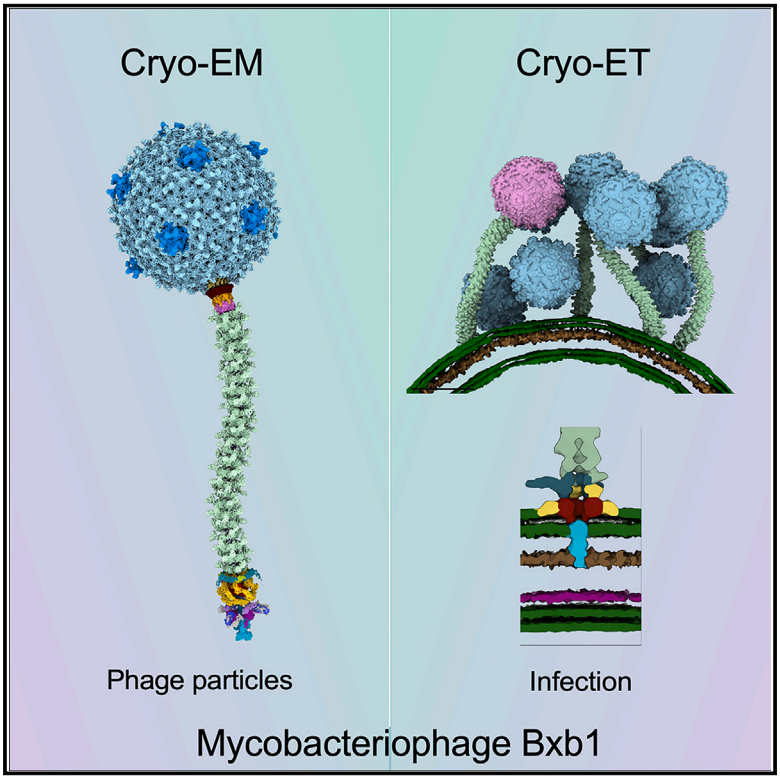
Cell

Article

Structure and infection dynamics of mycobacteriophage Bxb1

Graphical abstract



Highlights

- Atomic-level resolution of the complete mycobacteriophage Bxb1 structure
- Structural insights into unusual capsid and tail tube decorations
- Structural details of the tail-tip structure and protein interactions
- Conformational changes of the tail tip during infection of *Mycobacterium smegmatis*

Authors

Krista G. Freeman, Sudipta Mondal, Lourriel S. Macale, ..., James F. Conway, Donghyun Park, Graham F. Hatfull

Correspondence

dopark@scripps.edu (D.P.), gfh@pitt.edu (G.F.H.)

In brief

Cryo-EM and cryo-ET reveal the structural details and conformational dynamics of mycobacteriophage Bxb1 as it infects its *Mycobacterium smegmatis* host.

Freeman et al., 2025, Cell *188*, 1–18
May 29, 2025
https://doi.org/10.1016/j.cell.2025.03.027

CellPress

Cell

CellPress
OPEN ACCESS

Article

Structure and infection dynamics of mycobacteriophage Bxb1

Krista G. Freeman,[1] Sudipta Mondal,[2] Lourriel S. Macale,[3] Jennifer Podgorski,[4] Simon J. White,[4] Benjamin H. Silva,[2] Valery Ortiz,[2] Alexis Huet,[5] Ronelito J. Perez,[3] Joemark T. Narsico,[3] Meng-Chiao Ho,[3,6] Deborah Jacobs-Sera,[1] Todd L. Lowary,[3,6] James F. Conway,[5] Donghyun Park,[2,*] and Graham F. Hatfull[1,7,*]

[1]Department of Biological Sciences, University of Pittsburgh, Pittsburgh, PA, USA
[2]Department of Integrative Structural and Computational Biology, Scripps Research, La Jolla, CA, USA
[3]Institute of Biological Chemistry, Academia Sinica, Taipei 115, Taiwan
[4]Department of Molecular and Cell Biology, University of Connecticut, Storrs, CT, USA
[5]Department of Structural Biology, University of Pittsburgh School of Medicine, Pittsburgh, PA, USA
[6]Institute of Biochemical Sciences, National Taiwan University, Taipei, Taiwan
[7]Lead contact
*Correspondence: dopark@scripps.edu (D.P.), gfh@pitt.edu (G.F.H.)
https://doi.org/10.1016/j.cell.2025.03.027

SUMMARY

Mycobacteriophage Bxb1 is a well-characterized virus of *Mycobacterium smegmatis* with double-stranded DNA and a long, flexible tail. Mycobacteriophages show considerable potential as therapies for *Mycobacterium* infections, but little is known about the structural details of these phages or how they bind to and traverse the complex *Mycobacterium* cell wall. Here, we report the complete structure and atomic model of phage Bxb1, including the arrangement of immunodominant domains of both the capsid and tail tube subunits, as well as the assembly of the protein subunits in the tail-tip complex. The structure contains protein assemblies with 3-, 5-, 6-, and 12-fold symmetries, which interact to satisfy several symmetry mismatches. Cryoelectron tomography of phage particles bound to *M. smegmatis* reveals the structural transitions that occur for free phage particles to bind to the cell surface and navigate through the cell wall to enable DNA transfer into the cytoplasm.

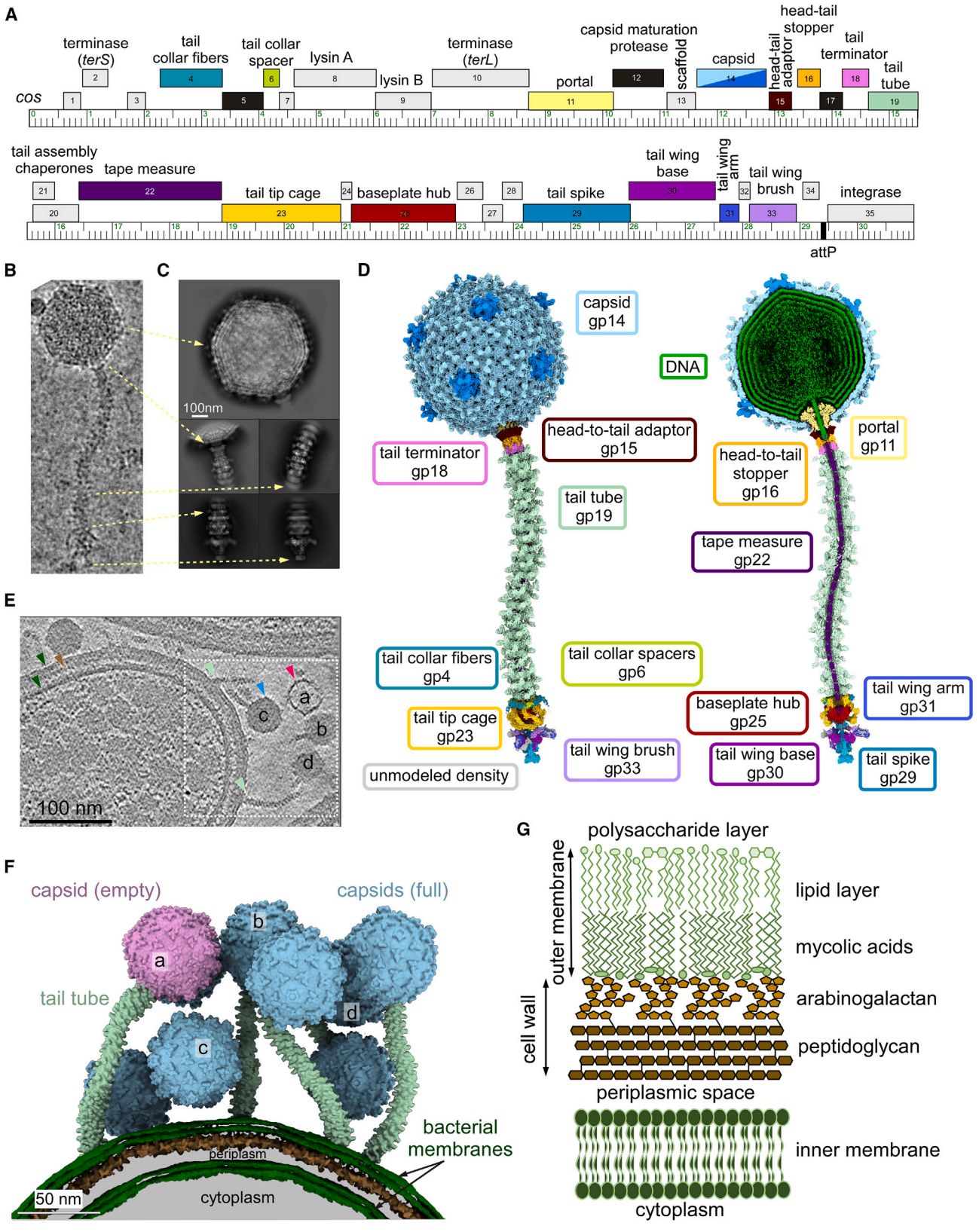
INTRODUCTION

Mycobacteria have complex cell walls that bacteriophages must navigate for both productive infection and efficient lysis.[1] These gram-positive organisms contain an unusual outer membrane ("mycomembrane") composed of mycolic acids, trehalose mycolates, and other glycolipids, separated from the cytoplasmic membrane by thick layers of peptidoglycan and arabinogalactan (see Figure 1).[2] Mycobacteriophages are viruses that infect *Mycobacterium* hosts and use specialized esterases for separation of the mycomembrane from the cell wall during lysis.[3,4] But little is known about mycobacteriophage receptors or how the virions circumvent mycobacterial cell wall barriers for entry,[5,6] although we recently showed that phages Muddy and BPs require cell wall trehalose polyphleates for infection of *Mycobacterium abscessus*.[7] Elucidation of the determinants of phage specificity and the mechanisms of DNA transfer to the cytoplasm are critical for advancing the potential of mycobacteriophages for the therapy of tuberculosis and non-tuberculous mycobacterium (NTM) infections.[8–15]

Mycobacteriophage diversity is well-defined, and the ~2,500 sequenced genomes can be grouped according to their genetic relationships into over 40 genome types (clusters and singletons).[1,16] All mycobacteriophages discovered to date are double-stranded DNA (dsDNA) tailed phages, and morphologically most are siphophages with long flexible non-contractile tails. Two genomic groups have myophage morphologies with contractile tails.[1] No *Mycobacterium* podophages with short stubby tails have been identified, presumably because these are geometrically incapable of traversing the thick mycobacterial cell wall.[1] Although there is great sequence diversity among these phages, the virion structure and assembly genes of the siphophages are highly syntenic.[1]

Advances in cryoelectron microscopy (cryo-EM) have enabled whole phage structural characterization at a near atomic level, and virions of several other hosts have been described, including podophages (HRP29[17] and Sf6[18] of Shigella, P22 of *Salmonella*,[19] PSCSP1u of *Cyanobacteria*,[20] SU10 of *Escherichia*,[21] ϕ29 of *Bacillus*,[22,23] and ΦM1 of *Pectobactia*[24]), myophages (Pam3 of *Cyanobacteria*,[25] XM1 of *Vibrio*,[26] and E217 of *Pseudomonas*[27]), and siphophages (χ of *Salmonella*[28,29] and DT57C,[30] T5[31] and λ[32,33] of *Eschericia*, vB_DshS-R4C, of *Dinoroseobacter*,[34] and 80α of *Staphylococcus*[35]). Likewise, developments in cryoelectron tomography (cryo-ET) have enabled subtomogram averaging and direct 3D visualization of host-bound phages.[19,36] Although recent cryo-EM studies have provided

Please cite this article in press as: Freeman et al., Structure and infection dynamics of mycobacteriophage Bxb1, Cell (2025), https://doi.org/10.1016/j.cell.2025.03.027



(legend on next page)

insights into the structure and evolution of mycobacteriophage capsids,[37–39] the tails and the tail tips—which play critical roles in host recognition, binding, and DNA transfer—have only been characterized at low resolution for one phage.[40] Moreover, no cryo-ET studies of mycobacteriophages have been described.

Bxb1 is a well-characterized cluster A1 mycobacteriophage[41,42] with a ~60 nm diameter isometric icosahedral capsid containing a 50.5 kbp genome, and a ~130 nm-long flexible tail. Genome analyses showed that Bxb1 virions contain unusual sequence-related ~90 amino acid C-terminal extensions (CTEs) on both the capsid and tail tube proteins and another three related copies of the sequence within a putative minor tail protein.[41,42] Bxb1 has been widely exploited as a source of biotechnological tools[1,43,44] and has been proposed as a vaccine platform in which large numbers of epitopes can be displayed on the capsid surface.[45] Here, we report the complete structure of both free Bxb1 virions (cryo-EM) and Bxb1 particles infecting *Mycobacterium smegmatis* (cryo-ET of cryofocused ion beam [cryo-FIB] milled lamellae). These structures give insights into virus architecture and assembly, the potential for Bxb1 as a vaccine platform, and the dynamic structural rearrangements associated with binding to host cells and DNA transfer. These findings will also advance mycobacteriophage therapies for tuberculosis and NTM infections.[14,15]

RESULTS

Overview of the virion structure of mycobacteriophage Bxb1

Bxb1 was one of the first mycobacteriophages to be genomically characterized,[42] and its virion structure and assembly genes constitute the genome left arm (Figure 1A). Like all ~800 sequenced cluster A mycobacteriophages,[46] the gene organization is unusual among siphophages in that some of the tail genes (and the lysis cassette) are situated in the extreme left arm of the genome, over 10 kbp away from the other tail genes (Figure 1A). We used mass spectrometry of highly purified phage particles to identify 18 of the 34 left arm gene products as virion components (Figure 1A; Table 1). We then used cryo-EM and single particle analysis (SPA) to determine the high-resolution structure of free phage particles (Figures 1B–1D and S1; Table S1). Finally, we employed cryo-FIB milling and cryo-ET, followed by subtomogram averaging (STA), to elucidate Bxb1 interactions with the *Mycobacterium* cell envelope (Figures 1E–1G and S2; Video S1; Table S2). These methods facilitate visualization of the structural dynamics of mycobacteriophage infection.

Atomic modeling into experimental SPA maps provided structural evidence for 15 of the 18 proteins identified by mass spectrometry, with nearly all of the density accounted for. Three proteins (gp12, gp5, and gp17) were not resolvable in the density map (Figure 1A; Table 1). The structure reveals interesting features of the Bxb1 virion (Figures 1B–1D), including the unusual arrangement of CTE domains displayed on both the capsid (gp14) and tail tube (gp19) surfaces. It also details portal (gp11) structural variations that accommodate symmetry mismatches in the capsid-tail connector region, which also contains the adaptor (gp15), stopper (gp16), and tail terminator (gp18). The tail-tip structure involves nine proteins, with some conserved structural features like the tape measure protein (TMP) (gp22), baseplate hub protein (gp25), and collar with fibers (gp4 and gp6), but also structures unique to Bxb1 like the tail-tip cage (gp23), spike protein (gp29), and wing complexes formed by gp30, 31, and 33 (Figure 1D).

To elucidate Bxb1's interactions with the *M. smegmatis* cell envelope, we produced lamellae with cryo-FIB milling and subjected them to cryo-ET followed by STA (Figures 1E and 1F; Table S2; Video S1). *M. smegmatis* cells were infected with Bxb1 at high (300) multiplicity of infection (MOI) to maximize the number of bound phage particles for analysis, and the mixture was vitrified 30 min after mixing phages with cells (a sufficient time for efficient adsorption but early in the lytic cycle[5]). We identified and extracted 3D sub-volumes of capsids, tail tubes, and tail tips interacting with the host membrane for subtomogram analysis. The maps reflect structures of both DNA-filled and empty host-bound phage particles (Figure 1F) and provide insights into the interaction of the Bxb1 tail tip with the mycobacterial cell envelope (Figure 1G).

These structures are discussed in detail in the sections below, while workflows and specifications for cryo-EM (Figure S1; Table S1) and cryo-ET (Figure S2; Table S2; Video S1) are provided in the supplemental information.

The Bxb1 capsid

The Bxb1 capsid was reconstructed with SPA with icosahedral symmetry (2.71 Å resolution), and a model of the capsid protein was built (Figures 2A and 2B). The Bxb1 capsid subunit adopts the HK97-fold, including the N-arm, P domain, E-loop, A domain, and G-loop (Figures 2A and 2C).[47] However, the structure

Figure 1. Mycobacteriophage Bxb1

(A) The left arm of the Bxb1 genome is represented by a white ruler (tick marks indicate every 100 bp and 1 kbp) with genes shown above as boxes. Individual genes are color-coded to structural components (the capsid protein is shown in light blue for hexon subunits and darker blue for penton subunits). Genes detected by mass spectrometry but not in the cryo-EM density are black, and genes whose products are absent in the virion are gray.

(B) Cryo-EM micrograph of a Bxb1 particle.

(C) 2D class averages of the Bxb1 capsid, portal complex, tail tube, tail tip, and tail spike components.

(D) A composite cryo-EM map of the entire Bxb1 particle, with density segments for each structural component labeled and colored as in (A). The left image shows the full volume, and the right is clipped to show a cross section.

(E) A representative tomographic slice showing an *M. smegmatis* cell infected with Bxb1. Mycobacterial membrane (green arrowhead), peptidoglycan/arabinogalactan layer (brown), DNA-containing capsids (blue), empty capsids (pink), and tail tube (light green) are indicated, and capsids are lettered.

(F) A 3D rendering of the boxed area in (E) (rotated 90° counterclockwise) displaying Bxb1 in both DNA-filled (teal) and empty capsid (pink) states. The model is rendered by mapping the subtomogram averages onto the tomogram (Video S1).

(G) Schematic representation of the *Mycobacterium* cell envelope.

See also Figures S1 and S2.

Please cite this article in press as: Freeman et al., Structure and infection dynamics of mycobacteriophage Bxb1, Cell (2025), https://doi.org/10.1016/j.cell.2025.03.027

Table 1. Virion proteins of Bxb1

Gene	Protein name	aa[a]	# spectra[b]	Modeled?	Interacts with[c]	# copies per phage[d]	Symmetry
4	tail collar fibers	356	183	yes	4, 6, 23	9	C3
5	unknown	236	23	no	N/A	N/A	N/A
6	tail collar spacer	86	27	yes	4, 23	3	C3
11	portal	488	800	yes	11, 14, 15	12	C12/C1
12	capsid mat. protease	294	75	no	N/A	N/A	N/A
14	capsid	397	1679	yes	11, 14	415	I/C1
15	head-to-tail adaptor	125	110	yes	11, 15, 16	12	C12
16	head-to-tail stopper	126	171	yes	15, 16, 18	6	C6
17	unknown	131	20	no	N/A	N/A	N/A
18	tail terminator	148	72	yes	16, 18, 19	6	C6
19	tail tube	283	809	yes	18, 19, 23	156[e]	C1
22	tape measure	823	405	yes	25	3[e]	C3
23	tail tip cage/Dit	685	285	yes	4, 6, 19, 23, 25	6	C3
25	baseplate hub	600	170	yes	22, 23, 25, 29	3	C3
29	tail spike	617	132	yes	25, 29, 30	3	C3
30	tail wing base	496	161	yes	29, 31, 33	3	C3
31	tail wing arm	106	86	yes	30, 31	6	C3
33	tail wing brush	267	71	yes	30	3	C3

[a]The number of amino acids in each protein subunit.
[b]The total number of spectra identified by LC-MS/MS are indicated.
[c]Other Bxb1 gene products that this protein subunit interacts with, as determined by ChimeraX interface mapping.
[d]The number of protein subunits per virion particle are shown for each protein; N/A, not available.
[e]The numbers of tail tube subunits are estimated from the approximate length of the tail tube and ring-to-ring distance, and the number of tape measure subunits is based on the fragments built into the cryo-EM density map.

deviates from the HK97-fold in two important ways: the P-loop and CTE domain. The P-loop is a 14-amino acid insertion in the P domain that forms trimeric "turrets" at the 3-fold axes (Figures 2A and 2C).[38] At these axes, each P-loop trimer interacts with three rings formed by hexameric E-loops to form a crosslinked capsid chainmail structure, with isopeptide bonding between asparagine-278 and lysine-74 in neighboring capsid subunits (Figure 2D).[48] The Bxb1 capsid CTE domains are displayed on the capsid external surface (Figures 2A–2C) and interact as three pairs atop capsid hexamers and as two pairs with a single monomeric domain at the pentamers. Resolving the CTE domain arrangement on pentons required an asymmetric local reconstruction (3.17 Å resolution) since icosahedral averaging smears the two dimer, one monomer configuration (Figures 2C and S3A). The CTE domains are highly immunogenic[45] and are positioned analogously to Ig-like domains on the capsid of *E. coli* phages ϕ29, P22, and Sf6, although the fold is structurally distinct from these (Figure S3B).[22,49–52] Instead, the CTE folds similarly to the C-terminal domain of a minor capsid protein of mycobacteriophage Patience (PDB: 8GIU-1, root-mean-square deviation [RMSD] 0.940 Å across 59 pruned atom pairs),[37] which is otherwise unrelated to Bxb1 (Figure S3C). Although the folds are structurally similar and both are on the capsid external surface, the Patience domain occupies the 3-fold vertices of its capsid (Figure S3D).[37]

STA maps (11.7 and 12.6 Å resolution for DNA-filled and empty capsids, respectively) show some minor differences in the Bxb1 capsids when DNA-filled versus empty (Figures 2E–2H). In particular, the CTE domains on top of capsid hexamers change position slightly in the empty capsids, with ~12 Å larger distances between adjacent pairs (Figures 2F and 2G). Additionally, the 6 Å pore formed at the center of the hexameric A domains in the capsid floor is somewhat larger in the DNA-emptied capsid reconstructions, suggesting potential changes in capsid geometry following the genome ejection (Figure 2H).

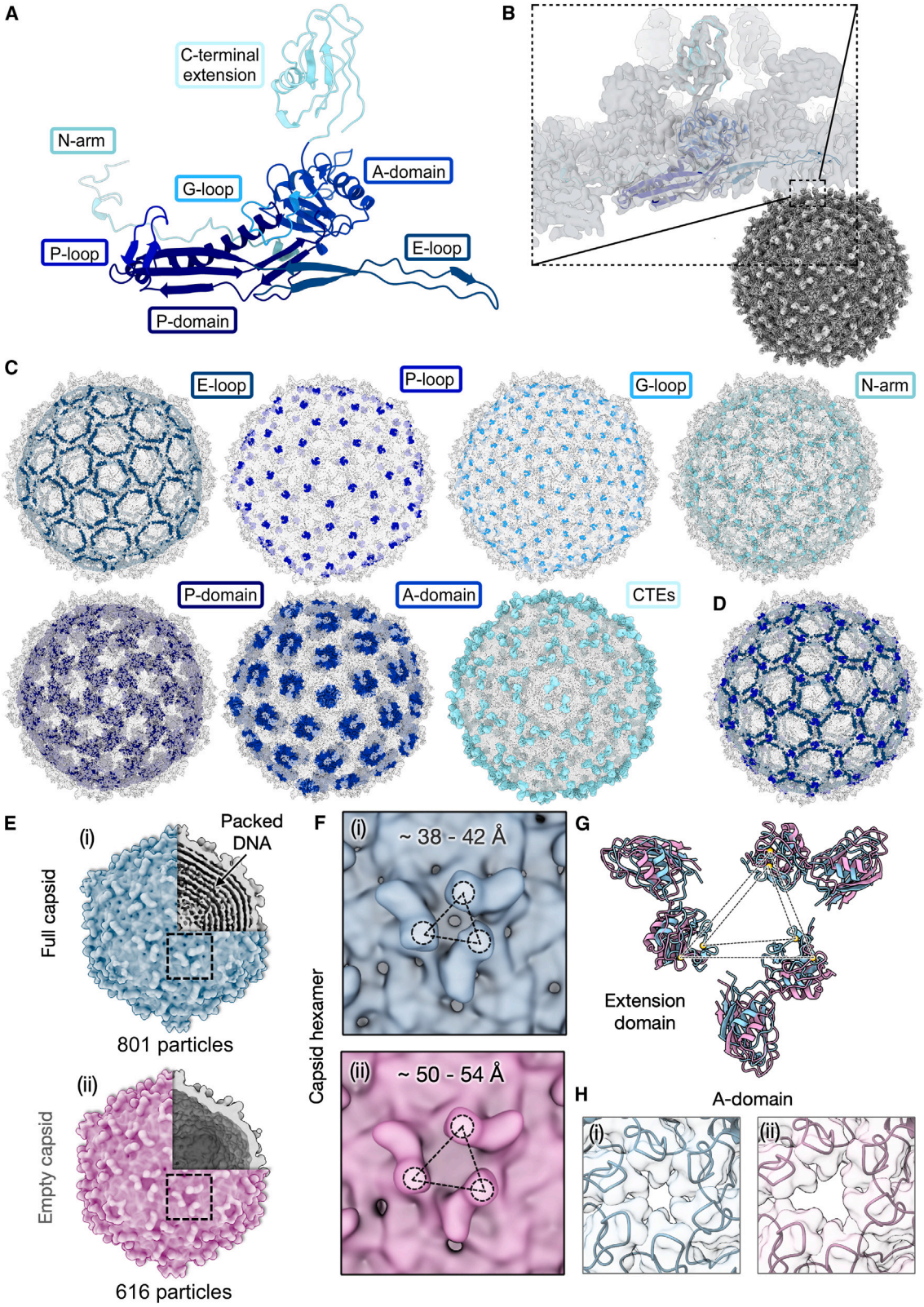
The portal and head-to-tail joining assembly

The head-to-tail region has several symmetry mismatches, most notably the association at one of the 5-fold capsid vertices of the 12-fold portal (gp11) and adaptor (gp15) complex. Below this, the adaptor protein links to the 6-fold head-to-tail stopper (gp16), which binds to the 6-fold tail terminator (gp18) that caps the top of the long hexameric but 3-fold symmetric tail tube protein (gp19). Some poorly resolved density in this region could correspond to gp17, which is particle associated (Table 1) but otherwise unmapped (Figure S4A). The distantly related p143 of phage T5 was reported to be similarly challenging to map to cryo-EM density.[31,53] The head-to-tail joining structures were resolved by symmetry-expanding the SPA dataset from the icosahedral capsid reconstruction and using two 3D classification steps to first identify the 5-fold vertex with connector density and then to isolate one of the five symmetry-related copies of each particle (Figures 3A and S1). Two approaches were taken to achieve the second step.

First, the portal vertices were classified to identify poses with well-aligned capsid and portal densities (Figure S1). We

Please cite this article in press as: Freeman et al., Structure and infection dynamics of mycobacteriophage Bxb1, Cell (2025), https://doi.org/10.1016/j.cell.2025.03.027

CellPress
OPEN ACCESS



(legend on next page)

Please cite this article in press as: Freeman et al., Structure and infection dynamics of mycobacteriophage Bxb1, Cell (2025), https://doi.org/10.1016/j.cell.2025.03.027

extracted subparticles with these orientations and locally refined both the entire map (3.44 Å resolution) and the isolated portal region (3.31 Å resolution) without symmetry (Figures 3B and S1C). This region has five capsid hexons surrounding the dodecameric portal complex, and modeling shows that the portal subunits interact with the ten closest capsid subunits, which are in either "surrounding" or "distal" orientations, as described for other phages (Figures 3B and S5A).[18,54] Two features of the model enable the 5- to 12-fold symmetry transition. First, two portal subunits on opposite sides of the complex form smaller interfaces with single capsid subunits relative to the remaining ten (Figure 3B). This is reminiscent of the 5- to 12-fold symmetry transition in HK97 prohead I, which is accommodated by ten portal subunits interacting with ten capsid scaffold domains, leaving two portal subunits free.[55] Second, each portal subunit exists in a different local capsid environment, necessitating subtle breaking of the portal's 12-fold symmetry (Figures 3C and S5A–S5C). For example, five of the Bxb1 portal subunits (subunits 2, 6, 7, 8, and 9, Figure 3C) form hydrogen bonds with a single capsid subunit (three surrounding capsid subunits and two distal subunits are partners in these interactions), and the other seven bond to two capsid subunits (one surrounding and one distal). Of those portal subunits that hydrogen bond to two capsid subunits, three (subunits 1, 4, and 11) bind to capsid subunits within the same hexamer while the other four (3, 5, 10, and 12) bridge two adjacent capsid hexamers (Figures 3C and S5B). The remodeling required to accommodate these different binding patterns occurs in the loops connecting beta strands in the wing domain (Figure S5C; Video S2), similar to gp20 of phage T4.[54]

Second, the portal vertices were classified with a focus mask over the area below the portal to resolve the entire connector complex (Figure S1). The entire portal-adaptor-stopper-terminator-tail tube assembly was locally refined with C3 symmetry (3.30 Å resolution), and then each symmetric piece was refined (resolutions of 2.69 Å for C12 region, 2.68 Å for C6, and 3.06 Å for C3) before these were merged to form a composite map (Figures 3D and S1; Table S1). A complete model was built into the composite map and shows an intricate network of interfaces (Figure 3E), with individual protein models largely homologous to similar previously described structures (Figures S5D–S5K; Table S3).[32,38,56,57]

Cryo-ET STA comparisons of the head-tail connection in particles that are host attached but have either DNA-filled or empty capsids reveal several differences (Figures 3G and 3H). About 75% of the particles have DNA-filled capsids and have a portal structure similar to that in the free virions, although the density presumed to be the top of the TMP is absent and replaced by a cylinder of new density that may correspond to dsDNA (Figures 3G and 3H). The other particles have empty capsids and have neither TMP nor dsDNA within the hollow core of the connector complex (Figures 3G and 3H). We presume these particles have completed the process of DNA release.

The flexible tail tube

The bottom of the connector complex transitions into the first of approximately 26 layers of the tail tube ring. The tail tube is highly flexible, and our SPA asymmetric reconstructions (resolutions of 3.11, 2.99, and 4.88 Å for consensus, local core, and local CTE domain maps, respectively) focused on resolving two adjoining layers at relative angles (Figure 4A). Each ring is formed by six subunits of the tail tube protein (gp19), which has conserved elements of other siphophage tail tubes, including a β-sandwich core and three loop domains: the outer loop, buried loop, and hinge loop (Figure 4B).[58] The buried loop aids in the formation of the ring by binding to adjacent subunits, while the hinge loop sticks out of the bottom to bind the lower adjoining ring (Figure 4C). The hinge loop can extend to accommodate adjoining rings at a relative tilt, mediating the tail tube's impressive flexibility (Figure 4D).[59,60] The outer loop is displayed on the external surface of the tail tube, as is the CTE domain that has sequence similarity to the capsid CTE domains (Figure S3E). The tail tube CTE domains interact in pairs reminiscent of the capsid hexon CTEs, with three interacting pairs in each tail ring giving the structure a 3-fold, rather than the typical 6-fold, symmetry (Figures 4A and 4C). Like on the capsid, the tail CTE domains can be readily removed without loss of phage viability, and their function is unclear.[45]

Within the ~40 Å central aperture of the tail tube is a hollow cylindrical density ~22 Å in diameter that presumably is the TMP (Figure 4A). However, apart from the C-terminal 22 residues situated in the tail tip (see below), the TMP cannot be readily modeled to this density due to the need for averaging from different parts of the tails in the reconstruction, as with other siphophage structures.[30,31]

Direct visualization of host-bound phages in tomographic slices revealed a darker central density in the tail tube of DNA-filled virions, compared with the hollow tail tube density in empty virions (Figure 4E). To compare changes in the tail tube at different stages of genome ejection, we performed STA on tail tube segments from cell-bound phages with genome-filled and emptied capsids (7 and 22 Å resolution, respectively; Figures 4F and

Figure 2. Bxb1 capsid structure

(A) The protein fold of Bxb1's capsid subunit, colored and labeled by domain.

(B) A single subunit of gp14 in the capsid density map, with a magnified view in the upper left showing the model (oriented and colored as in A) in a transparent density slab.

(C) Cryo-EM density maps of the capsid, with each panel showing the density contribution from a single domain in solid color with the remaining capsid density in transparent gray.

(D) E-loop and P-loop domains are shown as in (C) to illustrate the chainmail linkage of capsomeres.

(E) Subtomogram average maps of the capsids in pre-genome (teal) and post-genome (pink) translocation states.

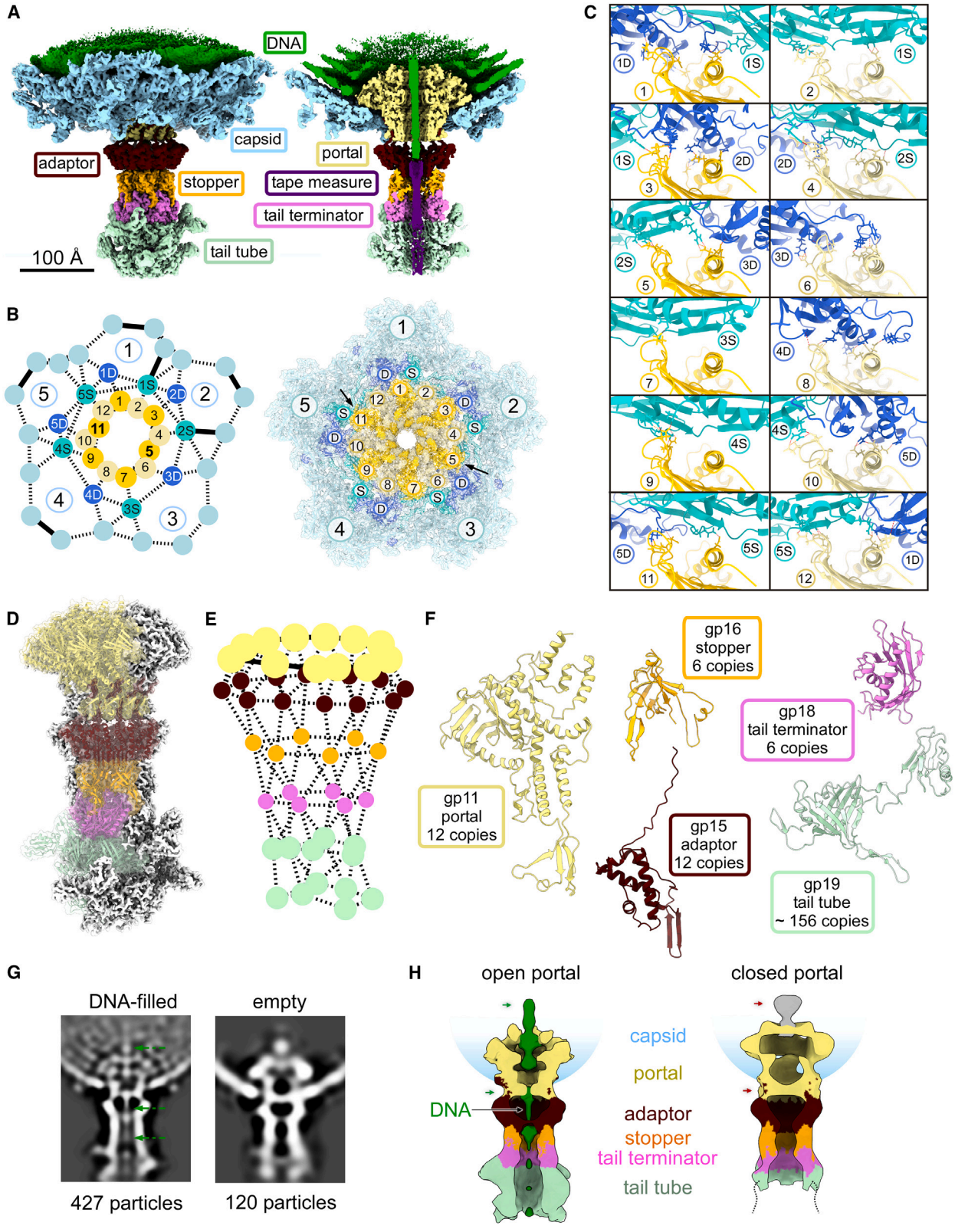
(F) Zoom-in views show the conformational rearrangement of the A-domain post-genome translocation.

(G and H) Rigid-body fitting of extension domain and A-domain models in the subtomogram average maps reveals a relaxed organization of the capsid post-genome translocation.

See also Figures S1, S2, S3, and S7.

Please cite this article in press as: Freeman et al., Structure and infection dynamics of mycobacteriophage Bxb1, Cell (2025), https://doi.org/10.1016/j.cell.2025.03.027

CellPress
OPEN ACCESS



(legend on next page)

Please cite this article in press as: Freeman et al., Structure and infection dynamics of mycobacteriophage Bxb1, Cell (2025), https://doi.org/10.1016/j.cell.2025.03.027

4G). STA revealed a 20 Å diameter solid cylinder of density in the tail tubes of infection state (host-attached, genome-filled) phages. This structure is slightly narrower than the hollow cylinder TMP density (22 Å diameter) observed in the pre-infection SPA structure (Figures 4F and 4G) and likely represents dsDNA. By contrast, STA of post-infection (host attached with empty capsids) phages showed a hollow tail tube, indicating that the DNA ejection process had been completed. Superimposing these two STA maps onto the SPA map of free phages reveals that the dimensions and torsion of the tail tube remain unchanged throughout the genome ejection process. Notably, neither STA structure of host-attached phages exhibited a density resembling the TMP in the pre-infection SPA map, indicating the absence of ordered TMP density after host binding (Figures 4F and 4G).

The virion tail-tip complex

At the end of the long, flexible tail tube is the tail-tip complex, for which we used a consensus C3 SPA reconstruction (2.45 Å resolution) and local C1 refinements (resolutions of 2.85, 3.98, and 3.65 Å for tail spike, wings, and collar/fibers, respectively) to generate a composite map with upper and lower regions (Figure 5A). Building a model into this map revealed eight separate minor tail proteins encoded by Bxb1, in addition to the tail tube and tape measure proteins (Figures 5B–5D and S6). Each protein is detailed by domain and compared with its closest structural match when such exists (Figure S6; Table S3). In the overall structure, the tail tube is at the top of the complex and binds to the tail-tip cage (gp23), which is surrounded at the top by a collar and fibers made of gp4 and gp6. The extreme C terminus of the TMP (gp22) is resolvable within the tail tube in this area and binds to the top of the baseplate hub formed by gp25. The hub is surrounded by the lower, spiraling arms of the tail-tip cage that lead to the tail wing structures comprised of the brush (gp33), base (gp31), and arm (gp32). At the center of the wings and connected to the bottom of the baseplate hub is the tail spike protein (gp29). These proteins form a complex network of interactions, with most protein subunits in the tail tip interacting with 2–4 other proteins (Figure 5E; Table 1).

The upper region of the hexameric tail-tip cage is formed by the gp23 N-terminal domains and is structurally similar to other distal tail (Dit) proteins (Figures S6A and S6B; Table S3).[23] The collar that surrounds the Dit-like structure has nine subunits of the tail collar fibers (gp4) separated into trimers by three subunits of the tail collar spacers (gp6) (Figures 5D and S6C–S6F). Only a small N-terminal part of the gp4 tail fiber is resolved in the density maps, and the remaining model is predicted by AlphaFold3 to form ~30 nm-long trimeric fibers (Figure S4B). AlphaFold3 also suggests that the N terminus of gp5 may associate with the unmodeled gp4 C-terminal ends, but gp5 appears too large to fit into the region of unmodeled density in the tail-tip wings where such a complex would likely reside (Figures S4E and S4F). The extreme C termini of the trimeric gp4 fibers may instead interact with the unmodeled N termini of the gp33 wing brushes, and this model fits well into the unmodeled density (Figures S4C and S4D).[61]

The lower region of the gp23 cage protein contains three globular domains with structure and sequence similarity to the capsid and tail tube CTE domains (Figures S3E, S3F, and S6B).[41] These three domains do not dimerize with neighboring subunits as seen with the capsid and tail tube CTEs but instead form spiraling arms wrapping around the baseplate hub, gp25 (Figures 5A and 5B). The interactions between gp23 and gp25 are intimate: (1) the top of the hub complex hydrogen bonds to all six gp23 Dit-like domains, (2) helical arms of gp25 protrude through holes in the gp23 upper cage structure, and (3) the middle of the hub bonds to the inside surfaces of three gp23 subunits (Figures 5A, 5B, and S6H). Within the core of gp25 is an alpha-helical bundle that interacts with the extreme C termini of three chains of the TMP (gp22), although only 22 residues at the extreme C terminus are well resolved (Figures 5A, 5B, 5D, 5E, and S6H).

The lower end of the gp25 baseplate hub connects to the lower tail-tip assembly, which comprises the tail spike protein, gp29, and the wing complex made up of the tail wing base (gp30), arm (gp31), and brush (gp33) (Figures 5A–5D and S6I–S6O). The spike protein (gp29) forms an elongated structure with N-terminal helices tightly bound to the lower region of the baseplate hub and extending to form a bulb domain at the very end of the tail-tip complex (Figures 5B, S6I, and S6J). The C-terminal part of gp29 forms a brush domain and a short beta loop that interacts intimately with the N-terminal domain of the tail wing base, gp30 (Figures 5B and S6J). The gp30 tail wing base is a particularly intriguing part of the tail-tip complex. Its N-terminal

Figure 3. Bxb1 portal complex structure

(A) A composite cryo-EM map of Bxb1's capsid, connector region, and tail tube shown in full volume (left) and clipped (right), colored and labeled as in Figure 1.
(B) An interface network (left) and the C1 cryo-EM map and model of the capsid-portal complex (right), shown from inside the capsid. The map is segmented and shown as transparent surfaces. Portal subunits 1–12 and portal-adjacent hexons 1–5 are numbered and colored in shades of gold and blue, respectively. Capsid subunits are labeled as either "S" (surrounding) or "D" (distal). Arrows point to portal subunits 5 and 11, which have smaller interface areas.
(C) Individual portal subunits and their hydrogen-bonded capsid subunits were extracted and aligned. The hydrogen-bonding residues of each complex are shown individually and labeled consistently with (B).
(D) A composite cryo-EM map of the connector region, with one-third of the map removed and replaced by the model shown as a cartoon with transparent surface.
(E) Interface network of the connector region, with color-coded circles showing each protein subunit and lines connecting subunits with interface areas ≥ 300 Å^2.
(F) Protein subunits for each component of the connector region, identified by gene number, structural annotation, and number of copies in the complete particle.
(G) 2D slices of subtomogram averages of the capsid-tail tube connector in the pre- (left) and post- (right) genome translocation states. Green arrows indicate putative DNA density.
(H) Cross-sectional 3D renderings of subtomogram average maps are shown in (G). Red arrows point to the closed channel in the post-genome translocation state map.
See also Figures S1, S2, S4, and S5.

Please cite this article in press as: Freeman et al., Structure and infection dynamics of mycobacteriophage Bxb1, Cell (2025), https://doi.org/10.1016/j.cell.2025.03.027

CellPress
OPEN ACCESS

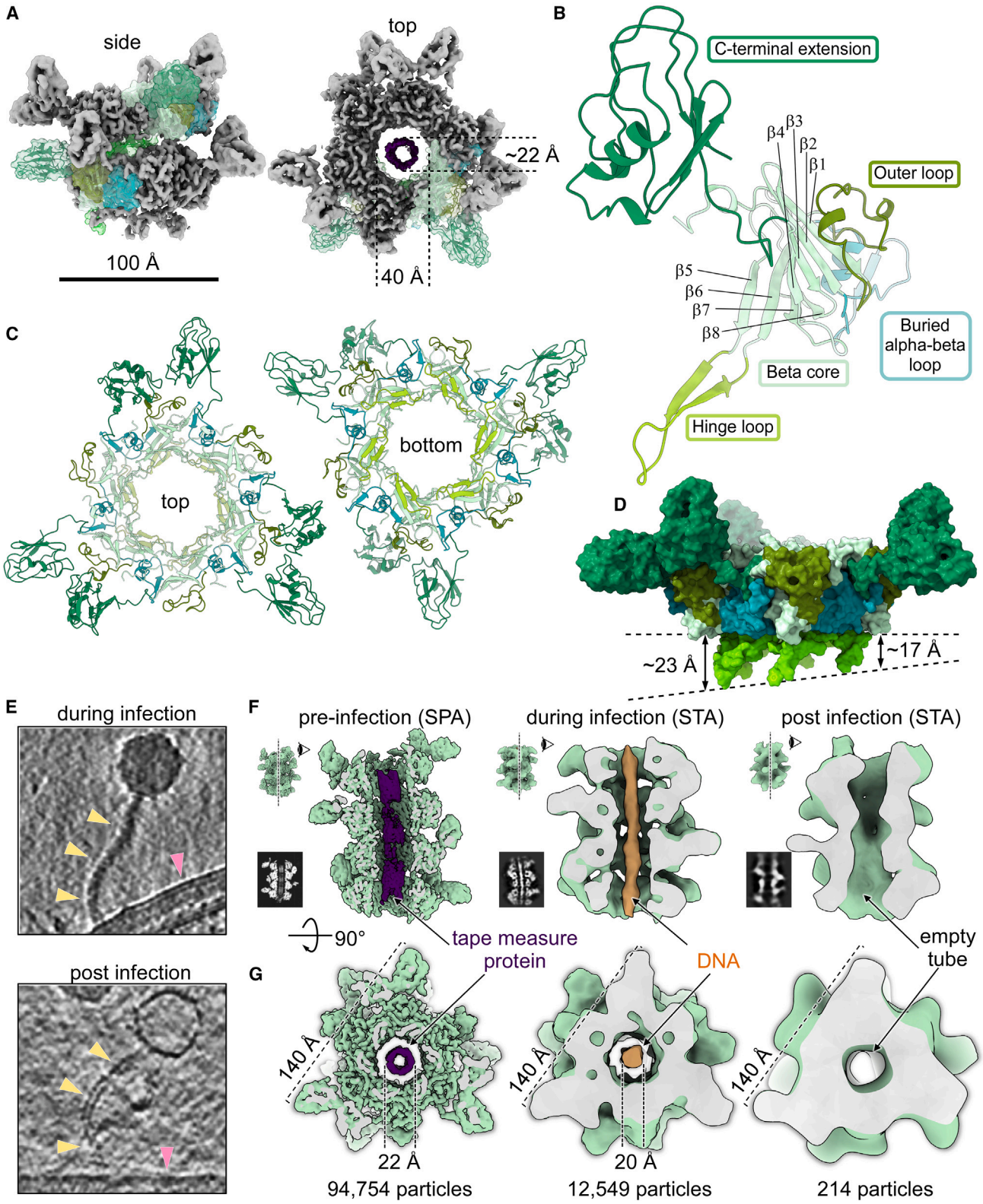


Figure 4. Bxb1 tail tube structure

(A) A C1 reconstruction of a segment of the Bxb1 tail tube is shown in gray from the side (left) and top (right), with the TMP density colored purple. One subunit from each tail tube ring has been subtracted from the density map and replaced by models shown as cartoons with transparent surfaces.

(legend continued on next page)

domain, which binds to the tail spike protein, has structural similarity to glycan-binding FN3 domains (Figures S6K and S6L). The C-terminal 42 residues of gp30 form an extended arm interacting closely with a homodimer of the gp31 tail wing arm protein (Figures 5A, 5E, and S6L). Between these N- and C-terminal domains, gp30 contains a predicted D-ala-D-ala carboxypeptidase catalytic domain, which presumably acts on a peptidoglycan target (Figures S6K and S6L). The remaining wing components are similarly interesting, with gp33 resembling cyanophage Pam3 fibers[25] and gp31 being structurally distinct from known protein structures (Figures S6M–S6O; Table S3).

Dynamic rearrangements of the tail-tip complex during infection

Prior cryo-EM and cryo-ET studies have described the two-membrane organization of the *Mycobacterium* envelope.[62,63] Our cryo-ET STA of the envelope structure from cryo-FIB milled cells shows clearly defined inner and outer leaflets of both the cytoplasmic membrane and the outer mycomembrane with unprecedented detail (Figures 6A and 6B). Between these are several layers of density, including an inner dense layer referred to as the granular layer,[62] which is likely composed of membrane-anchored phosphatidylinositol mannosides.[2,62] The remaining layers likely correspond to the peptidoglycan-arabinogalactan complex (Figures 1G and 6B). The overall dimension of the wall (38 nm) is similar to that reported previously (42.4 nm).[62] The distance from the outside of the cell to the outer layer of the cytoplasmic membrane—to which the phage tail tip likely needs to interact—is about 31.7 nm. Without considering further rearrangements, the full phage tail-tip structure would extend about 27.5 nm (from the bottom of the tail tube to the bottom of the tail spike; Figure 5A), positioning the bottom of the spike protein near but not at the inner membrane.

The cryo-ET analysis of cell-bound Bxb1 particles reveals striking conformational changes in the tail tip, which was reconstructed at ~12 Å resolution with STA (Figure 6C). Clearly recognizable above the cell surface is the CTE-studded tail tube, the tail collar and fibers, and the Dit-like domains of the tail-tip cage (Figure 6C). An overlay of the collar density from the cryo-EM and cryo-ET maps showed good agreement, although the fibers pivot 45° upward in the infection state, becoming parallel to the cell surface (Figure 6C). The top of the baseplate hub and the lower portion of the tail-tip cage are clearly discernible immediately adjacent to or even within the mycobacterial outer membrane (Figures 6C and 6D). The infection-state tail-tip structure has an empty core absent of TMP density or the lower, closed portion of the baseplate hub seen in the cryo-EM map (Figures 6D and 6E). The lower tail-tip cage remarkably flattens into a spiraling disk, with the lower portion lifting up and out to occupy the plane of the outer membrane (Figures 6C–6E; Video S3). We note that, while these averaged sub-tomograms provide exceptional views of the phage tail tip embedded in the mycobacterial cell wall, the resolution is insufficient for building atomic models. Nonetheless, a feasible model was generated by rigid-body fitting gp23 domains into the subtomogram averages, with the CTE domains rationally oriented to interface positively charged surfaces against the outer membrane (Figures 6F and 6G).

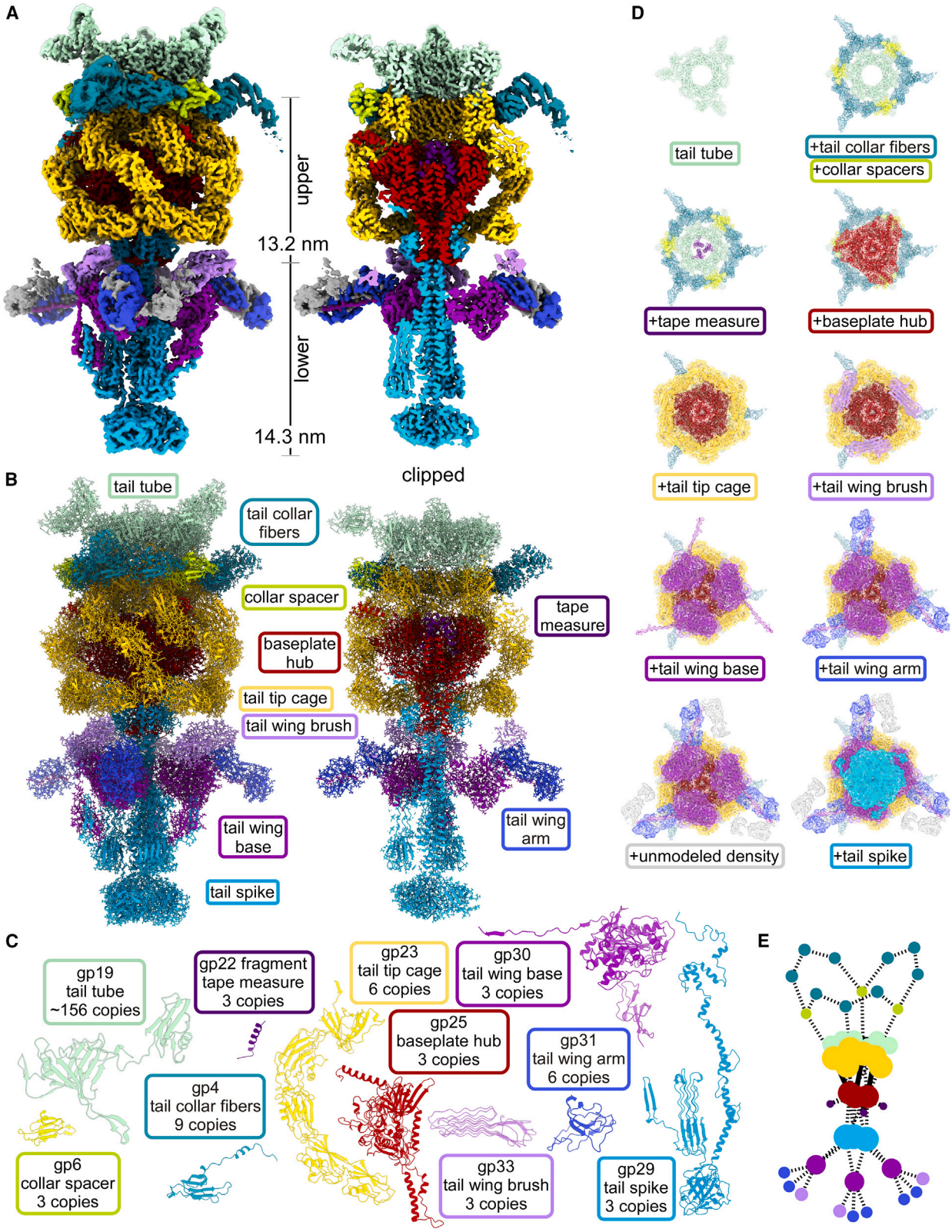
We did not observe clear evidence for the tail spike or wing structures within the cell wall, possibly because of low occupancy (and release upon attachment) or disorder due to flexibility or conformational heterogeneity. To investigate this, we performed 3D classification and identified a subset of particles with density below the baseplate hub within the cell wall. This subset was subjected to focused refinement to yield a ~33 Å map with a weak conical density directly underneath the tail assembly (Figure 6H). This density crosses the outer membrane and penetrates into and possibly through the peptidoglycan layer of the cell wall (Figure 6H). Although this structure is low resolution, it likely represents the tail spike and wing proteins after dynamic rearrangements to facilitate the translocation of DNA from the phage into the host cytoplasm. This structure is consistent with the tail wing components at or close to the peptidoglycan-arabinogalactan layer (Figures 5B and 6H), supporting the proposed D-ala-D-ala-carboxypeptidase activity of gp30 (Figure S6L), the tail wing base, which may cleave the peptidoglycan-arabinogalactan layer to promote access of the spike tip to the inner membrane.

Given its proposed positioning within the cell wall, we tested whether the gp30 FN3-like N-terminal domain can bind to members of the cell wall glycome of mycobacteria (Figures 6I and S6L; Table S4). A fusion protein of the gp30 N-terminal domain to His-GST (glutathione S-transferase) was incubated with a previously validated array of 72 mycobacterial cell glycans[64] and showed binding to a series of arabinofuranose-containing components of arabinomannan, lipoarabinomannan, and arabinogalactan (Figure 6I; Table S4). The GST tag alone does not bind to any array glycans (Figure 6I). These observations are consistent with the model that the Bxb1 tail tip can maneuver close to the inner membrane and bind to cell wall constituents. These dynamic rearrangements could release the gp30 N-terminal domain from its interactions with the spike protein, allowing the gp30 FN3 domain to serve as a carbohydrate-binding module[65] for the gp30 D-ala-D-ala-carboxypeptidase domain, thus facilitating peptidoglycan cleavage.

(B) Model of the tail tube protein with domains colored and labeled.
(C) One complete ring of the tail tube model from the top (left) and bottom (right), colored by domain as in (B).
(D) Density from one ring of the tail tube reconstruction is colored by domain. The variable length of the hinge loop is shown as dashed lines tracing the ring floor and the lowest point of each domain.
(E) Representative tomographic slices of Bxb1 during infection (top) and post-genome translocation (bottom). The Bxb1 tail and mycobacterial surface are annotated with yellow and red arrowheads, respectively.
(F and G) Vertical (F) and horizontal (G) cross-sections showing internal tail tube structures in the pre-infection SPA (left, tape measure density), infection-state STA (middle, dsDNA density), and post-infection STA (right, hollow) maps. 2D central slices are shown as insets.
See also Figures S1, S2, and S3.

Please cite this article in press as: Freeman et al., Structure and infection dynamics of mycobacteriophage Bxb1, Cell (2025), https://doi.org/10.1016/j.cell.2025.03.027

CellPress
OPEN ACCESS



(legend on next page)

Please cite this article in press as: Freeman et al., Structure and infection dynamics of mycobacteriophage Bxb1, Cell (2025), https://doi.org/10.1016/j.cell.2025.03.027

Structural insights into Bxb1 as a phage display vaccine platform

The capsid structures of Bxb1 and its vaccine derivative, BaDAS-1 (<2.67 Å resolution; Figure S7; Table S1), offer insight into the use of Bxb1 as a vaccine display platform.[45] With prior engineering efforts, we observed that although removal of the CTE domains on the capsid and tail tube does not alter viral viability, the domains could not be simply replaced with similarly sized segments of the SARS-CoV-2 spike protein.[45] This may be a consequence of the interactions between the CTE domains: on both hexons and pentons of the Bxb1 capsid, pairs of CTEs from neighboring subunits interact to form a ~180 Å^2 interface (Figures S7A and S7B). This interface contains close contacts between the backbone and sidechain of one CTE's D351 and the backbones of the partner domains' D352 and V354 (Figure S7B). Although the CTE domains can be removed without obvious loss of viability in ideal lab conditions, it is plausible that this intimate interface stabilizes the capsids in less favorable conditions. Removal of the domains may also make either capsid assembly or stability more sensitive to perturbations such as CTE replacement with foreign antigens, especially if these replacements introduce steric clashes that are destabilizing.

Although replacement of the CTE domains was not possible, a vaccine candidate (BaDAS-1) was constructed by extending the extreme capsid subunit C terminus situated near the capsid floor. Only 30 amino acids could be appended, and larger peptides were not tolerated, likely due to space constraints. The C-terminal ends of adjacent Bxb1 subunits are separated by just ~16 Å, leaving only a small cavity into which additional amino acids can be extended (Figure S7C). A volume subtraction between BaDAS-1 and Bxb1 icosahedral SPA maps confirms that the displayed SARS-CoV-2 antigens fill these cavities, with excess density crowded between the CTE domains (Figure S7D). Display of larger peptides likely disrupts capsid assembly and stability. These structural insights should facilitate engineering designs in which the CTE can be readily substituted with desirable epitopes, for example, by choosing foreign antigens that can dimerize to provide stability like the CTEs or engineering a stabilizing minor capsid protein into Bxb1's genome. Alternatively, adding a flexible linker between the C terminus and the displayed peptide may reduce steric clashes from larger peptides while keeping the CTEs in place.

DISCUSSION

Here, we have reported structures of mycobacteriophage Bxb1 derived from cryo-EM and cryo-ET analyses. The Bxb1 structure has some common features of other structurally defined siphophages (χ, lambda, T5, DT57C, 80α, and vB_DshS-R4C), although we note that only 80α infects a gram-positive bacterial host.[28,30–32,34,35] These phages share common features of icosahedral capsids, hexameric tail tubes, and symmetry mismatches of the portal-capsid assemblies. They also share overall 3-fold symmetries in the tail-tip complexes and a Dit-like protein connecting the tail tube and hub assemblies. Bxb1, T5, DT57C, and 80α also all have collar ring structures and (often disordered) fibers attached to their Dit proteins. However, Bxb1 has numerous differences from these other phages, including an unusual cage-like protein (gp23) that undergoes dynamic changes during mycobacterial infection and a bulb-ended tail tip lacking a needle-like structure. Bxb1 also possesses wing structures (formed by gp30, gp31, and gp33), which are unique in both sequence and structure. These include an enzymatic domain in the tail wing base (gp30), which cryo-ET analysis suggests is positioned close to its presumed peptidoglycan target in the cell wall after dynamic changes in interactions with other tail-tip proteins. Notably, we observed weak density for protrusion through the cell wall during infection, but unlike previous reports on podophage P22 and myophages P1 and T4,[19,36,66] this density did not reach the inner membrane.

The extensive coverage of the capsid and tail with related C-terminal globular domains is a particularly unusual feature of the Bxb1 virion. Surprisingly, related domains are also present in gp23, the tail-tip cage (Figure S3E). Approximately 25% of actinobacteriophages have sequence-related capsids to Bxb1, but only ~1% of these have CTE domains.[41,42] Tail tube CTEs are only present in ~26% of related phages. The CTEs are reminiscent of IgG-like domains that are similarly located on many phages,[67] including some actinobacteriophages, but may be exclusively on the capsid[68] or the tail, as in phage lambda.[59] These Ig-like domains have been proposed to enhance phage association with mucosal surfaces[69] but could also be involved in enhancing capsid and tail stability like other decoration proteins.[70] The Bxb1 CTE domains are highly immunogenic and can be readily removed via genomic engineering with no apparent defect in phage viability.[45] The CTE domains cannot, however, be readily replaced perhaps due to destabilization, and there are constraints on extending them, likely due to crowding (Figure S7). Phages provide potentially powerful vaccine platforms,[71,72] and structural studies provide important clues as to how to better engineer mycobacteriophage-based vaccine platforms.[45]

Mass spectrometry of highly purified virus particles identified a total of 18 proteins that are likely to be virion structural components. However, gp5, gp12, and gp17 could not be modeled into the cryo-EM density maps. Bxb1 gp12 is the protease and may

Figure 5. Bxb1 tail-tip complex structure

(A) A composite cryo-EM map of the entire Bxb1 tail tip, with density segments colored as in Figure 1 (unmodeled density is shown in gray). Full density map on the left, clipped map on the right to show the interior organization.

(B) The complete protein model built into the density map is displayed as cartoons with atoms shown as sticks and balls. The model is in the same pose as in (A), and ten distinct proteins are labeled and color-coded to (A). Full model on the left, clipped view on the right to show the interior organization.

(C) Gene products present in the tail-tip complex are shown as monomers. Each subunit is labeled, and its copy number per phage particle is indicated.

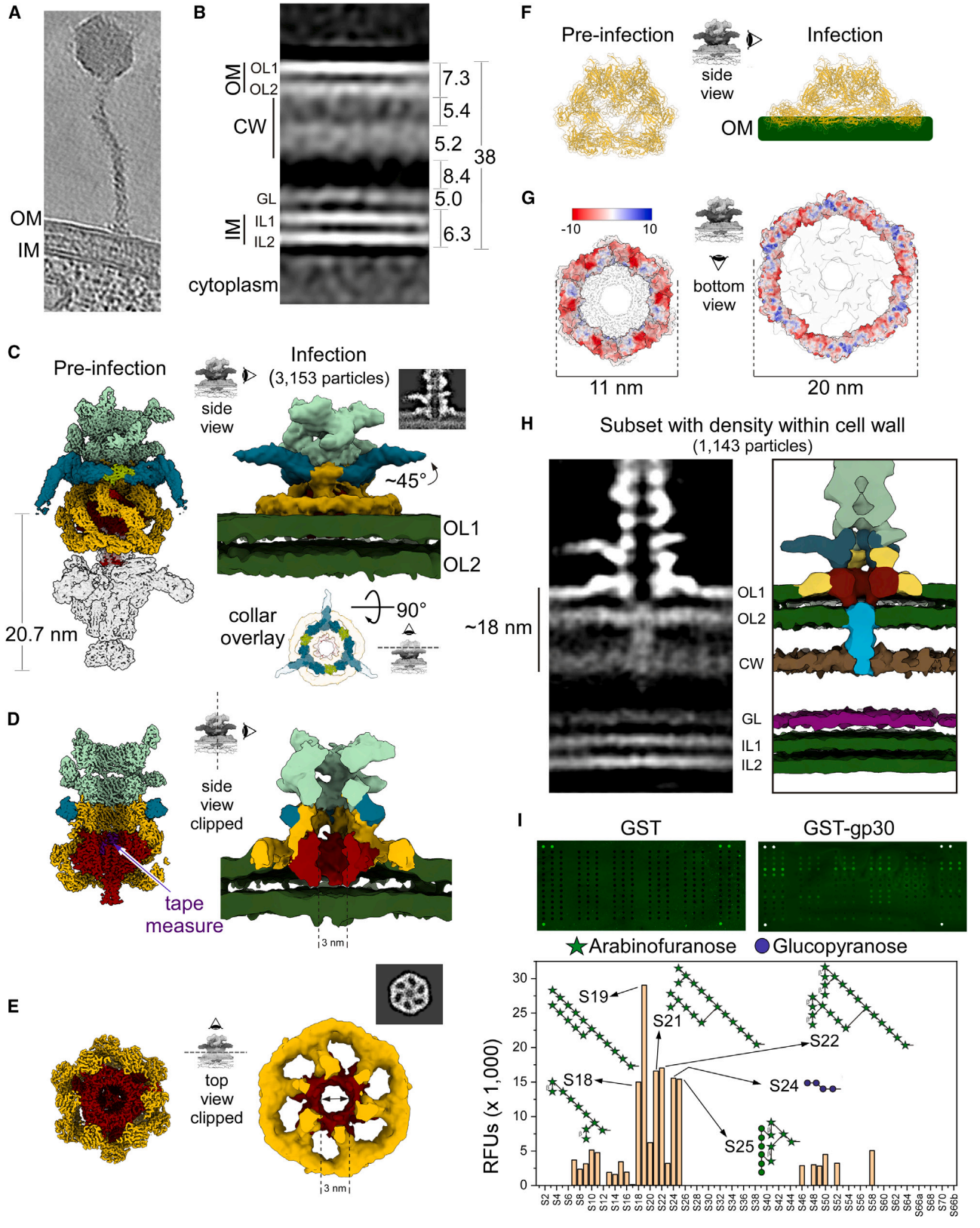
(D) The tail-tip complex is viewed from the bottom and displayed piece by piece, starting from the tail tube and moving toward the tail spike, as labeled. Segmented cryo-EM densities are shown as transparent surfaces, and protein models are shown as cartoons for each component.

(E) A color-coded interface map of the tail-tip complex. The circles represent individual protein subunits, and the lines connect subunits with interface areas ≥300 Å^2.

See also Figures S1, S2, S3, S4, and S6.

Please cite this article in press as: Freeman et al., Structure and infection dynamics of mycobacteriophage Bxb1, Cell (2025), https://doi.org/10.1016/j.cell.2025.03.027

CellPress
OPEN ACCESS



(legend on next page)

Please cite this article in press as: Freeman et al., Structure and infection dynamics of mycobacteriophage Bxb1, Cell (2025), https://doi.org/10.1016/j.cell.2025.03.027

be associated with a small subset of incompletely assembled particles in the phage preparation (~3% of capsids were empty and may belong to this subset). Bxb1 gp5 and gp17 are virion components but could not be resolved in the density (Figure S4F). The products of six genes (24, 26, 27, 28, 32, and 34) embedded within the structural gene operon are not present in Bxb1 virions (Figure 1A), and three of these (gp27, gp28, and gp34) are predicted to harbor transmembrane domains and may be associated with the host cell envelope. In general, these genes are as well conserved as their flanking tail genes among the 215 sequenced cluster A1 phages and may play a role in phage-phage or phage-host dynamics (such as defense or counter-defense),[73] although Bxb1 lacks encapsidated anti-defense proteins.[74,75]

Combining cryo-EM and cryo-ET provides structural insights into phage-host interactions. Bound phages with empty capsids lack density within the tail tube and have presumably completed DNA injection, whereas the tail tubes of host-attached phages with DNA-filled capsids appear to be filled with DNA rather than the TMP. With the absence of a distinct inner membrane channel, this may represent an infection intermediate, where DNA transfer is paused during the infection process. Even more strikingly, cryo-ET analysis suggests that the tail-tip cage rotates and flattens to form a disc in the plane of the outer membrane upon host binding. This dramatic conformational change is associated with a loss of ordered density from the lower tail tip, suggesting the tail spike protein and wing structures undergo dynamic rearrangements. Indeed, a subset of particles was found to have density beneath the tail that spans the outer membrane and punctures through the cell wall, possibly reflecting tail spike rearrangements that facilitate DNA translocation into the cytoplasm.

These structural analyses suggest that several Bxb1 proteins likely contribute to host specificity, including the baseplate hub (gp25) and tail-tip cage (gp23), which are closely associated with the mycomembrane during infection. The gp4 protein is predicted to form a long, fiber-like structure but appears somewhat flexible and poorly resolved, although it is positioned to interact with the outer cell surface. A domain on the tail spike protein (gp29; residues I217-D408) is at the extreme "bottom" of the Bxb1 particle and is likely near the inner membrane during infection and may also confer host specificity. Interestingly, host range mutants with amino acid substitutions within phage tail spike proteins have been isolated in other mycobacteriophages,[7] but the substitutions map in an upper part of the spike protein where rearrangements of Bxb1 are needed to provide an opening for TMP and DNA ejection. This raises the possibility that host range is also dependent on the dynamic rearrangements of the tail tip observed by cryo-ET.

Mycobacteriophages are highly diverse, and it is unclear what parts of the Bxb1 structure—especially the tail tip—are likely shared within the broader group of phages. We note that several Bxb1 virion proteins are only found within cluster A phages (e.g., the gp4 tail collar fiber, gp6 collar spacer, gp19 tail tube, gp23 tail-tip cage, and gp31 tail wing arm), whereas others are broadly distributed among the actinobacteriophages (including the gp14 capsid subunit, gp25 baseplate hub, and gp30 tail wing base). Curiously, gp30—which contains an arabinan-binding FN3-like domain and a D-ala-D-ala carboxypeptidase domain and is likely important for infection—is not universally present among mycobacteriophages. It is present in only two of the 21 subclusters within cluster A and in some representatives of seven other clusters. Phages lacking a gp30-like protein do not appear to have an alternative protein with equivalent enzymatic activity that could facilitate the infection process.

A major limitation to the therapeutic use of mycobacteriophages is their relatively narrow specificity for clinical *Mycobacterium* isolates.[14] Specificity could be engineered through the swapping of tail proteins among different phages, potentially combining the epitope presentation potential of the Bxb1 capsid and tail tube proteins with the tail-tip proteins of therapeutically useful phages such as BPs and Muddy.[8] The Bxb1 structure illustrates the complexity of this by mapping the intricate network of protein interactions in the tail tip. However, the structures also provide a framework for such engineering. Structural characterization of additional mycobacteriophages will be needed, but the swapping of tail parts, including substitutions of the tail spike, baseplate hub, cage, and wing proteins, can now be contemplated.

Cryo-EM and cryo-ET along with recently developed bioinformatic tools provide powerful approaches for complete structural analyses of bacteriophages. There are numerous mycobacteriophages that warrant similar investigation, including

Figure 6. Remodeling of the tail tip upon attachment

(A) A representative tomogram slice of Bxb1 infecting *M. smegmatis*.

(B) Subtomogram averaged density section of the *M. smegmatis* cell envelope. Outer mycomembrane (OM) with outer leaflets 1 and 2 (OL1 and OL2), cell wall, granular layer (GL), and inner membrane (IM) with inner leaflets 1 and 2 (IL1 and IL2) are indicated. The *M. smegmatis* cell envelope is approximately 38 nm thick.

(C) Comparison of the pre-infection (SPA, left) and infection (subtomogram average, right) state structures of the tail-tip complex, colored as in Figure 5. Components in gray in the pre-infection state are not resolved in the infection-state structure. The upper right inset shows a 2D density section of the subtomogram average map. At the lower right, the pre-infection collar density is overlaid with the infection-state map and shown from the top.

(D) Clipped maps showing tape measure density, a closed baseplate hub for the pre-infection map (left), and an open tail-tip core in the infection state (right).

(E) Clipped maps of the isolated tail-tip cage and baseplate hub densities in the pre-infection (left) and infection (right) state, shown from the top. A 2D section is shown at the upper right.

(F) Side views of tail-tip cage models in the pre-infection and infection states, shown as cartoons with transparent surfaces.

(G) Bottom views of the tail-tip cage models are shown as electrostatic surfaces, and a key indicates the surface potential.

(H) A subset of particles (1,143/3,153) with density below the baseplate hub were subjected to focused refinement, revealing weak density within the cell wall (2D section, left). This density was segmented and tentatively assigned as the tail spike protein (cyan in the 3D map, right). Layers of the mycobacterial wall are labeled as in (B).

(I) Binding of Bxb1 gp30-GST and GST alone to a glycan array (top), with structures (below) for fragments with the highest gp30-GST signal intensity (see Table S4).

See also Figure S2.

those proposed for therapeutic use against tuberculosis and NTM infections.[14,15,76] The comparative structural analyses of these phages are expected to provide further insights into host range and infection structural dynamics and advance the use of mycobacteriophages for therapeutic and biotechnological applications.

Limitations of the study

There are some limitations to the structures described here, most notably the resolution of the cryo-ET structures relative to those determined by cryo-EM. Although the ~12 Å resolution of the cryo-ET structures represents an improvement over previous phage tail-tip subtomogram average structures,[19,36,66] protein structures could still only be modeled with rigid-body fitting of domains from the cryo-EM structure. Technological advances will undoubtedly improve the resolution of cryo-ET analysis, enabling detailed molecular-level analysis of all tail proteins and their disposition during infection. Furthermore, despite significant efforts, the structural complexity of the *Mycobacterium* cell wall and its glycome remains incompletely understood; its further definition will assist in understanding phage-*Mycobacterium* interactions. Finally, Bxb1 is the first mycobacteriophage to be structurally defined, and the genomes of the >2,600 sequenced mycobacteriophages reveal enormous sequence diversity among their virion structure and assembly genes that remain to be explored. The structural characterization of representatives of these phages promises rapid and productive advances that will guide phage engineering.

RESOURCE AVAILABILITY

Lead contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by the lead contact, Graham Hatfull (gfh@pitt.edu).

Materials availability

This study did not generate new, unique reagents.

Data and code availability

- Cryo-EM reconstructions and protein models have been deposited at EMDB and PDB, respectively, and accession numbers are listed in the key resources table. Raw cryo-EM micrographs and raw cryo-ET tilt series have been deposited at EMPIAR, and accession numbers are listed in the key resources table. Subtomogram average maps have been deposited at EMDB, and accession numbers are listed in the key resources table. Mass spectrometry proteomics data have been deposited to PRIDE, and accession numbers are listed in the key resources table.
- This paper does not report original code.
- Any additional information required to reanalyze the data reported in this paper is available from the lead contact upon request.

ACKNOWLEDGMENTS

We thank Colin Lewis for excellent technical assistance. We greatly appreciate Christian Gauthier and Lawrence Abad for their help with setting up a workstation for cryo-EM analyses and protein structure prediction. This work was supported by grants to G.F.H. from NIH GM116884 and from the Howard Hughes Medical Institute GT12053, to K.G.F. from NIH K99AI173544, and to D.P. from the Scripps Research start-up fund and Bachrach Family Foundation grant. The University of Pittsburgh Titan Krios microscope and Falcon 3 camera were supported by the Office of the Director, National Institutes of Health, under award numbers S10 OD025009 and S10 OD019995, respectively, to J.F.C. Cryo-ET data collection was performed using equipment supported by NIH S10OD032467. A portion of this research was supported by NIH grant U24GM129547 and performed at the PNCC at OHSU and accessed through EMSL (grid.436923.9), a DOE Office of Science User Facility sponsored by the Department of Energy Office of Science and Office Biological and Environmentla Research. T.L.L. and M.-C.H. are grateful for support from Academia Sinica (AS-IDR-110-06).

AUTHOR CONTRIBUTIONS

Conceptualization, K.G.F., D.P., and G.F.H.; methodology, K.G.F., S.M., S.J.W., A.H., R.J.P., J.T.N., J.F.C., D.P., T.L.L., M.-C.H., and D.J.-S.; validation, K.G.F., S.J.W., J.F.C., and D.P.; formal analysis, K.G.F., S.M., and D.P.; investigation, K.G.F., S.M., L.S.M., J.P., B.H.S., V.O., A.H., R.J.P., J.T.N., T.L.L., M.-C.H., and D.P.; resources, K.G.F., S.J.W., T.L.L., J.F.C., D.P., and G.F.H.; data curation, K.G.F., S.M., L.S.M., S.J.W., J.F.C., and D.P.; writing – original draft, K.G.F., S.M., T.L.L., D.P., and G.F.H.; writing – review and editing, K.G.F., S.M., L.S.M., S.J.W., J.P., B.H.S., V.O., A.H., R.J.P., J.T.N., M.-C.H., T.L.L., J.F.C., D.P., D.J.-S., and G.F.H.; visualization, K.G.F., S.M., and D.P.; supervision, S.J.W., T.L.L., M.-C.H., J.F.C., D.P., and G.F.H.; project administration, K.G.F. and G.F.H.; and funding acquisition, K.G.F., S.J.W., T.L.L., M.-C.H., J.F.C., D.P., and G.F.H.

DECLARATION OF INTERESTS

The authors declare no competing interests.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS
  - Bacterial strains and phages
- METHOD DETAILS
  - Proteomics
  - Cryo-electron micrograph collection
  - Cryo-EM reconstructions and model building
  - Cryo-electron tomograph collection
  - Cryo-ET subtomogram averaging
  - Cloning, expression, and purification of Bxb1 gp30 FN3 domain
  - Glycan binding using the glycan microarray
  - Microarray Scan and Data Analysis

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.cell.2025.03.027.

Received: August 21, 2024
Revised: November 27, 2024
Accepted: March 14, 2025
Published: April 15, 2025

REFERENCES

1. Hatfull, G.F. (2020). Actinobacteriophages: Genomics, dynamics, and applications. Annu. Rev. Virol. *7*, 37–61. https://doi.org/10.1146/annurev-virology-122019-070009.
2. Dulberger, C.L., Rubin, E.J., and Boutte, C.C. (2020). The mycobacterial cell envelope — a moving target. Nat. Rev. Microbiol. *18*, 47–59. https://doi.org/10.1038/s41579-019-0273-7.
3. Payne, K., Sun, Q., Sacchettini, J., and Hatfull, G.F. (2009). Mycobacteriophage Lysin B is a novel mycolylarabinogalactan esterase. Mol. Microbiol. *73*, 367–381. https://doi.org/10.1111/j.1365-2958.2009.06775.x.

4. Singh, A.K., Gangakhedkar, R., Thakur, H.S., Raman, S.K., Patil, S.A., and Jain, V. (2023). Mycobacteriophage D29 Lysin B exhibits promising antimycobacterial activity against drug-resistant Mycobacterium tuberculosis. Microbiol. Spec. *11*, e04597–e04522. https://doi.org/10.1128/spectrum.04597-22.

5. Hatfull, G.F. (2018). Mycobacteriophages. Microbiol. Spectr. *6*. https://doi.org/10.1128/microbiolspec.GPP3-0026-2018.

6. Chen, J., Kriakov, J., Singh, A., Jacobs, W.R., Jr., Besra, G.S., and Bhatt, A. (2009). Defects in glycopeptidolipid biosynthesis confer phage I3 resistance in *Mycobacterium smegmatis*. Microbiology *155*, 4050–4057. https://doi.org/10.1099/mic.0.033209-0.

7. Wetzel, K.S., Illouz, M., Abad, L., Aull, H.G., Russell, D.A., Garlena, R.A., Cristinziano, M., Malmsheimer, S., Chalut, C., Hatfull, G.F., et al. (2023). Therapeutically useful mycobacteriophages BPs and Muddy require trehalose polyphleates. Nat. Microbiol. *8*, 1717–1731. https://doi.org/10.1038/s41564-023-01451-6.

8. Dedrick, R.M., Smith, B.E., Cristinziano, M., Freeman, K.G., Jacobs-Sera, D., Belessis, Y., Whitney Brown, A., Cohen, K.A., Davidson, R.M., van Duin, D., et al. (2023). Phage Therapy of Mycobacterium Infections: Compassionate Use of Phages in 20 Patients With Drug-Resistant Mycobacterial Disease. Clin. Infect. Dis. *76*, 103–112. https://doi.org/10.1093/cid/ciac453.

9. Dedrick, R.M., Freeman, K.G., Nguyen, J.A., Bahadirli-Talbott, A., Smith, B.E., Wu, A.E., Ong, A.S., Lin, C.T., Ruppel, L.C., Parrish, N.M., et al. (2021). Potent antibody-mediated neutralization limits bacteriophage treatment of a pulmonary Mycobacterium abscessus infection. Nat. Med. *27*, 1357–1361. https://doi.org/10.1038/s41591-021-01403-9.

10. Dedrick, R.M., Guerrero-Bustamante, C.A., Garlena, R.A., Russell, D.A., Ford, K., Harris, K., Gilmour, K.C., Soothill, J., Jacobs-Sera, D., Schooley, R.T., et al. (2019). Engineered bacteriophages for treatment of a patient with a disseminated drug-resistant *Mycobacterium abscessus*. Nat. Med. *25*, 730–733. https://doi.org/10.1038/s41591-019-0437-z.

11. Strathdee, S.A., Hatfull, G.F., Mutalik, V.K., and Schooley, R.T. (2023). Phage therapy: From biological mechanisms to future directions. Cell *186*, 17–31. https://doi.org/10.1016/j.cell.2022.11.017.

12. Nick, J.A., Dedrick, R.M., Gray, A.L., Vladar, E.K., Smith, B.E., Freeman, K.G., Malcolm, K.C., Epperson, L.E., Hasan, N.A., Hendrix, J., et al. (2022). Host and pathogen response to bacteriophage engineered against Mycobacterium abscessus lung infection. Cell *185*, 1860–1874.e12. https://doi.org/10.1016/j.cell.2022.04.024.

13. Little, J.S., Dedrick, R.M., Freeman, K.G., Cristinziano, M., Smith, B.E., Benson, C.A., Jhaveri, T.A., Baden, L.R., Solomon, D.A., and Hatfull, G.F. (2022). Bacteriophage treatment of disseminated cutaneous Mycobacterium chelonae infection. Nat. Commun. *13*, 2313. https://doi.org/10.1038/s41467-022-29689-4.

14. Hatfull, G.F. (2022). Mycobacteriophages: From Petri dish to patient. PLoS Pathog. *18*, e1010602. https://doi.org/10.1371/journal.ppat.1010602.

15. Hatfull, G.F. (2023). Phage therapy for nontuberculous mycobacteria: Challenges and opportunities. Pulm. Ther. *9*, 91–107. https://doi.org/10.1007/s41030-022-00210-y.

16. Russell, D.A., and Hatfull, G.F. (2017). PhagesDB: the actinobacteriophage database. Bioinformatics *33*, 784–786. https://doi.org/10.1093/bioinformatics/btw711.

17. Subramanian, S., Bergland Drarvik, S.M., Tinney, K.R., and Parent, K.N. (2024). Cryo-EM structure of a Shigella podophage reveals a hybrid tail and novel decoration proteins. Structure *32*, 24–34.e4. https://doi.org/10.1016/j.str.2023.10.007.

18. Li, F., Hou, C.D., Yang, R., Whitehead, R., Teschke, C.M., and Cingolani, G. (2022). High-resolution cryo-EM structure of the Shigella virus Sf6 genome delivery tail machine. Sci. Adv. *8*, eadc9641. https://doi.org/10.1126/sciadv.adc9641.

19. Wang, C., Tu, J., Liu, J., and Molineux, I.J. (2019). Structural dynamics of bacteriophage P22 infection initiation revealed by cryo-electron tomography. Nat. Microbiol. *4*, 1049–1056. https://doi.org/10.1038/s41564-019-0403-z.

20. Cai, L., Liu, H., Zhang, W., Xiao, S., Zeng, Q., and Dang, S. (2023). Cryo-EM structure of cyanophage P-SCSP1u offers insights into DNA gating and evolution of T7-like viruses. Nat. Commun. *14*, 6438. https://doi.org/10.1038/s41467-023-42258-7.

21. Šiborová, M., Füzik, T., Procházková, M., Nováček, J., Benešík, M., Nilsson, A.S., and Plevka, P. (2022). Tail proteins of phage SU10 reorganize into the nozzle for genome delivery. Nat. Commun. *13*, 5622. https://doi.org/10.1038/s41467-022-33305-w.

22. Xu, J., Wang, D., Gui, M., and Xiang, Y. (2019). Structural assembly of the tailed bacteriophage ϕ29. Nat. Commun. *10*, 2366. https://doi.org/10.1038/s41467-019-10272-3.

23. Veesler, D., Robin, G., Lichière, J., Auzat, I., Tavares, P., Bron, P., Campanacci, V., and Cambillau, C. (2010). Crystal Structure of Bacteriophage SPP1 Distal Tail Protein (gp19.1): A BASEPLATE HUB PARADIGM IN GRAM-POSITIVE INFECTING PHAGES. J. Biol. Chem. *285*, 36666–36673. https://doi.org/10.1074/jbc.M110.157529.

24. Eruera, A.-R., Hodgkinson-Bean, J., Rutter, G.L., Hills, F.R., Kumaran, R., Crowe, A.J.M., Jadav, N., Chang, F., McJarrow-Keller, K., Jorge, F., et al. (2024). Ejectosome of Pectobacterium bacteriophage ΦM1. PNAS Nexus *3*, pgae416. https://doi.org/10.1093/pnasnexus/pgae416.

25. Yang, F., Jiang, Y.-L., Zhang, J.-T., Zhu, J., Du, K., Yu, R.-C., Wei, Z.-L., Kong, W.-W., Cui, N., Li, W.-F., et al. (2023). Fine structure and assembly pattern of a minimal myophage Pam3. Proc. Natl. Acad. Sci. USA *120*, e2213727120. https://doi.org/10.1073/pnas.2213727120.

26. Wang, Z., Fokine, A., Guo, X., Jiang, W., Rossmann, M.G., Kuhn, R.J., Luo, Z.-H., and Klose, T. (2023). Structure of Vibrio Phage XM1, a Simple Contractile DNA Injection Machine. Viruses *15*, 1673. https://doi.org/10.3390/v15081673.

27. Li, F., Hou, C.D., Lokareddy, R.K., Yang, R., Forti, F., Briani, F., and Cingolani, G. (2023). High-resolution cryo-EM structure of the Pseudomonas bacteriophage E217. Nat. Commun. *14*, 4052. https://doi.org/10.1038/s41467-023-39756-z.

28. Sonani, R.R., Esteves, N.C., Scharf, B.E., and Egelman, E.H. (2024). Cryo-EM structure of flagellotropic bacteriophage Chi. Structure *32*, 856–865.e3. https://doi.org/10.1016/j.str.2024.03.011.

29. Iglesias, S.M., Lokareddy, R.K., Yang, R., Li, F., Yeggoni, D.P., David Hou, C.F., Leroux, M.N., Cortines, J.R., Leavitt, J.C., Bird, M., et al. (2023). Molecular Architecture of Salmonella Typhimurium Virus P22 Genome Ejection Machinery. J. Mol. Biol. *435*, 168365. https://doi.org/10.1016/j.jmb.2023.168365.

30. Ayala, R., Moiseenko, A.V., Chen, T.H., Kulikov, E.E., Golomidova, A.K., Orekhov, P.S., Street, M.A., Sokolova, O.S., Letarov, A.V., and Wolf, M. (2023). Nearly complete structure of bacteriophage DT57C reveals architecture of head-to-tail interface and lateral tail fibers. Nat. Commun. *14*, 8205. https://doi.org/10.1038/s41467-023-43824-9.

31. Linares, R., Arnaud, C.A., Effantin, G., Darnault, C., Epalle, N.H., Boeri Erba, E., Schoehn, G., and Breyton, C. (2023). Structural basis of bacteriophage T5 infection trigger and E. coli cell wall perforation. Sci. Adv. *9*, eade9674. https://doi.org/10.1126/sciadv.ade9674.

32. Wang, C., Duan, J., Gu, Z., Ge, X., Zeng, J., and Wang, J. (2024). Architecture of the bacteriophage lambda tail. Structure *32*, 35–46.e3. https://doi.org/10.1016/j.str.2023.10.006.

33. Xiao, H., Tan, L., Tan, Z., Zhang, Y., Chen, W., Li, X., Song, J., Cheng, L., and Liu, H. (2023). Structure of the siphophage neck–Tail complex suggests that conserved tail tip proteins facilitate receptor binding and tail assembly. PLoS Biol. *21*, e3002441. https://doi.org/10.1371/journal.pbio.3002441.

34. Huang, Y., Sun, H., Wei, S., Cai, L., Liu, L., Jiang, Y., Xin, J., Chen, Z., Que, Y., Kong, Z., et al. (2023). Structure and proposed DNA delivery mechanism of a marine roseophage. Nat. Commun. *14*, 3609. https://doi.org/10.1038/s41467-023-39220-y.

35. Kizziah, J.L., Manning, K.A., Dearborn, A.D., and Dokland, T. (2020). Structure of the host cell recognition and penetration machinery of a Staphylococcus aureus bacteriophage. PLoS Pathog. *16*, e1008314. https://doi.org/10.1371/journal.ppat.1008314.

36. Liu, J., Chen, C.-Y., Shiomi, D., Niki, H., and Margolin, W. (2011). Visualization of bacteriophage P1 infection by cryo-electron tomography of tiny *Escherichia coli*. Virology *417*, 304–311. https://doi.org/10.1016/j.virol.2011.06.005.

37. Podgorski, J., Calabrese, J., Alexandrescu, L., Jacobs-Sera, D., Pope, W., Hatfull, G., and White, S. (2020). Structures of Three Actinobacteriophage Capsids: Roles of Symmetry and Accessory Proteins. Viruses *12*, 294. https://doi.org/10.3390/v12030294.

38. Podgorski, J.M., Freeman, K., Gosselin, S., Huet, A., Conway, J.F., Bird, M., Grecco, J., Patel, S., Jacobs-Sera, D., Hatfull, G., et al. (2023). A structural dendrogram of the actinobacteriophage major capsid proteins provides important structural insights into the evolution of capsid stability. Structure *31*, 282–294.e5. https://doi.org/10.1016/j.str.2022.12.012.

39. Podgorski, J.M., Podgorski, J., Abad, L., Jacobs-Sera, D., Freeman, K.G., Brown, C., Hatfull, G., Luque, A., and White, S.J. (2023). A novel stabilization mechanism accommodating genome length variation in evolutionarily related viral capsids. Preprint at bioRxiv. https://doi.org/10.1101/2023.11.03.565530.

40. Sassi, M., Bebeacua, C., Drancourt, M., and Cambillau, C. (2013). The First Structure of a Mycobacteriophage, the Mycobacterium abscessus subsp. bolletii Phage Araucaria. J. Virol. *87*, 8099–8109. https://doi.org/10.1128/JVI.01209-13.

41. Hatfull, G.F. (2005). Mycobacteriophages. In The Bacteriophages, R.L. Calendar and S.T. Abedon, eds. (Oxford University Press), pp. 603–620.

42. Mediavilla, J., Jain, S., Kriakov, J., Ford, M.E., Duda, R.L., Jacobs, W.R., Hendrix, R.W., and Hatfull, G.F. (2000). Genome organization and characterization of mycobacteriophage Bxb1. Mol. Microbiol. *38*, 955–970. https://doi.org/10.1046/j.1365-2958.2000.02183.x.

43. Hosur, V., Low, B.E., and Wiles, M.V. (2022). Programmable RNA-Guided Large DNA Transgenesis by CRISPR/Cas9 and Site-Specific Integrase Bxb1. Front. Bioeng. Biotechnol. *10*, 910151. https://doi.org/10.3389/fbioe.2022.910151.

44. Balabaskaran-Nina, P., and Desai, S.A. (2018). Diverse target gene modifications in Plasmodium falciparum using Bxb1 integrase and an intronic attB. Parasit. Vectors *11*, 548. https://doi.org/10.1186/s13071-018-3129-5.

45. Freeman, K.G., Wetzel, K.S., Zhang, Y., Zack, K.M., Jacobs-Sera, D., Walters, S.M., Barbeau, D.J., McElroy, A.K., Williams, J.V., and Hatfull, G.F. (2021). A Mycobacteriophage-Based Vaccine Platform: SARS-CoV-2 Antigen Expression and Display. Microorganisms *9*, 2414. https://doi.org/10.3390/microorganisms9122414.

46. Hatfull, G.F., and Sarkis, G.J. (1993). DNA sequence, structure and gene expression of mycobacteriophage L5: a phage system for mycobacterial genetics. Mol. Microbiol. *7*, 395–405. https://doi.org/10.1111/j.1365-2958.1993.tb01131.x.

47. Duda, R.L., and Teschke, C.M. (2019). The amazing HK97 fold: versatile results of modest differences. Curr. Opin. Virol. *36*, 9–16. https://doi.org/10.1016/j.coviro.2019.02.001.

48. Wikoff, W.R., Liljas, L., Duda, R.L., Tsuruta, H., Hendrix, R.W., and Johnson, J.E. (2000). Topologically Linked Protein Rings in the Bacteriophage HK97 Capsid. Science *289*, 2129–2133. https://doi.org/10.1126/science.289.5487.2129.

49. Morais, M.C., Choi, K.H., Koti, J.S., Chipman, P.R., Anderson, D.L., and Rossmann, M.G. (2005). Conservation of the Capsid Structure in Tailed dsDNA Bacteriophages: the Pseudoatomic Structure of ϕ29. Mol. Cell *18*, 149–159. https://doi.org/10.1016/j.molcel.2005.03.013.

50. Parent, K.N., Gilcrease, E.B., Casjens, S.R., and Baker, T.S. (2012). Structural evolution of the P22-like phages: Comparison of Sf6 and P22 procapsid and virion architectures. Virology *427*, 177–188. https://doi.org/10.1016/j.virol.2012.01.040.

51. Parent, K.N., Khayat, R., Tu, L.H., Suhanovsky, M.M., Cortines, J.R., Teschke, C.M., Johnson, J.E., and Baker, T.S. (2010). P22 Coat Protein Structures Reveal a Novel Mechanism for Capsid Maturation: Stability without Auxiliary Proteins or Chemical Crosslinks. Structure *18*, 390–401. https://doi.org/10.1016/j.str.2009.12.014.

52. Zhao, H., Li, K., Lynn, A.Y., Aron, K.E., Yu, G., Jiang, W., and Tang, L. (2017). Structure of a headful DNA-packaging bacterial virus at 2.9 Å resolution by electron cryo-microscopy. Proc. Natl. Acad. Sci. USA *114*, 3601–3606. https://doi.org/10.1073/pnas.1615025114.

53. Chen, W., Xiao, H., Wang, L., Wang, X., Tan, Z., Han, Z., Li, X., Yang, F., Liu, Z., Song, J., et al. (2021). Structural changes in bacteriophage T7 upon receptor-induced genome ejection. Proc. Natl. Acad. Sci. USA *118*, e2102003118. https://doi.org/10.1073/pnas.2102003118.

54. Fang, Q., Tang, W.-C., Tao, P., Mahalingam, M., Fokine, A., Rossmann, M.G., and Rao, V.B. (2020). Structural morphing in a symmetry-mismatched viral vertex. Nat. Commun. *11*, 1713. https://doi.org/10.1038/s41467-020-15575-4.

55. Huet, A., Oh, B., Maurer, J., Duda, R.L., and Conway, J.F. (2023). A symmetry mismatch unraveled: How phage HK97 scaffold flexibly accommodates a 12-fold pore at a 5-fold viral capsid vertex. Sci. Adv. *9*, eadg8868. https://doi.org/10.1126/sciadv.adg8868.

56. Xiao, H., Zhou, J., Yang, F., Liu, Z., Song, J., Chen, W., Liu, H., and Cheng, L. (2023). Assembly and Capsid Expansion Mechanism of Bacteriophage P22 Revealed by High-Resolution Cryo-EM Structures. Viruses *15*, 355. https://doi.org/10.3390/v15020355.

57. Orlov, I., Roche, S., Brasilès, S., Lukoyanova, N., Vaney, M.-C., Tavares, P., and Orlova, E.V. (2022). CryoEM structure and assembly mechanism of a bacterial virus genome gatekeeper. Nat. Commun. *13*, 7283. https://doi.org/10.1038/s41467-022-34999-8.

58. Zinke, M., Schröder, G.F., and Lange, A. (2022). Major tail proteins of bacteriophages of the order Caudovirales. J. Biol. Chem. *298*, 101472. https://doi.org/10.1016/j.jbc.2021.101472.

59. Campbell, P.L., Duda, R.L., Nassur, J., Conway, J.F., and Huet, A. (2020). Mobile Loops and Electrostatic Interactions Maintain the Flexible Tail Tube of Bacteriophage Lambda. J. Mol. Biol. *432*, 384–395. https://doi.org/10.1016/j.jmb.2019.10.031.

60. Zinke, M., Sachowsky, K.A.A., Öster, C., Zinn-Justin, S., Ravelli, R., Schröder, G.F., Habeck, M., and Lange, A. (2020). Architecture of the flexible tail tube of bacteriophage SPP1. Nat. Commun. *11*, 5759. https://doi.org/10.1038/s41467-020-19611-1.

61. Emsley, P., and Cowtan, K. (2004). Coot: model-building tools for molecular graphics. Acta Crystallogr. D Biol. Crystallogr. *60*, 2126–2132. https://doi.org/10.1107/S0907444904019158.

62. Zuber, B., Chami, M., Houssin, C., Dubochet, J., Griffiths, G., and Daffé, M. (2008). Direct Visualization of the Outer Membrane of Mycobacteria and Corynebacteria in Their Native State. J. Bacteriol. *190*, 5672–5680. https://doi.org/10.1128/JB.01919-07.

63. Hoffmann, C., Leis, A., Niederweis, M., Plitzko, J.M., and Engelhardt, H. (2008). Disclosure of the mycobacterial outer membrane: Cryo-electron tomography and vitreous sections reveal the lipid bilayer structure. Proc. Natl. Acad. Sci. USA *105*, 3963–3967. https://doi.org/10.1073/pnas.0709530105.

64. Zheng, R.B., Jégouzo, S.A.F., Joe, M., Bai, Y., Tran, H.A., Shen, K., Saupe, J., Xia, L., Ahmed, M.F., Liu, Y.H., et al. (2017). Insights into Interactions of Mycobacteria with the Host Innate Immune System from a Novel Array of Synthetic Mycobacterial Glycans. ACS Chem. Biol. *12*, 2990–3002. https://doi.org/10.1021/acschembio.7b00797.

65. Hayes, S., Vincentelli, R., Mahony, J., Nauta, A., Ramond, L., Lugli, G.A., Ventura, M., van Sinderen, D., and Cambillau, C. (2018). Functional carbohydrate binding modules identified in evolved dits from siphophages infecting various Gram-positive bacteria. Mol. Microbiol. *110*, 777–795. https://doi.org/10.1111/mmi.14124.

66. Hu, B., Margolin, W., Molineux, I.J., and Liu, J. (2015). Structural remodeling of bacteriophage T4 and host membranes during infection initiation. Proc. Natl. Acad. Sci. USA *112*, E4919–E4928. https://doi.org/10.1073/pnas.1501064112.

67. Fraser, J.S., Yu, Z., Maxwell, K.L., and Davidson, A.R. (2006). Ig-like domains on bacteriophages: a tale of promiscuity and deceit. J. Mol. Biol. *359*, 496–507. https://doi.org/10.1016/j.jmb.2006.03.043.

68. Zheng, J., Chen, W., Xiao, H., Yang, F., Li, X., Song, J., Cheng, L., and Liu, H. (2022). A Capsid Structure of Ralstonia solanacearum podoviridae GP4 with a Triangulation Number T = 9. Viruses *14*, 2431. https://doi.org/10.3390/v14112431.

69. Barr, J.J., Auro, R., Furlan, M., Whiteson, K.L., Erb, M.L., Pogliano, J., Stotland, A., Wolkowicz, R., Cutting, A.S., Doran, K.S., et al. (2013). Bacteriophage adhering to mucus provide a non-host-derived immunity. Proc. Natl. Acad. Sci. USA *110*, 10771–10776. https://doi.org/10.1073/pnas.1305923110.

70. Dedeo, C.L., Teschke, C.M., and Alexandrescu, A.T. (2020). Keeping It Together: Structures, Functions, and Applications of Viral Decoration Proteins. Viruses *12*, 1163. https://doi.org/10.3390/v12101163.

71. Zhu, J., Ananthaswamy, N., Jain, S., Batra, H., Tang, W.C., Lewry, D.A., Richards, M.L., David, S.A., Kilgore, P.B., Sha, J., et al. (2021). A universal bacteriophage T4 nanoparticle platform to design multiplex SARS-CoV-2 vaccine candidates by CRISPR engineering. Sci. Adv. *7*, eabh1547. https://doi.org/10.1126/sciadv.abh1547.

72. Zhu, J., Tao, P., Chopra, A.K., and Rao, V.B. (2024). Bacteriophage T4 as a Protein-Based, Adjuvant- and Needle-Free, Mucosal Pandemic Vaccine Design Platform. Annu. Rev. Virol. *11*, 395–420. https://doi.org/10.1146/annurev-virology-111821-111145.

73. Cumby, N., Edwards, A.M., Davidson, A.R., and Maxwell, K.L. (2012). The bacteriophage HK97 gp15 moron element encodes a novel superinfection exclusion protein. J. Bacteriol. *194*, 5012–5019. https://doi.org/10.1128/JB.00843-12.

74. Rifat, D., Wright, N.T., Varney, K.M., Weber, D.J., and Black, L.W. (2008). Restriction endonuclease inhibitor IPI* of bacteriophage T4: a novel structure for a dedicated target. J. Mol. Biol. *375*, 720–734. https://doi.org/10.1016/j.jmb.2007.10.064.

75. Bair, C.L., Rifat, D., and Black, L.W. (2007). Exclusion of glucosyl-hydroxymethylcytosine DNA containing bacteriophages is overcome by the injected protein inhibitor IPI*. J. Mol. Biol. *366*, 779–789. https://doi.org/10.1016/j.jmb.2006.11.049.

76. Guerrero-Bustamante, C.A., Dedrick, R.M., Garlena, R.A., Russell, D.A., and Hatfull, G.F. (2021). Toward a phage cocktail for tuberculosis: Susceptibility and tuberculocidal action of mycobacteriophages against diverse *Mycobacterium tuberculosis* strains. mBio *12*, e00973–e00921. https://doi.org/10.1128/mBio.00973-21.

77. Snapper, S.B., Melton, R.E., Mustafa, S., Kieser, T., and Jacobs, W.R., Jr. (1990). Isolation and characterization of efficient plasmid transformation mutants of *Mycobacterium smegmatis*. Mol. Microbiol. *4*, 1911–1919. https://doi.org/10.1111/j.1365-2958.1990.tb02040.x.

78. Punjani, A., Rubinstein, J.L., Fleet, D.J., and Brubaker, M.A. (2017). cryoSPARC: algorithms for rapid unsupervised cryo-EM structure determination. Nat. Methods *14*, 290–296. https://doi.org/10.1038/nmeth.4169.

79. Afonine, P.V., Poon, B.K., Read, R.J., Sobolev, O.V., Terwilliger, T.C., Urzhumtsev, A., and Adams, P.D. (2018). Real-space refinement in PHENIX for cryo-EM and crystallography. Acta Crystallogr. D Struct. Biol. *74*, 531–544. https://doi.org/10.1107/S2059798318006551.

80. Jamali, K., Käll, L., Zhang, R., Brown, A., Kimanius, D., and Scheres, S.H.W. (2024). Automated model building and protein identification in cryo-EM maps. Nature *628*, 450–457. https://doi.org/10.1038/s41586-024-07215-4.

81. Jumper, J., Evans, R., Pritzel, A., Green, T., Figurnov, M., Ronneberger, O., Tunyasuvunakool, K., Bates, R., Žídek, A., Potapenko, A., et al. (2021). Highly accurate protein structure prediction with AlphaFold. Nature *596*, 583–589. https://doi.org/10.1038/s41586-021-03819-2.

82. Abramson, J., Adler, J., Dunger, J., Evans, R., Green, T., Pritzel, A., Ronneberger, O., Willmore, L., Ballard, A.J., Bambrick, J., et al. (2024). Accurate structure prediction of biomolecular interactions with AlphaFold 3. Nature *630*, 493–500. https://doi.org/10.1038/s41586-024-07487-w.

83. Pettersen, E.F., Goddard, T.D., Huang, C.C., Meng, E.C., Couch, G.S., Croll, T.I., Morris, J.H., and Ferrin, T.E. (2021). UCSF ChimeraX: Structure visualization for researchers, educators, and developers. Protein Sci. *30*, 70–82. https://doi.org/10.1002/pro.3943.

84. Goddard, T.D., Huang, C.C., Meng, E.C., Pettersen, E.F., Couch, G.S., Morris, J.H., and Ferrin, T.E. (2018). UCSF ChimeraX: Meeting modern challenges in visualization and analysis. Protein Sci. *27*, 14–25. https://doi.org/10.1002/pro.3235.

85. Croll, T.I. (2018). ISOLDE: a physically realistic environment for model building into low-resolution electron-density maps. Acta Crystallogr. D Struct. Biol. *74*, 519–530. https://doi.org/10.1107/S2059798318002425.

86. Chen, M., Bell, J.M., Shi, X., Sun, S.Y., Wang, Z., and Ludtke, S.J. (2019). A complete data processing workflow for cryo-ET and subtomogram averaging. Nat. Methods *16*, 1161–1168. https://doi.org/10.1038/s41592-019-0591-8.

87. Mastronarde, D.N. (2005). Automated electron microscope tomography using robust prediction of specimen movements. J. Struct. Biol. *152*, 36–51. https://doi.org/10.1016/j.jsb.2005.07.007.

88. Eisenstein, F., Yanagisawa, H., Kashihara, H., Kikkawa, M., Tsukita, S., and Danev, R. (2023). Parallel cryo electron tomography on in situ lamellae. Nat. Methods *20*, 131–138. https://doi.org/10.1038/s41592-022-01690-1.

89. Winkler, H. (2007). 3D reconstruction and processing of volumetric data in cryo-electron tomography. J. Struct. Biol. *157*, 126–137. https://doi.org/10.1016/j.jsb.2006.07.014.

90. Delano, W. (2006). Pymol Home Page. pymol.sourceforge.net.

91. Kremer, J.R., Mastronarde, D.N., and McIntosh, J.R. (1996). Computer visualization of three-dimensional image data using IMOD. J. Struct. Biol. *116*, 71–76. https://doi.org/10.1006/jsbi.1996.0013.

92. Zheng, S., Wolff, G., Greenan, G., Chen, Z., Faas, F.G.A., Bárcena, M., Koster, A.J., Cheng, Y., and Agard, D.A. (2022). AreTomo: An integrated software package for automated marker-free, motion-corrected cryo-electron tomographic alignment and reconstruction. J. Struct. Biol. X *6*, 100068. https://doi.org/10.1016/j.yjsbx.2022.100068.

93. Tivol, W.F., Briegel, A., and Jensen, G.J. (2008). An Improved Cryogen for Plunge Freezing. Microsc. Microanal. *14*, 375–379. https://doi.org/10.1017/S1431927608080781.

94. Mastronarde, D.N., and Held, S.R. (2017). Automated tilt series alignment and tomographic reconstruction in IMOD. J. Struct. Biol. *197*, 102–113. https://doi.org/10.1016/j.jsb.2016.07.011.

95. Zhang, K. (2016). Gctf: Real-time CTF determination and correction. J. Struct. Biol. *193*, 1–12. https://doi.org/10.1016/j.jsb.2015.11.003.

96. Morado, D.R., Hu, B., and Liu, J. (2016). Using Tomoauto: A Protocol for High-throughput Automated Cryo-electron Tomography. J. Vis. Exp. e53608. https://doi.org/10.3791/53608.

97. Yu, R.-C., Yang, F., Zhang, H.-Y., Hou, P., Du, K., Zhu, J., Cui, N., Xu, X., Chen, Y., Li, Q., et al. (2024). Structure of the intact tail machine of Anabaena myophage A-1(L). Nat. Commun. *15*, 2654. https://doi.org/10.1038/s41467-024-47006-z.

98. van Kempen, M., Kim, S.S., Tumescheit, C., Mirdita, M., Lee, J., Gilchrist, C.L.M., Söding, J., and Steinegger, M. (2024). Fast and accurate protein structure search with Foldseek. Nat. Biotechnol. *42*, 243–246. https://doi.org/10.1038/s41587-023-01773-0.

Please cite this article in press as: Freeman et al., Structure and infection dynamics of mycobacteriophage Bxb1, Cell (2025), https://doi.org/10.1016/j.cell.2025.03.027

CellPress
OPEN ACCESS

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Bacterial and virus strains		
Mycobacterium smegmatis mc^{2}155	Snapper et al.[77]	Genbank: NC_008596
Mycobacteriophage Bxb1	Mediavilla et al.[42]	Genbank: AF271693.1
Mycobacteriophage BaDAS-1	Freeman et al.[45]	N/A
Chemicals, peptides, and recombinant proteins		
Recombinant GST-gp30 Protein	This study	N/A
Deposited data		
Cryo-EM Bxb1 Capsid Composite Map	This study	EMDB: EMD-46685
Cryo-EM Bxb1 Capsid – Icosahedral Consensus Map	This study	EMDB: EMD-46682
Cryo-EM Bxb1 Capsid – Penton C1 Local Refinement	This study	EMDB: EMD-46684
Cryo-EM Bxb1 C1 Capsid and Portal Composite Map	This study	EMDB: EMD-46681
Cryo-EM Bxb1 C1 Capsid and Portal – C1 Consensus Map	This study	EMDB: EMD-46678
Cryo-EM Bxb1 C1 Capsid and Portal – Portal C1 Local Refinement	This study	EMDB: EMD-46679
Cryo-EM Bxb1 Portal and Connector Complex Composite Map	This study	EMDB: EMD-46669
Cryo-EM Bxb1 Portal and Connector Complex – C3 Consensus Map	This study	EMDB: EMD-46663
Cryo-EM Bxb1 Portal and Connector Complex – Portal and Adaptor C12 Local Refinement	This study	EMDB: EMD-46666
Cryo-EM Bxb1 Portal and Connector Complex – Stopper and Tail Terminator C6 Local Refinement	This study	EMDB: EMD-46667
Cryo-EM Bxb1 Portal and Connector Complex – Tail Tube C3 Local Refinement	This study	EMDB: EMD-46668
Cryo-EM Bxb1 Tail Tube Segment Composite Map	This study	EMDB: EMD-46676
Cryo-EM Bxb1 Tail Tube Segment – C1 Consensus Map	This study	EMDB: EMD-46673
Cryo-EM Bxb1 Tail Tube Segment – Core C1 Local Refinement	This study	EMDB: EMD-46674
Cryo-EM Bxb1 Tail Tube Segment – CTE Domains C1 Local Refinement	This study	EMDB: EMD-46675
Cryo-EM Bxb1 Tail Tip Complex Composite Map	This study	EMDB: EMD-46661
Cryo-EM Bxb1 Tail Tip Complex – C3 Consensus Map	This study	EMDB: EMD-46656
Cryo-EM Tail Tip Complex – Tail Spike C1 Local Refinement	This study	EMDB: EMD-46657
Cryo-EM Tail Tip Complex – Tail Wings C1 Local Refinement	This study	EMDB: EMD-46658
Cryo-EM Tail Tip Complex – Tail Collar and Fibers C1 Local Refinement	This study	EMDB: EMD-46659
Cryo-EM BaDAS-1 Capsid Icosahedral Map	This study	EMDB: EMD-46687
Cryo-EM Bxb1 Capsid Hexon and Penton Model	This study	PDB: 9D9X
Cryo-EM Bxb1 C1 Capsid and Portal Model	This study	PDB: 9D9W
Cryo-EM Bxb1 Portal and Head-to-Tail Connector Model	This study	PDB: 9D94
Cryo-EM Bxb1 Tail Tube Segment Model	This study	PDB: 9D9L
Cryo-EM Bxb1 Tail Tip Complex Model	This study	PDB: 9D93
Cryo-EM Raw Micrographs for BaDAS-1	This study	EMPIAR: EMPIAR-12291
Cryo-EM Raw Micrographs for Bxb1	This study	EMPIAR: EMPIAR-12280

(*Continued on next page*)

Please cite this article in press as: Freeman et al., Structure and infection dynamics of mycobacteriophage Bxb1, Cell (2025), https://doi.org/10.1016/j.cell.2025.03.027

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Subtomogram averaging of Bxb1 phage capsid (DNA-filled) during the infection state	This study	EMDB: EMD-46702
Subtomogram averaging of Bxb1 phage capsid (emptied) during the infection state	This study	EMDB: EMD-46704
Subtomogram averaging of Bxb1 phage tail tube (DNA-filled) during the infection state	This study	EMDB: EMD-46703
Subtomogram averaging of Bxb1 phage tail tube (emptied) during the infection state	This study	EMDB: EMD-46705
Subtomogram averaging of Bxb1 phage portal (DNA-filled) during the infection state	This study	EMDB: EMD-46706
Subtomogram averaging of Bxb1 phage portal (emptied) during the infection state	This study	EMDB: EMD-46707
Subtomogram averaging of Bxb1 phage tail tip (infection state)	This study	EMDB: EMD-48066
Subtomogram averaging of Bxb1 phage tail tip (infection state, extra density within cell wall)	This study	EMDB: EMD-48067
Cryo-ET raw tilt-series	This study	EMPIAR: EMPIAR-12238
Mass spectrometry of Bxb1 particles	This study	PRIDE: PXD056154
Software and algorithms		
CryoSPARC	Punjani et al.[78]	https://cryosparc.com/
Phenix	Afonine et al.[79]	https://phenix-online.org/
Model Angelo	Jamali et al.[80]	https://github.com/3dem/model-angelo
Alphafold2	Jumper et al.[81]	https://github.com/google-deepmind/alphafold
Alphafold3	Abramson et al.[82]	https://alphafoldserver.com/
Coot	Emsley and Cowtan[61]	https://www2.mrc-lmb.cam.ac.uk/personal/pemsley/coot/
ChimeraX	Pettersen et al.[83]; Goddard et al.[84]	https://www.cgl.ucsf.edu/chimerax/
Isolde	Croll[85]	https://tristanic.github.io/isolde/static/isolde/doc/tools/gui/getting_started.html
EMAN2	Chen et al.[86]	https://blake.bcm.edu/emanwiki/EMAN2
SerialEM	Mastronarde[87]	https://bio3d.colorado.edu/SerialEM/
PACEtomo	Eisenstein et al.[88]	https://github.com/eisfabian/PACEtomo
I3	Winkler[89]	https://electrontomography.org/
PyMol	DeLano[90]	http://www.pymol.org/pymol
IMOD	Kremer et al.[91]	https://bio3d.colorado.edu/imod/
AreTomo	Zheng et al.[92]	https://msg.ucsf.edu/software

EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS

Bacterial strains and phages

Bxb1 is a mycobacteriophage infecting the bacterial strain *M. smegmatis* mc^{2}155.[77] BaDAS-1 is a previously described derivative of Bxb1 with peptides of 30 amino acids from the SARS-CoV-2 receptor binding motif appended to the C-terminal end of gp14. *M. smegmatis* mc^{2}155 was used to amplify the phages used in these studies. *M. smegmatis* mc^{2}155 was grown on Middlebrook 7H10 (Difco) solid medium and a single colony was cultured in Middlebrook 7H9 liquid medium supplemented with albumin dextrose complex (ADC) and 0.05% Tween 80; all growth steps occur at 37 °C. This initial culture was sub-cultured in 7H10 supplemented with 0.2% glucose and 0.05% Tween 80 and grown to saturation (2 – 4 days). To amplify phages, cafeteria-style trays were used as very large petri dishes and filled with 2 L each of Middlebrook 7H10 supplemented with 0.5% glycerol, 0.2% glucose, and 1 mM $CaCl_2$. On top of this medium was poured a mixture of ~1-5 x 10^5 pfu of phage, 20 mL of the saturated *M. smegmatis* sub-culture, and 200 mL of Middlebrook Top Agar (MBTA; Middlebrook 7H9, 3.5 g/L BactoAgar, 1 mM $CaCl_2$). The trays were incubated at 37 °C for 24 hours before the top agar was harvested and centrifuged to pellet agar and bacteria (8,000 rpm for 15 minutes). The clarified supernatant

was cleared of DNA (via addition of DNase I to10 μg/mL, $MgSO_4$ to 5 mM, and $CaCl_2$ to 5 mM and incubation at room temperature for 30 minutes) then filtered through a 0.2 μm membrane. The lysates were then centrifuged at 100,000 x g for 1 hour to pellet the phages, which were then resuspended into a total of 10 mL of phage buffer (10 mM Tris, pH 7.5, 10 mM $MgSO_4$, 68 mM NaCl) supplemented with 1 mM $CaCl_2$. Cesium chloride was added to the sample to 850 g/L and the sample was centrifuged at 38,000 rpm for 16 hours; a bright band of phage formed in the gradient during centrifugation and was extracted through side puncture. This band was resuspended into phage buffer with cesium chloride (850 g/L) and put through a second round of centrifugation, and the resulting band was removed via side puncture as the final sample. Before cryo-EM or cryo-ET, the phage was extensively dialyzed against a modified phage buffer (10 mM Tris, pH 7.5, 1 mM $MgSO_4$, 68 mM NaCl) supplemented with 1 mM $CaCl_2$. Phages were stored at 4 °C during and after dialysis.

METHOD DETAILS

Proteomics

Approximately 10^{11} pfu of highly purified phage samples (as described in the Experimental Models section) were pelleted by centrifugation at 13,000 rpm for 20 minutes. The supernatant was removed and the pellets were flash frozen on a dry ice / ethanol bath before freezing at -80 °C and transferring to the University of Pittsburgh Mass Spectrometry Core (https://www.msc.pitt.edu/). There, pellets were resuspended in 50 μl lysis buffer (5% SDS in 50 mM TEAB), probe sonicated and centrifuged at 13,000 x g for 10 minutes at room temperature prior to quantification of total protein by micro-BCA assay (Thermo Scientific #23235). Protein digestion was carried out on 10 μg of protein from each sample according to the S-trap (Protifi) protocol. Samples were subject to reduction with 20 mM dithiothreitol, heated to 95 °C for 10 min, cooled to room temperature, and alkylated by incubation with 40 mM iodoacetamide in darkness for 30 min at room temperature. Samples were centrifuged at 13,000 x g for 8 min, acidified with 12% phosphoric acid at 1:10 concentration, and diluted six-fold in binding buffer containing 90% methanol and 100 mM TEAB at pH 7.1. Samples were dispensed onto S-trap columns which were washed with binding buffer, centrifuged at 4, 000 x g for 1 minute, and transferred to clean tubes for incubation at 47 °C in 50 mM TEAB for trypsin digestion at 1:20 enzyme/substrate. Samples were eluted with a series of solutions including 50 mM TEAB, 0.2% formic acid, and 50% acetonitrile with 0.2% formic acid, and underwent centrifugation at 1,000 x g for 1 min after each eluent was added. Samples were then dried in a speedvac and resuspended in a solution of 3% acetonitrile and 0.1% TFA and desalted using Pierce Peptide Desalting Spin Columns (Thermo Scientific # 89851). Eluants were dried in a speedvac and resusupended in a solution of 3% acetonitrile and 0.1% formic acid to a final concentration of 0.2 μg/μL.

Mass spectrometry analysis was conducted on a Bruker timsTOF Pro2 coupled to a NanoElute 2. Approximately 200 ng of peptide was loaded onto an Aurora Ultimate C18 column (1.7 μm, 25 cm x 75 μm, IonOpticks) and eluted at 300 nl/min over a 60-minute gradient. The timsTOF Pro2 was set to PASEF scan mode and DDA with a scan range of 100–1700 m/z with 10 PASEF ramps. The TIMS was set to a 100 ms ramp and accumulation time (100% duty cycle) with a ramp rate of 9.43 Hz. Linear precursor repetitions were set to a target intensity of 20,000 and a target threshold of 2500 and active exclusion was set to 0.40 min. Collision energy was set to a base of 1.60 $1/K_0$ [V-s/cm^2] at 59 eV and 0.60 $1/K_0$ [V-s/cm^2] at 20 eV. An isolation width was set to 2 m/z for <700 m/z and 3 m/z for >800 m/z.

The collected MS data were analyzed using FragPipe v22.0 and searched against a database of the Bxb1 six-frame translation plus the *M. smegmatis* (RefSeq NC_008596) CDS regions (missed cleavages = 2; minimum peptide length = 5; precursor mass tolerance = 20 ppm; fragment mass tolerance = 20 ppm; variable modifications: Oxidation / + 15.9949 Da, N-terminal acetylation / + 41.0106 Da; fixed modifications: carbamidomethylation / +57.02146 Da). Peptide spectral matches were filtered with a FDR of 1%.

Cryo-electron micrograph collection

The highly purified phage samples described above in the "Experimental Models" section were pelleted by centrifugation at 18,000 rpm for 30 minutes and the pellets were resuspended in $1/10^{th}$ of the initial volume to yield a sample of ~10^{13} pfu/mL. For Bxb1, 3 μL of such a sample was deposited onto a freshly glow-discharged Quantifoil R2/1 grid (Quantifoil Micro Tools GmbH, Großlöbicha, Germany) and plunge-frozen with a Vitrobot Mk IV (Thermo Fisher Scientific, Waltham, Massachusetts, USA) into a 50:50 mixture of liquid ethane:propane.[93] For BaDAS-1, 5 μL of phage was deposited onto Au-flat 2/2 grid using a Vitrobot Mk IV. Grids were blotted for 5 s with a force of 5 before being plunged into liquid ethane. Data were collected on 300 keV Titan Krios microscopes (Thermo Fisher Scientific, Waltham, Massachusetts, USA) at either the University of Pittsburgh (Bxb1) or the Pacific Northwest Center for Cryo-EM (BaDAS-1) using either a Falcon 3 direct electron detector or a Gatan K3 direct electron detector, respectively. Table S1 details data collection parameters for each phage.

Cryo-EM reconstructions and model building

The workflows for cryo-EM reconstructions are described in detail in Figures S1 and S3. In general, we used motion correction and single particle analysis (SPA) in CryoSPARC[78] to reconstruct phage virion structures. The capsids structures of Bxb1 and BaDAS-1 were readily reconstructed to < 2.7 Å resolution using icosahedral symmetry, and the 5-fold vertex of Bxb1 was separately refined without symmetry to resolve the flexible CTE domains (Figures 2, S3, and S7; Table S1). The head-to-tail connector region was resolved from the BaDAS-1 dataset using a symmetry-expanded capsid particle dataset which was subjected to rounds of masked

Please cite this article in press as: Freeman et al., Structure and infection dynamics of mycobacteriophage Bxb1, Cell (2025), https://doi.org/10.1016/j.cell.2025.03.027

3D classification to identify the portal vertex, isolate the best of the five symmetry-related portal classes, and isolate complete portal-to-tail tube connector complexes. These structures were further refined with particle subtraction and local refinements both in C1 and with appropriate symmetry imposed (Figures 3 and S1; Table S1). To reconstruct the tail tube, particles of 200 Å (containing ~4 rings of the tail tube hexamer) were extracted and subjected to extensive 2D classifications, then refined in C1 with non-uniform refinement, 3D classification, reference-based motion correction, particle subtraction, and separate local refinements of the core and CTE domains (Figures 4 and S1; Table S1). Finally, the tail tip was reconstructed by picking and isolating tail tip particles with exhaustive 2D and 3D classifications and performing non-uniform refinements of the entire complex in C3 followed by reference-based motion correction. To resolve the tail spike, tail tip wings, and tail collar structures, we symmetry expanded the dataset of the tail tip complex and performed particle subtraction and asymmetric local refinements (Figures 5 and S1; Table S1). For each map, the gold-standard method was used to estimate the resolution with an FSC (Fourier Shell Correlation) cutoff of 0.143. ChimeraX[83] was used to create composite maps of all the local reconstructions and these regional maps were assembled together into a composite density map of the complete Bxb1 particle (Figure 1D). Each local map was autosharpened with Phenix[79] and then ModelAngelo,[80] Alphafold2,[81] Coot,[61] and Isolde[85] facilitated model building into the cryo-EM density maps. Specifically, ModelAngelo was used to make primary assignments of protein subunits into the density and to establish stoichiometry. For fragmented chain segments that were not built well by ModelAngleo, Alphafold-predicted models were aligned to help assign model segments to appropriate chains. These segments were manually stitched together in Coot to form rough models, which were then refined using Coot to build missing residues and improve fit, Phenix to improve geometry and clash scores, and Isolde to relax the refined models into the density and polish geometry and clash metrics. Finally, Phenix was used to validate the models and map-model correlation.

Cryo-electron tomograph collection

The *M. smegmatis* mc^2 155 strain was grown to an OD_{600} of ~1 in liquid 7H9 medium with 1 mM $CaCl_2$ at 37 °C with shaking at 225 rpm. An appropriate volume of highly purified Bxb1 phages, sufficient to infect *M. smegmatis* cells at an MOI of 300, was added to the bacterial culture. The mixture was incubated on ice for 10 minutes to enhance phage binding, followed by 30 minutes of incubation at 37 °C with shaking at 200 rpm. The infection culture was concentrated to an OD_{600} of 6. 5 μl of the concentrated samples were deposited onto freshly glow-discharged (Pelco easiGlow; 25 seconds glow at 15 mA) Quantifoil R2/1 copper 200 mesh grids for 1 minute. The grids were then back-side blotted with Whatman Grade 1 filter paper and frozen in a liquid ethane and propane mixture using a gravity-driven homemade plunger apparatus at room temperature (Figure S2A). The vitrified grids were later clipped with cryo-FIB autogrids (Thermo Fisher Scientific) prior to milling.

Cryo-FIB milling was performed using the Aquilos 2 cryo-FIB/SEM (Thermo Fisher Scientific). The surface of the grid was sputtered with metallic platinum (20 mA, 10 pA, 10–15 seconds), followed by a coat of organo-platinum using a gas injection system (GIS) (40–45 seconds), and finally, a last layer of metallic platinum (20 mA, 10 pA, 10–15 seconds). After selecting milling positions in MAPS software (Thermo Fisher Scientific), the automated cryo-FIB milling software AutoTEM (Thermo Fisher Scientific) was used to mill multiple sites at a target milling angle of 8°. First, stress relief cuts (1 μm × 10 μm × 6 μm) were made 4 μm away from the lamella, which is ~11 μm wide, using a milling current of 0.3 nA. Rough milling then reduced the lamella to 1 μm thick using a rectangular pattern on both sides simultaneously with a milling current of 0.3 nA. Medium milling further reduced the lamella to 500 nm using the cleaning cross-section pattern on each side alternately with a milling current of 0.1 nA. Fine milling thinned the lamella to 300 nm using the same pattern as medium milling with a milling current of 50 pA. Each lamella was then manually polished to less than 200 nm using a rectangular pattern and a milling current of 30 pA (Figure S2A). Post-milling, the lamellae were sputtered with metallic platinum to create small beads for tilt-series alignment (10 mA, 10 pA, 20–25 seconds).

The lamellae were imaged using a 300 keV TEM microscope, Titan Krios (Thermo Fisher Scientific), equipped with a field emission gun, an energy filter (slit width of 20 eV), and a direct detection device (Gatan K3). Tilt-series data were collected using the SerialEM[87] data collection package with PACEtomo[88] scripts, with 35 image stacks from +51° to −51° (3° increments) using a dose-symmetric scheme, resulting in a total electron dose of ~100 $e^-/Å^2$ (Figure S2A). Data were collected at a magnification yielding a pixel size of 1.665 Å/pixel and a nominal defocus of ~-5 μm. Image stacks were motion-corrected using MotionCor2 and assembled into drift-corrected stacks using IMOD.[91] The drift-corrected stacks were aligned using both marker-dependent alignment in IMOD[94] (for images with sufficient beads) and patch-track alignment in AreTomo2[92] (for images with insufficient beads).

Cryo-ET subtomogram averaging

Subtomogram averaging of the capsid, portal, and tail filament was performed using the EMAN2[86] data processing pipeline. Both full and empty capsids were selected in bin 4 tomograms using the convolutional neural network (CNN)-based particle picker module. Initial classification without alignment was performed to separate full and empty capsid particles. Bin 4 particles were first aligned in C1 until icosahedral features emerged, and they were further refined with icosahedral symmetry applied. The aligned bin 4 particles were then re-extracted in bin 1 for further refinement with icosahedral symmetry applied. DNA-filled and empty tail filaments were picked separately in a semi-manual manner using the filament tracing module, and particles were extracted with a curve overlap of 0.75. Bin 4 particles were aligned in C1. The resolution of the tail filament maps was gradually improved by reducing the binning factor from 4 to 2 to 1. To average the portal, DNA-filled and empty capsids that clearly showed the capsid-to-bacterial membrane connection were manually selected. Bin 4 particles were aligned in C1 until C3 symmetry features emerged. The center of the

particles was then shifted to the center of the portal and further refined using bin 2, followed by bin 1 particles, with C3 symmetry applied. The gold-standard method was used to estimate the resolution with an FSC cutoff of 0.143 (Figure S2D).

To acquire the structure of the tail-tip complex, CTF estimation of tilt images was performed using gCTF,[95] and CTF correction was done using IMOD ctfphaseflip prior to reconstructing tomograms. Phage tails that directly interact with the bacterial membrane were manually selected with pre-defined orientation using TomoPick.[96] Subtomograms were averaged using i3 (0.9.9.3)[89] software.

Bin 4 particles were initially aligned in C1, and C3 symmetry was applied once it emerged. Final refinement was performed using bin 2 particles due to the limited number of particles and to maximize contrast. Even/odd particles were aligned against the same reference, and the resolution was estimated. To minimize potential overfitting, we low-pass filtered our reference and estimated the resolution in a moderate way based on an FSC cutoff of 0.5 (Figure S2D).

Subtomogram average maps were visualized and segmented using ChimeraX.[84] Rigid body fitting of the capsid and gp23 models to the corresponding subtomogram average maps was done using ChimeraX and PyMol.[90]

Cloning, expression, and purification of Bxb1 gp30 FN3 domain

The FN3 domain of Bxb1 gp30 (amino acids 1–57), fused with an N-terminal His-GST tag was cloned into a modified pET28 vector with a TEV cutting site between the tag and desired protein domain. The plasmid was subsequently transformed into *Escherichia coli* BL21 (DE3) cells, which were grown at 37 °C in LB medium supplemented with 50 mg/L Kanamycin until the OD_{600} reached ~0.8. The protein was overexpressed by the addition of IPTG to a final concentration of 0.5 mM. Post induction incubation overnight at 16 °C, the culture was harvested by centrifugation at 6,000 rpm for 30 min. Cell pellets were re-suspended in lysis buffer (50 mM Tris-HCl, 500 mM NaCl, 10 mM imidazole, 10 mM β-mercaptoethanol, 10 % glycerol, pH 7.4). The suspension was lysed by sonication and centrifuged at 15,000 rpm at 4 °C for 1 hour. The supernatant was loaded onto a nickel-affinity column pre-equilibrated with wash buffer (50 mM Tris at pH 7.4 and 500 mM NaCl). The protein was washed with wash buffer with various concentrations of imidazole and finally eluted with the wash buffer in the range of 50–500 mM imidazole. The eluted protein was further purified by size-exclusion chromatography using the buffer (50 mM Tris-HCl, 200 mM NaCl, 2% glycerol, pH 7.4). The protein was concentrated to 10 mg/ml and stored under –80 °C for further use.

Glycan binding using the glycan microarray

Glycan microarray slides[64] were pre-hydrated with TMST buffer [20 mM Tris-base (pH 7.4), 0.15 mM NaCl, 2 mM $CaCl_2$, 2 mM $MgCl_2$, 0.05% Tween 20] for 5 min at room temperature, then blocked with TMST–BSA buffer [20 mM Tris-base (pH 7.4), 0.15 mM NaCl, 2 mM $CaCl_2$, 2 mM $MgCl_2$, 0.05% Tween 20, 1% BSA] at 37 °C for 30 min. After blocking, the Bxb1 gp30 FN3 domain sample (100 μL, 100 μg/mL in TMST–BSA buffer) was loaded into the wells of the slide. The wells were covered with aluminum adhesive seal and incubated for 2 h at room temperature with shaking at 150 rpm to allow for protein binding. The protein sample were pipetted out, and the slides were washed three times with TMST buffer. The primary antibody (100 μL of anti-6X his epitope tag mouse monoclonal IgG1 (kappa) antibody) was loaded into the wells. Wells were then sealed with adhesive tape then incubated for 45 min at room temperature with shaking at 150 rpm. The antibody was pipetted out, and the slides were washed three times with TMST buffer. The secondary antibody (100 μL of Cy3-conjugated rabbit anti-mouse IgG1 (Gamma 1 chain) antibody) was then loaded into the wells. The wells were then sealed with adhesive tape then incubated for 45 min at room temperature with shaking at 150 rpm. The antibody was pipetted out, and the slides were washed sequentially with TMST buffer, followed by TMS buffer [20 mM Tris-base (pH 7.4), 0.15 mM NaCl, 2 mM $CaCl_2$, 2 mM $MgCl_2$]. Prior to scanning, slides were dried by centrifugation.

Microarray Scan and Data Analysis

Glycan microarrays were scanned at 10 μM pixel size using GenePix® 4000B Molecular Devices Scanner. The fluorescent signal from the Cy3 dye was detected at 532 nm. Laser power was set at 100% and photomultiplier tube gain was set at 500. Pixel density (intensity) of each spot was quantified using GenePix® Pro 7 Software. The fluorescence intensities for each spot and the local background were calculated. Relative fluorescence units were computed by subtracting the local background signal from the signal of each spot. The mean of triplicate readings was calculated using Microsoft Excel.

CellPress
OPEN ACCESS

Supplemental figures

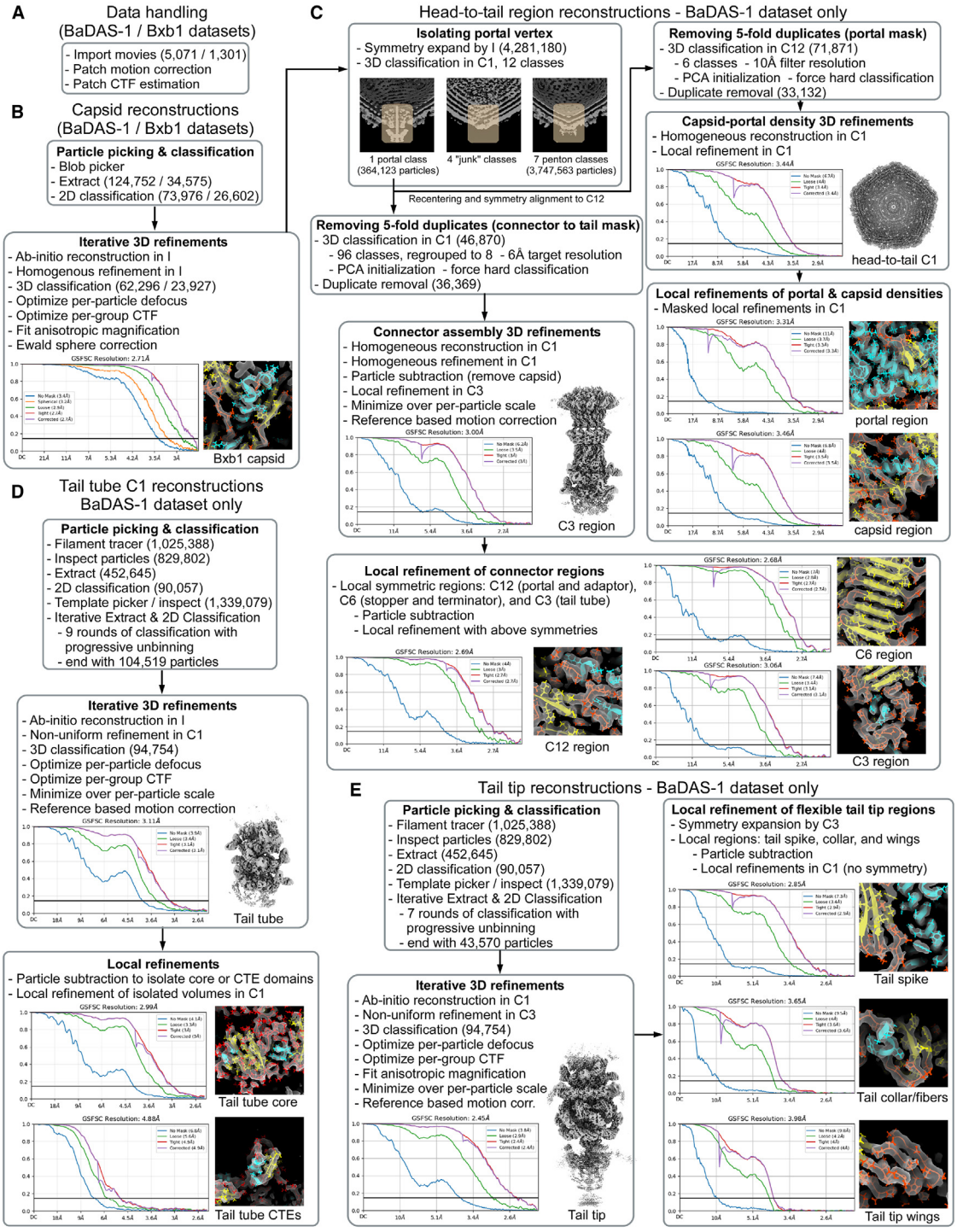


Figure S1. Workflow of cryo-EM reconstructions, related to Figures 1, 2, 3, 4, and 5

(A) Bxb1 and BaDAS-1 micrographs were imported into CryoSPARC,[78] where they were motion corrected, and the contrast transfer function was estimated. (B) Bxb1 and BaDAS-1 capsids were isolated with blob picker and 2D classification, then reconstructed with homogeneous refinement with icosahedral symmetry imposed. The gold-standard Fourier shell correlation (GSFSC) curve is shown for the Bxb1 capsid reonstruction. A zoomed-in map segment is

(legend continued on next page)

displayed, with a model showing regions of helices (teal), beta sheets (yellow), and loops (red). This format is preserved throughout the figure for each reconstruction.

(C) The head-to-tail region was reconstructed from symmetry-expanded BaDAS-1 capsid particles. These were subjected to iterative 3D classifications to find the portal vertex, isolate the capsid-portal complex in C1, and isolate the portal-adaptor-stopper-tail-terminator-tail tube complex in C3. These particle classes were homogeneously reconstructed and then locally refined in C1 (capsid-portal region) or C3 (connector region). Further masked local refinements of isolated regions followed the global refinements with no symmetry (capsid-portal regions) or C12, C6, or C3 symmetry (connector region) imposed, as shown.

(D) Tail tube particles were isolated with filament tracer to generate initial 2D classes and then template picker to isolate higher quality particles. These were refined with extensive 2D classification and 3D non-uniform refinements, followed by reference-based motion correction and local refinement of the tail tube core and tail tube CTE domain regions.

(E) Tail tips were isolated from the same template picker job as tail tubes but from a different set of 2D classes. These were refined in 3D with non-uniform refinement in C3 followed by reference-based motion correction. The resulting particles were symmetry expanded and subjected to three separate particle subtraction jobs to isolate the tail spike, tail collar and fibers, and tail-tip wing densities, then each region was locally refined in C1.

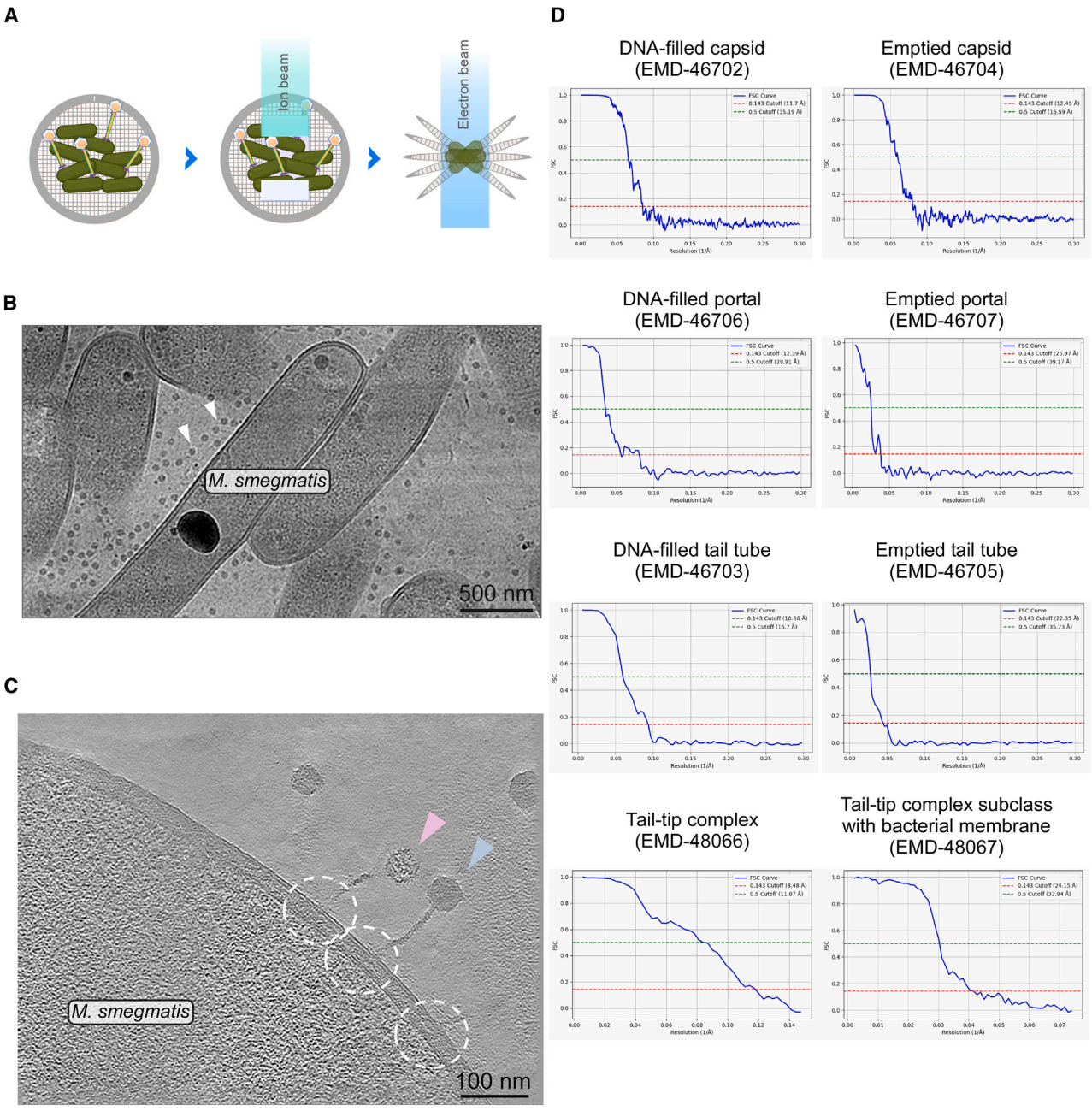


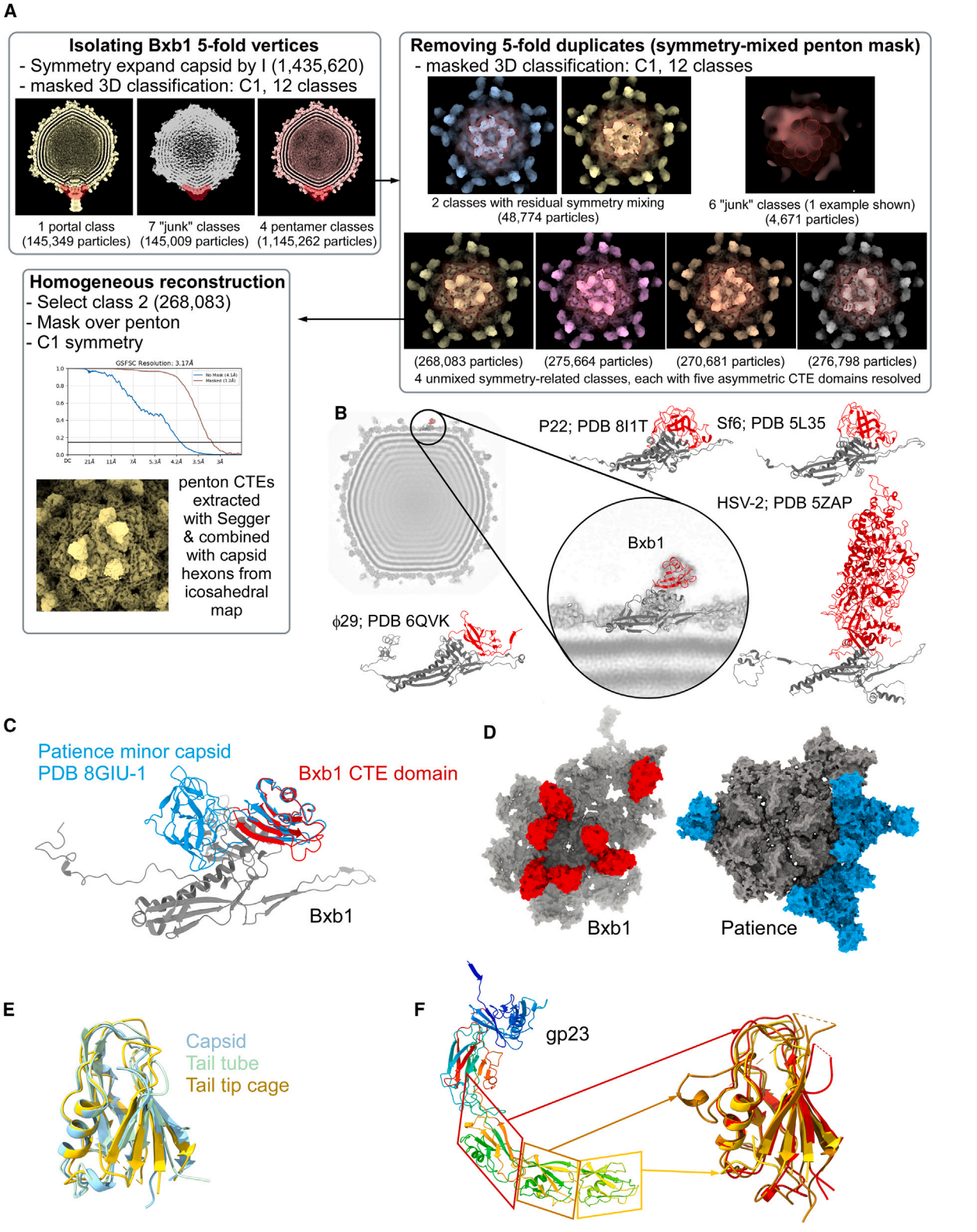
Figure S2. Cryo-ET pipeline for imaging Bxb1-*M. smegmatis* interactions, related to Figures 1, 2, 3, 4, 5, and 6

(A) Cartoon model depicting the workflow. *M. smegmatis* is infected with Bxb1 at an MOI of 300 for 30 min, then applied to cryo-EM grids for vitrification. The frozen hydrated samples are milled to <200 nm thickness using a cryofocused ion beam (cryo-FIB) milling machine. Tilt series are collected, and cryo-lamellae are imaged under a 300 keV electron microscope. Images are reconstructed into 3D tomograms and analyzed computationally.

(B) Low magnification (3,600×) of the lamella showing multiple *M. smegmatis* cells infected with Bxb1. White arrows point to representative Bxb1 particles.

(C) Representative high magnification (51,000×) tomogram showing Bxb1 infecting *M. smegmatis*. Pink and teal arrowheads point to representative empty and full capsids, respectively. White circles highlight areas where the phage tail is inserted into the bacterial outer membrane.

(D) FSC curves of all the subtomogram averages. The gold-standard method (FSC = 0.143) was used to estimate the resolutions of volumes aligned with EMAN2 software (EMD-46702, EMD-46703, EMD-46704, EMD-46705, EMD-46706, and EMD-46707). The reference supplementary method (FSC = 0.5) was used to estimate the resolutions of volumes aligned with i3 software (EMD-48066 and EMD-48067).



(legend on next page)

Figure S3. The CTE domains on Bxb1's capsid, tail, and tail-tip cage, related to Figures 2, 4, and 5

(A) Workflow for Bxb1 penton asymmetric reconstruction and refinement. The GSFSC is shown for the final map, along with a zoomed-in view of the penton.

(B) Bxb1 capsid in gray with a blow-up of the Bxb1 gp14 model in the density. The capsid floor domains are colored gray and the CTE domain is colored red. Around the capsid and inset are capsid subunits from other viruses with domains positioned analogously to the Bxb1 CTE.

(C) Alignment of the CTE domain (red) of Bxb1 g14 (gray) with the minor capsid protein from mycobacteriophage Patience (blue).

(D) Position of Bxb1's CTE domain (gray and red, left) and the related minor capsid protein in Patience (gray and blue, right).

(E) The CTE domains of the Bxb1 capsid, tail tube, and tail-tip cage aligned.

(F) Bxb1 gp23 subunit colored in rainbow from N to C termini, with boxes around the CTE-like domains (left). The boxed domains are shown color-coded and aligned to the right.[82,97]

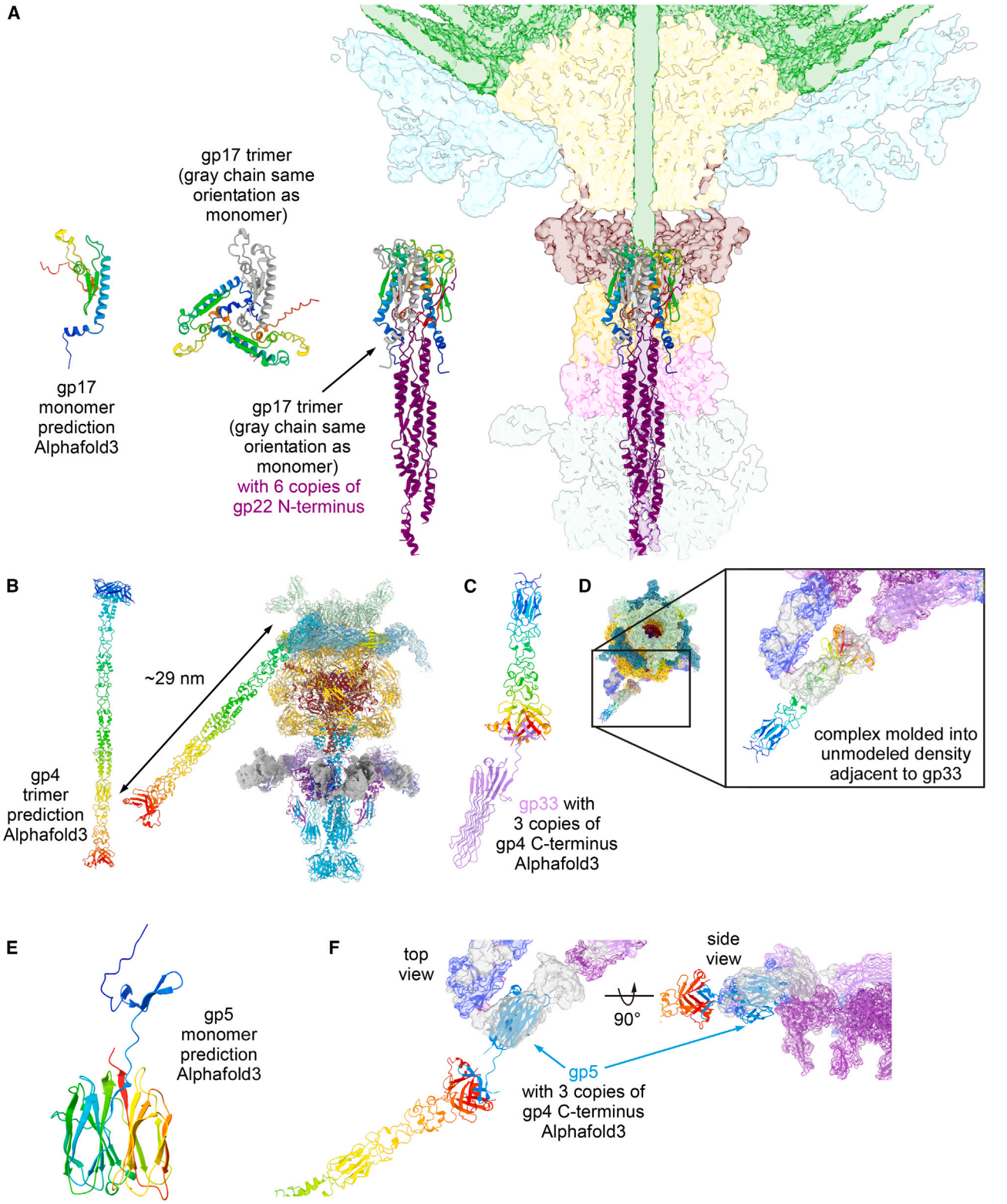


Figure S4. Alphafold models for gp5 and gp17 and their fit into density, related to Figures 3 and 5

(A) Alphafold3 predicted models of gp17 as a monomer or trimer, or as a trimer folded with a hexamer of the N-terminal end of the TMP, gp22, as labeled. Although we also used Alphafold3 to predict the model of gp17 with a gp22 N-terminal trimer, this structure looked very different and did not fit into the unmodeled cryo-EM

(legend continued on next page)

density at all; thus, we show the gp17-TMP hexamer model. This is not entirely unconventional, as other studies have reported six copies of the TMP at the N-terminal end near the portal but resolved just three copies of the TMP C terminus within the tail-tip complex.[97] To the right, the 3Xgp17-6Xgp22 multimer complex is shown in the density of the head-tail connector region, indicating that the dimensions are approximately correct for gp17 to be potentially enclosed within the connector tube.

(B) Alphafold3-predicted model of the gp4 trimeric fiber structure (left) and this 29 nm model aligned to the built model of gp4 in the tail-tip model (right).

(C) Alphafold3-predicted model of gp33 interacting with the C-terminal segment of the gp4 trimeric fibers.

(D) Fit of the multimer model shown in (C) into the unmodeled density of the tail wings.

(E) Monomer of Bxb1 gp5, as predicted by Alphafold3.

(F) An Alphafold3-predicted multimer complex of gp4 trimeric fibers bound to gp5 is shown fit into the unmodeled density of the Bxb1 tail wings. The view is shown from the top and from the side, showing that gp5 is not the same shape as the unmodeled density.

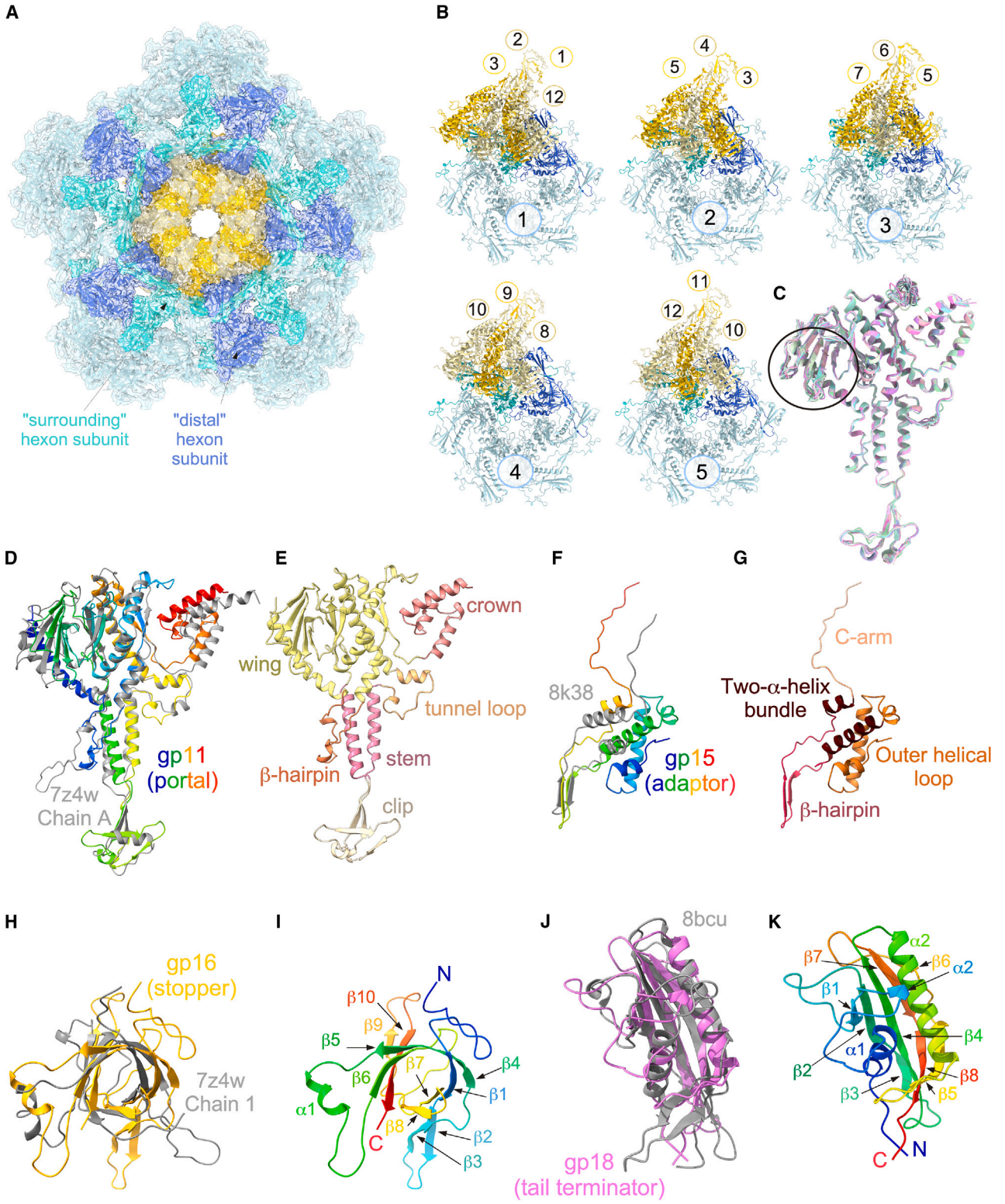
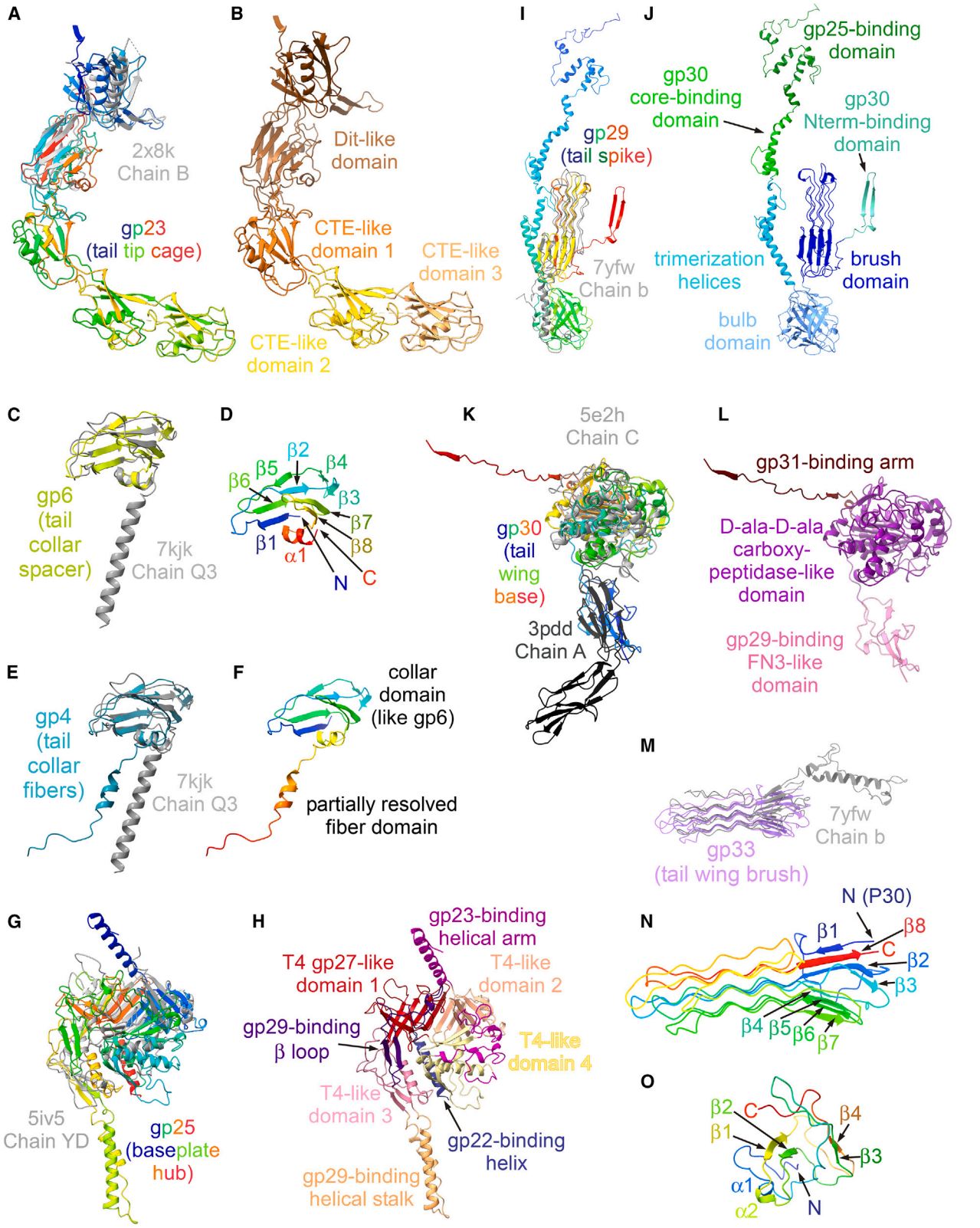


Figure S5. Further details on the 5- to 12-fold transition and connector assembly proteins, related to Figure 3

(A) The C1 capsid-portal map shown from the outside (looking up toward the capsid from the tail tube), with "surrounding" and "distal" capsid subunits indicated with labels and arrows.

(legend continued on next page)

(B) Portal-adjacent hexons and their hydrogen-bonded portal subunits were extracted and aligned. Each complex is shown individually and labeled consistently with Figure 3B.
(C) The 12 portal subunits aligned and shown in different colors, and the area of remodeling in the wing domain is circled.
(D) The portal protein, gp11, is shown in rainbow from N to C termini and aligned with its Foldseek[98] hit 7z4w in gray.
(E) The gp11 model is colored and labeled according to domains.
(F) The adaptor protein, gp15, is shown in rainbow from N to C termini and aligned with its Foldseek hit 8k38 in gray.
(G) The gp15 model is colored and labeled according to domains.
(H) The stopper protein, gp16, is shown in yellow and aligned with its Foldseek hit 7z4w in gray.
(I) The gp16 model is colored in rainbow from N to C termini with all secondary structure features labeled.
(J) The tail terminator protein, gp18, is shown in pink and aligned with its Foldseek hit 7z4w in gray.
(K) The gp18 model is colored in rainbow from N to C termini with all secondary structure features labeled.



(legend on next page)

Figure S6. Further details on the tail-tip proteins, related to Figure 5

(A) The tail-tip cage protein, gp23, is shown in rainbow from N to C termini and aligned with its Foldseek hit 2x8k in gray.
(B) The gp23 model is colored and labeled according to domains.
(C) The tail collar spacer, gp6, is shown in yellow-green and aligned with its Foldseek hit 7kjk in gray.
(D) The gp6 model is colored in rainbow from N to C termini with all secondary structure features labeled.
(E) The tail collar fiber, gp4, is shown in yellow-green and aligned with its Foldseek hit 7kjk in gray.
(F) The gp4 model is colored in rainbow from N to C termini with all secondary structure features labeled.
(G) The base plate hub protein, gp25, is shown in rainbow from N to C termini and aligned with its Foldseek hit 5iv5 in gray.
(H) The gp25 model is colored and labeled according to domains.
(I) The tail spike protein, gp29, is shown in rainbow from N to C termini and aligned with its Foldseek hit 7yfw in gray.
(J) The gp29 model is colored and labeled according to domains.
(K) The tail wing base protein, gp30, is shown in rainbow from N to C termini and aligned with its Foldseek hit 3pdd in gray.
(L) The gp30 model is colored and labeled according to domains.
(M) The tail wing brush, gp33, is shown in lavender and aligned with its Foldseek hit 7yfw in gray.
(N) The gp33 model is colored in rainbow from N to C termini with all secondary structure features labeled.
(O) The tail wing arm protein, gp33, model is colored in rainbow from N to C termini with all secondary structure features labeled.

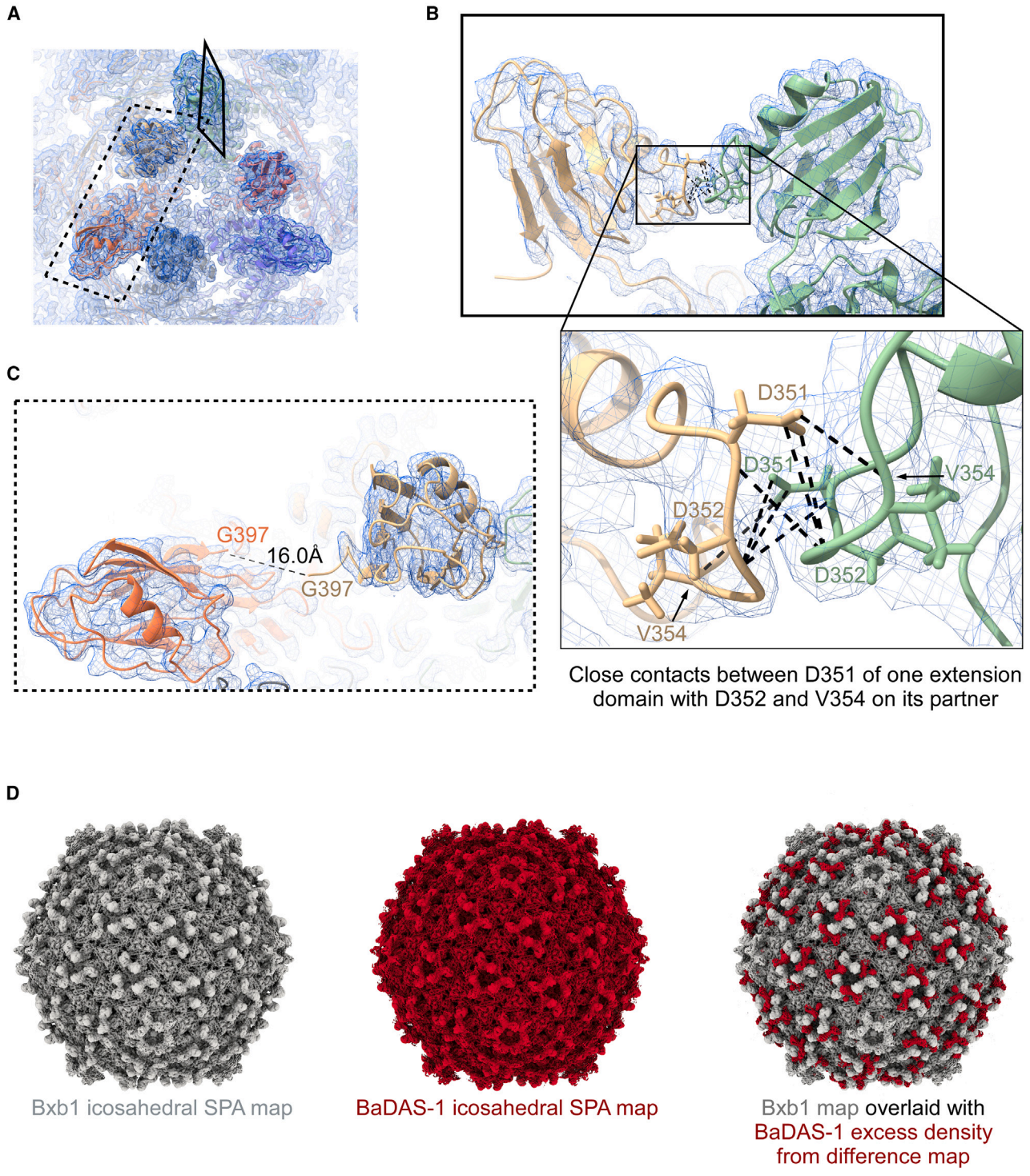


Figure S7. Structural limitations of the Bxb1 capsid as a vaccine platform, related to Figure 2

(A) An overview of the Bxb1 hexon, with cryo-EM density shown as a blue mesh and six capsid protein subunits represented as colored cartoons. Solid and dashed boxes indicate viewpoints for (B) and (C), respectively.

(B) A side view of the interface between two interacting extension domains. The area within the interior box is shown zoomed in (top right) to illustrate close contacts between neighboring chains.

(C) The view from above two neighboring, non-interacting subunits, with the terminal residue G396 shown for each chain.

(D) The Bxb1 capsid icosahedral SPA map in gray (left), the BaDAS-1 capsid icosahedral SPA map in red (middle), and an overlay (right) of the Bxb1 capsid (gray) with the excess density (generated via map subtraction in ChimeraX[84]) from BaDAS-1 icosahedral capsid (red).