 2 differing in their decorating proteins. MCP, mCP, LP1 and LP2 are colored white, light pink, pink and yellow, respectively. Pentamer is decorated by five copies of LP2, hexamer1 is decorated by six copies of LP1 and hexamer 2 is decorated by five copies of LP2 and a single LP1. **d** Cryo-EM density corresponding to the  $\beta$ -sheet Ig-type domain of LP1, hanging out of hexamer 2 near the neck (Pink and indicated by red arrow). **e** Cryo-EM density corresponding to the  $\beta$ -sheet Ig-type domain of LP1 at pentamer and hexamer 2 hanging out of the capsid (Pink colored and indicated by red arrows). **f** Different conformations of P-domain, E-loop and N-arm of MCP (gp9) protomers in pentamer (orange) compared to those in hexamer 2 (grey).



**Fig. 5 Structural classification of proteins equivalent to Milano neck 1 found in various bacteriophages.** The structures for T4 and Mu phages are predicted by AlphaFold. The helical domain (white, black, grey),  $\beta$ -sandwich domain (green), C-terminal extension (orange) and stopper loop (cyan) are colored separately.

negatively charged DNA. The low-resolution density likely corresponding to the DNA near this diaphragm provides evidence for such an interaction (Supplementary Fig. S10). The diaphragm-like structure must be opened-up or widened for the release of the genome. The lumen protrusion loop of the neck 1 homolog in XM1 phage adopts an orientation parallel to the neck's central axis and provides a surface for neck 2 interaction<sup>11</sup>. The Milano neck 1 diaphragm might be adopting a similar structure in an open state to facilitate the genome gating.

The lumen of portal, neck 1, and neck 2 possesses a positive electrostatic potential resulting in a positive charge of the neck lumen (Supplementary Fig. S1)—suitable to hold and position the end of the negatively charged linear dsDNA until the genome release. In contrast, the tail terminator ring possesses a negatively charged lumen, repelling the DNA from the wall and thereby lubricating its passage (Supplementary Fig. S1).

The collar of Milano is structurally distinct from that of T4 and XM1 bacteriophage. In both, T4 and XM1, the collar is made up of one ring/disk that surrounds the neck and carries fibers or spikes. In contrast, the Milano collar consists of four rings (4 x gp13<sub>15</sub>) surrounding the whole neck and displaying no fibers or spike-like structures. The T4 collar whiskers (gp wac) are important for assembly of tail fibers and control of infection<sup>42</sup>. Based on those findings, XM1 collar spikes (gp40) were proposed to have a role in host sensing and attachment<sup>11</sup>. The biological function of the Milano collar is currently unknown. An elaborate

network of disulfide bonds in the collar suggests that they play a role in the overall robustness of Milano particles against mechanical stress (Supplementary Fig. S6), supported by the observation that this stability is lost under reducing conditions. Additionally, the Milano collar connects the sheath to the diaphragm-like structure formed by the neck 1 lumen protrusion loop (Fig. 3d), allowing for the sheath contraction signal to reach to the genome directly, instead of indirectly through the neck proteins, in preparation for genome release during the attachment to the host cell.

Finely, we note that the MCP and mCP are doubly linked: (1) by direct disulfide bonds between MCP and mCP, and (2) through linking proteins LP1 and LP2, which are bound to both MCP and mCP. This likely provides additional robustness to the phage particle.

## Methods

**Cultivation and purification of Milano particles.** Bacteriophage Milano stock was prepared using an overlay-plate method. *Agrobacterium tumefaciens* C58 cells were grown in TYC (5 g/l Bacto tryptone, 3 g/l yeast extract, 0.86 g/l CaCl<sub>2</sub> 2xH<sub>2</sub>O) to an OD<sub>600</sub> of 0.6. Motility was verified via phase contrast microscopy. Motile cells were infected with Milano phage (10<sup>6</sup> pfu) and mixed with 0.35% soft agar TYC, layered on 1.5% agar plates and incubated overnight at 30 °C until plaques formed. Phage

particles were recovered by overlaying plaque plates with 5 ml TM buffer (50 mM Tris-HCl, pH 7.4 and 10 mM MgSO<sub>4</sub>), and rocking overnight at 4 °C. Phage particles were released from cells by the addition of 1% chloroform and precipitated with 1 M NaCl and 10% PEG8000 under overnight stirring. Phage particles were purified by 10–50% OptiPrep™ linear density gradient ultracentrifugation at 200,000 × g for 2 h. Purified phages were extracted and dialyzed against TM buffer in a 10,000 MWCO Thermo Scientific Slide-A-Lyzer® dialysis cassette. Purified phage sample was titered by plaque assay.

### Cryo-electron microscopy

**Cryo-EM Sample preparation, data collection, and image pre-processing.** Holey C-flat carbon grids (1.2/1.3, 400 mesh, copper) were glow-discharged for 30 s in a GATAN Solarus Plasma Cleaner. A 3 µL phage sample in TM buffer was applied to glow-discharged grid, blotted for 3 s and plunge-frozen in liquid ethane (−180 °C) using an EM GP Plunge Freezer (Leica). Cryo-EM data was collected on a 300 keV Titan Krios with a Gatan K3 camera (University of Virginia). Total exposure per movie was ~50 e Å<sup>−2</sup> with a pixel size of ~1.08 Å. Micrographs were preprocessed by ‘patch motion correction’ and ‘patch CTF estimation’ jobs in cryoSPARC for motion correction and contrast transfer function (CTF) estimation, respectively<sup>43–45</sup>.

**Reconstruction of various neck regions.** The 3D reconstructions for various neck regions were achieved by single particle analysis (SPA) using cryoSPARC<sup>45</sup>. Neck region of Milano was manually boxed in several micrographs to generate 2D class averages for automatic template-based particle picking by ‘Template picker’. Bad picks were removed by iterative ‘2D-classification’ and ‘selection of 2D-classes’. The ab initio 3-D model was generated by ‘ab-initio reconstruction’ job. The ab initio map was further refined by imposing C3 symmetry to the highest possible resolution by ‘homogenous refinement’ job. The visual inspection of C3 map suggested different symmetries for different regions of neck. The C3 map was then further refined by imposing/relaxing symmetries (C1, C3, C5, C12 and C15) to generate the highest possible resolution maps for the whole neck assembly, neck devoid of capsid (portal, neck 1, neck 2, tail terminator, tail tube and collar), neck-capsid junction (neck 1, collar and capsid), portal only and collar only. Refinement was performed by iterative cycles of ‘homogenous refinement’, ‘nonuniform homogenous refinement’ and/or ‘local CTF refinement’. The map for the neck-tail junction was refined by ‘focused refinement’ with C3 symmetry.

**Reconstruction of capsid.** Choice and selection of Milano capsid particles was done according to a similar procedure used for the neck. Due to the large size of the capsid, the pixel size was down sampled to 1.44 Å before reconstruction. The ab initio model of capsid was generated by ‘ab-initio reconstruction’ job and refined by iterative cycles of ‘homogenous refinement’ and ‘CTF refinement’ jobs imposing icosahedral and C5 symmetries to highest possible resolutions.

The statistics of data collection and processing are detailed in Table 1. The resolution of maps and map-to-map Fourier shell correlation curves generated from cryoSPARC are provided in Supplementary Fig. S11 for all reconstructions.

**Model building.** Three-dimensional structures for all 127 proteins of Milano were predicted by AlphaFold<sup>46</sup>. The protein structure models were rigid-body fitted into 3D maps manually and by DeepTracer, as needed<sup>47</sup>. The rigid-body fit models were refined against density map by iterative cycles of automatic

refinement using PHENIX<sup>48</sup> and interactive refinement in Coot<sup>49</sup>. Model building exercise and sequence analysis revealed the error in ORF annotations in the published Milano genome<sup>22</sup>. The density for the N-terminal 56 residues of collar (Uniport# A0A482MGH3) was not present in the map, but instead appeared as a separate protein, for which density had been found in the capsid. Therefore, annotated collar ORF (Uniport# A0A482MGH3) actually starts at Met-57 and its previously annotated residues 1–38 represents another capsid protein (gp128) in a different reading frame that we named linking protein 2 (personal communication with Author of Reference<sup>22</sup>) (Supplementary Fig. S8). The statistics of model building are detailed in Table 1.

**Phage disintegration assay.** *Sonication-induced disintegration:* Bacteriophage samples ( $1 \times 10^{10}$  pfu mL<sup>−1</sup> in TM buffer) were ultra-sonicated on ice by the XL-2000 series probe ultra-sonicator (MISONIX, Farmingdale, NY) with pulse 1:1 s (on:off) in the presence or absence of 10 mM DTT. Samples at different time intervals (5, 10, 20 and 30 s) were withdrawn for phage morphology analysis by NS-TEM.

*Heat-induced disintegration:* Bacteriophage samples were incubated at 60 °C in the presence or absence of 10 mM DTT in a T-100 Thermal Cycler (BioRad). Samples at different time intervals (90, 180 and 300 s) were withdrawn for phage morphology analysis by NS-TEM.

**Negative-staining and transmission electron microscopy.** A 3 microliter phage sample was applied to plasma-cleaned carbon film grids, stained by 2% uranyl acetate, and imaged on a Tecnai T12 electron microscope with standard settings. NS-TEM micrographs were used to analyze the intactness of phage particles after sonication or heating.

**Structural and sequence analysis.** Structural and sequence similarity analysis were performed by DALI<sup>37</sup> and NCBI-BLAST (<https://blast.ncbi.nlm.nih.gov/Blast.cgi>), respectively. Structural analysis was performed and displayed by ChimeraX<sup>50</sup>. The surface electrostatic potential was calculated by APBS—Adaptive Poisson-Boltzmann Solver<sup>51</sup>.

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

### Data availability

The atomic models and three-dimensional reconstructions described in this paper are available in the Protein Data Bank (8FWB, 8FWC, 8FWE, 8FWG, 8FWM, 8FXR, 8FXP) and Electron Microscopy Data Bank (29500, 29501, 29503, 29504, 29512, 29541, 29540), respectively. All other data are available from the corresponding author on reasonable request.

Received: 26 July 2023; Accepted: 28 August 2023;

Published online: 08 September 2023

### References

- Venturini, C., Petrovic Fabijan, A., Fajardo Lubian, A., Barbirz, S. & Iredell, J. Biological foundations of successful bacteriophage therapy. *EMBO Mol. Med.* **14**, e12435 (2022).
- Eskenazi, A. et al. Combination of pre-adapted bacteriophage therapy and antibiotics for treatment of fracture-related infection due to pandrug-resistant *Klebsiella pneumoniae*. *Nat. Commun.* **13**, 302 (2022).
- Ferry, T. et al. Personalized bacteriophage therapy to treat pandrug-resistant spinal *Pseudomonas aeruginosa* infection. *Nat. Commun.* **13**, 4239 (2022).



4. Adriaenssens, E. M. et al. Taxonomy of prokaryotic viruses: 2018–2019 update from the ICTV Bacterial and Archaeal Viruses Subcommittee. *Arch. Virol.* **165**, 1253–1260 (2020).
5. Ackermann, H. W. 5500 Phages examined in the electron microscope. *Arch. Virol.* **152**, 227–243 (2007).
6. Hardy, J. M., Dunstan, R. A., Lithgow, T. & Coulibaly, F. Tall tails: cryo-electron microscopy of phage tail DNA ejection conduits. *Biochem. Soc. Trans.* **50**, 459–22W (2022).
7. Simpson, A. A. et al. Structure of the bacteriophage  $\phi$ 29 DNA packaging motor. *Nature* **408**, 745–750 (2000).
8. Li, F. et al. High-resolution cryo-EM structure of the Shigella virus Sf6 genome delivery tail machine. *Sci. Adv.* **8**, eadc9641 (2022).
9. Lhuillier, S. et al. Structure of bacteriophage SPP1 head-to-tail connection reveals mechanism for viral DNA gating. *Proc. Natl Acad. Sci.* **106**, 8507–8512 (2009).
10. Fokine, A. et al. The molecular architecture of the bacteriophage T4 neck. *J. Mol. Biol.* **425**, 1731–1744 (2013).
11. Wang, Z. et al. Structure of Vibrio phage XM1, a simple contractile DNA injection machine. *Viruses* **15**, 1673 (2023).
12. Leiman, P., Kanamaru, S., Mesyanzhinov, V., Arisaka, F. & Rossmann, M. Structure and morphogenesis of bacteriophage T4. *Cell. Mol. Life Sci. CMLS* **60**, 2356–2370 (2003).
13. Black, L. W. & Rao, V. B. Structure, assembly, and DNA packaging of the bacteriophage T4 head. *Adv. Virus Res.* **82**, 119–153 (2012).
14. Tavares, P., Zinn-Justin, S. & Orlova, E. V. Genome gating in tailed bacteriophage capsids. *Adv. Exp. Med. Biol.* **726**, 585–600 (2012).
15. Tavares, P. The Bacteriophage Head-to-Tail Interface. In *Subcellular Biochemistry* Vol. 88 (eds Harris, J. R. & Bhella, D.) 305–328 (Springer, 2018).
16. Orlov, I. et al. CryoEM structure and assembly mechanism of a bacterial virus genome gatekeeper. *Nat. Commun.* **13**, 1–12 (2022).
17. Cuervo, A. et al. Structures of T7 bacteriophage portal and tail suggest a viral DNA retention and ejection mechanism. *Nat. Commun.* **10**, 1–11 (2019).
18. Fang, Q. et al. Structural morphing in a symmetry-mismatched viral vertex. *Nat. Commun.* **11**, 1–11 (2020).
19. Wikoff, W. R. et al. Topologically linked protein rings in the bacteriophage HK97 capsid. *Science* **289**, 2129–2133 (2000).
20. Fokine, A. et al. Structural and functional similarities between the capsid proteins of bacteriophages T4 and HK97 point to a common ancestry. *Proc. Natl Acad. Sci.* **102**, 7163–7168 (2005).
21. Hua, J. et al. Capsids and genomes of jumbo-sized bacteriophages reveal the evolutionary reach of the HK97 fold. *MBio* **8**, 01579–17 (2017).
22. Nittolo, T., Ravindran, A., Gonzalez, C. F. & Ramsey, J. Complete genome sequence of agrobacterium tumefaciens myophage Milano. *Microbiol. Resour. Announc.* **8**, e00587–19 (2019).
23. Ge, P. et al. Action of a minimal contractile bactericidal nanomachine. *Nature* **580**, 658–662 (2020).
24. Olia, A. S., Prevelige, P. E., Johnson, J. E. & Cingolani, G. Three-dimensional structure of a viral genome-delivery portal vertex. *Nat. Struct. Mol. Biol.* **18**, 597–603 (2011).
25. Zinke, M. et al. Architecture of the flexible tail tube of bacteriophage SPP1. *Nat. Commun.* **11**, 5759 (2020).
26. Cardarelli, L. et al. The crystal structure of bacteriophage HK97 gp6: defining a large family of head–tail connector proteins. *J. Mol. Biol.* **395**, 754–768 (2010).
27. Helgstrand, C. et al. The refined structure of a protein catenane: the HK97 bacteriophage capsid at 3.44 Å resolution. *J. Mol. Biol.* **334**, 885–899 (2003).
28. Hardy, J. M. et al. The architecture and stabilisation of flagellotropic tailed bacteriophages. *Nat. Commun.* **11**, 3748 (2020).
29. Iwasaki, T. et al. Three-dimensional structures of bacteriophage neck subunits are shared in Podoviridae, Siphoviridae and Myoviridae. *Genes Cells* **23**, 528–536 (2018).
30. Pintilie, G., Chen, D.-H., Haase-Pettingell, C. A., King, J. A. & Chiu, W. Resolution and probabilistic models of components in CryoEM maps of mature P22 bacteriophage. *Biophys. J.* **110**, 827–839 (2016).
31. Sun, L. et al. Cryo-EM structure of the bacteriophage T4 portal protein assembly at near-atomic resolution. *Nat. Commun.* **6**, 1–11 (2015).
32. Xu, J., Wang, D., Gui, M. & Xiang, Y. Structural assembly of the tailed bacteriophage  $\phi$ 29. *Nat. Commun.* **10**, 2366 (2019).
33. Kropinski, A. M. et al. Genome and proteome analysis of 7-7-1, a flagellotropic phage infecting Agrobacterium sp H13-3. *Virol. J.* **9**, 1–8 (2012).
34. Sathaliyawala, T. et al. Functional analysis of the highly antigenic outer capsid protein, Hoc, a virus decoration protein from T4-like bacteriophages. *Mol. Microbiol.* **77**, 444–455 (2010).
35. Morais, M. C. et al. Conservation of the capsid structure in tailed dsDNA bacteriophages: the pseudoatomic structure of  $\phi$ 29. *Mol. Cell* **18**, 149–159 (2005).
36. Dedeo, C. L., Teschke, C. M. & Alexandrescu, A. T. Keeping it together: structures, functions, and applications of viral decoration proteins. *Viruses* **12**, 1163 (2020).
37. Holm, L. Dali server: structural unification of protein families. *Nucleic Acids Res.* **50**, W210–W215 (2022).
38. Wong, E. et al. The Vibrio cholerae colonization factor GbpA possesses a modular structure that governs binding to different host surfaces. *PLoS Pathog.* **8**, e1002373 (2012).
39. Hare, B. J. et al. Structure, specificity and CDR mobility of a class II restricted single-chain T-cell receptor. *Nat. Struct. Biol.* **6**, 574–581 (1999).
40. Lowe, E. C., Baslé, A., Czjzek, M., Firbank, S. J. & Bolam, D. N. A scissor blade-like closing mechanism implicated in transmembrane signaling in a Bacteroides hybrid two-component system. *Proc. Natl Acad. Sci.* **109**, 7298–7303 (2012).
41. Olia, A. S. et al. Binding-induced stabilization and assembly of the phage P22 tail accessory factor gp4. *J. Mol. Biol.* **363**, 558–576 (2006).
42. Wood, W. B. & Conley, M. P. Attachment of tail fibers in bacteriophage T4 assembly: role of the phage whiskers. *J. Mol. Biol.* **127**, 15–29 (1979).
43. Rohou, A. & Grigorieff, N. CTFFIND4: Fast and accurate defocus estimation from electron micrographs. *J. Struct. Biol.* **192**, 216–221 (2015).
44. Zheng, W. et al. Refined Cryo-EM structure of the T4 tail tube: exploring the lowest dose limit. *Structure* **25**, 1436–1441.e2 (2017).
45. Punjani, A., Rubinstein, J. L., Fleet, D. J. & Brubaker, M. A. cryoSPARC: algorithms for rapid unsupervised cryo-EM structure determination. *Nat. Methods* **14**, 290–296 (2017).
46. Mirdita, M. et al. ColabFold: making protein folding accessible to all. *Nat. Methods* **19**, 679–682 (2022).
47. Chang, L. et al. DeepTracer-ID: de novo protein identification from cryo-EM maps. *Biophys. J.* **121**, 2840–2848 (2022).
48. Afonine, P. V. et al. Real-space refinement in PHENIX for cryo-EM and crystallography. *Acta Crystallogr. Sect. D: Struct. Biol.* **74**, 531–544 (2018).
49. Emsley, P. & Cowtan, K. Coot: model-building tools for molecular graphics. *Acta Crystallogr. Sect. D: Biol. Crystallogr.* **60**, 2126–2132 (2004).
50. Goddard, T. D. et al. UCSF ChimeraX: meeting modern challenges in visualization and analysis. *Protein Sci.: A Publ. Protein Soc.* **27**, 14–25 (2018).
51. Jurrus, E. et al. Improvements to the APBS biomolecular solvation software suite. *Protein Sci.: A Publ. Protein Soc.* **27**, 112–128 (2018).

## Acknowledgements

This work was supported by NIH GM122150 (to E.H.E.). B.E.S. is supported by NSF IOS 2054392, the Dean's Discovery Fund at Virginia Tech, and N.C.E. by a Virginia Tech Center for Emerging, Zoonotic, and Arthropod-borne Pathogens mini-grant. P.G.L. is supported by NIH GM139034. Electron Microscopy was performed at the University of Virginia Molecular Electron Microscopy Core facility. We thank Carlos Gonzalez for phage Milano.

## Author contributions

R.R.S., E.H.E. and B.E.S. designed the study, while P.G.L. provided additional suggestions for experiments. The phage sample preparation was conducted by N.C.E., A.A.H., A.L.S. and R.J.K. under the guidance of B.E.S. Cryo-EM grid preparation, experiment, image-analysis, reconstructions, and model building were conducted by R.R.S., with suggestions from E.H.E., P.G.L., F.W. and M.A.B.K. The phage disintegration experiment was performed by R.R.S. with suggestions from B.E.S. and E.H.E. Manuscript and revisions were prepared by R.R.S., E.H.E., P.G.L. and B.E.S. Additional suggestions on the manuscript were provided by M.A.B.K.

## Competing interests

The authors declare no competing interests.

## Additional information

**Supplementary information** The online version contains supplementary material available at <https://doi.org/10.1038/s42003-023-05292-1>.

**Correspondence** and requests for materials should be addressed to Petr G. Leiman, Birgit E. Scharf or Edward H. Egelman.

**Peer review information** This manuscript has been previously reviewed at another Nature Portfolio journal. The manuscript was considered suitable for publication without further review at *Communications Biology*. Primary Handling Editor: Gene Chong.

**Reprints and permission information** is available at <http://www.nature.com/reprints>

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2023