PNAS RESEARCH ARTICLE | MICROBIOLOGY

Structural basis for Lamassu-based antiviral immunity and its evolution from DNA repair machinery

Matthieu Haudiquet[a,b,1], Arpita Chakravarti[c,1], Zhiying Zhang[c,1], Josephine L. Ramirez[d], Alba Herrero del Valle[a], Paul Dominic B. Olinares[e], Rachel Lavenir[a], Massilia Aït Ahmed[a], M. Jason de la Cruz[c], Brian T. Chait[e], Samuel H. Sternberg[d,f], Aude Bernheim[a,2], and Dinshaw J. Patel[c,2]

Affiliations are included on p. 12.

Edited by Anthony Kossiakoff, The University of Chicago, Chicago, IL; received July 21, 2025; accepted October 16, 2025

Bacterial immune systems exhibit remarkable diversity and modularity, as a consequence of the continuous selective pressures imposed by phage predation. Despite recent mechanistic advances, the evolutionary origins of many antiphage immune systems remain elusive, especially for those that encode homologs of the structural maintenance of chromosomes (SMC) superfamily, which are essential for chromosome maintenance and DNA repair across domains of life. Here, we elucidate the structural basis and evolutionary emergence of Lamassu, a bacterial immune system family featuring diverse effectors but a core conserved SMC-like sensor. Using cryo-EM, we determined structures of the *Vibrio cholerae* Lamassu complex in both apo- and dsDNA-bound states, revealing unexpected stoichiometry and topological architectures. We further demonstrate how Lamassu specifically senses dsDNA ends in vitro and phage replication origins in vivo, thereby triggering the formation of LmuA tetramers that activate its Cap4 nuclease domain. Our findings reveal that Lamassu evolved via exaptation of the bacterial Rad50-Mre11 DNA repair system to form a compact, modular sensor for viral replication, exemplifying how cellular machinery can be co-opted for novel immune functions.

bacterial immunity | SMC-like proteins | cryo-EM structures | phages | evolutionary repurposing

Defense systems are highly diverse, reflecting the extensive and ongoing phage-bacteria coevolution (1, 2). While there has been continued progress in the deciphering of molecular mechanisms of bacterial immune systems, the emergence and evolution of some of these immune systems have remained elusive. Evolutionary studies have until now primarily focused on CRISPR-Cas systems, uncovering their evolutionary origins in transposable elements machinery or toxin–antitoxin systems (3, 4), but the evolutionary origins of hundreds of other identified antiphage systems remain unclear.

Structural maintenance of chromosome (SMC) proteins are a conserved family of ATPases critical for DNA processing across all domains of life (5, 6). They mediate distinct processes, such as condensation and organization of chromatin, sister chromatid cohesion, DNA repair, and innate immunity (5–7). A widespread and highly diverse defense system encoding an SMC-like core component is the Lamassu family (8–11). It comprises LmuB, an SMC-like protein, LmuC, a small protein with a domain of unknown function, and LmuA, which encodes a range of diverse effectors that include Cap4 nucleases. The extensive diversity of LmuA effectors contributes to the functional versatility of the Lamassu family. Lamassu from the El Tor *Vibrio cholerae* strain (also known as DNA defense module DdmABC), the most studied Lamassu homolog, is thought to be triggered by palindromic sequences from phages and plasmids leading to cell death, thus protecting the bacterial population from further infection (8, 9).

From a structural perspective, a number of structural modeling and domain analysis of the Lamassu components have revealed the presence of a SMC scaffold for the LmuB subunit (8), whose coiled coils are approximately half of the length (~20 nm) of conventional SMC proteins (12), such as condensins (7), cohesins (13–15), Smc5/6 (16), MukBEF (17, 18), and Wadjet JetC (19). This noncanonical feature of the Lamassu complex might play a key role in its immune activity. The functional implications of this shorter coiled-coil and its influence on Lamassu complex assembly and specificity, remain unclear. How Lamassu complexes detect foreign DNA, trigger effector activation, and evolve their modular architecture remains unknown—posing fundamental questions about the structural and evolutionary principles underlying this widespread antiphage system.

Here we set out to understand the phylogeny and structural basis of Lamassu complex formation and activation and how these structural features illuminate Lamassu evolution across species and its mechanism of action.

Significance

Lamassu is a widespread antiviral system in bacteria that uses structural maintenance of chromosomes-like proteins, typically associated with chromosome maintenance, to detect and halt phage infection. By solving cryo-EM structures of the Lamassu complex in apo, dsDNA-bound, and activated states, we reveal how it senses viral DNA ends and activates its nuclease effector through oligomerization. Phylogenetic and structural analyses further show that Lamassu evolved from the SbcCD DNA repair machinery, repurposing genome surveillance tools for immune defense. These findings uncover a molecular mechanism by which bacteria distinguish viral from host DNA and illuminate a broader evolutionary principle: the transformation of essential cellular machinery into specialized immune systems.

Author contributions: M.H., A.C., Z.Z., J.L.R., S.H.S., A.B., and D.J.P. designed research; M.H., A.C., Z.Z., J.L.R., P.D.B.O., and M.A.A. performed research; M.H., A.C., Z.Z., and J.L.R. contributed new reagents/analytic tools; M.H., A.C., Z.Z., J.L.R., A.H.d.V., P.D.B.O., R.L., M.A.A., M.J.d.l.C., B.T.C., S.H.S., A.B., and D.J.P. analyzed data; and M.H., A.C., Z.Z., J.L.R., S.H.S., A.B., and D.J.P. wrote the paper.

The authors declare no competing interest.

This article is a PNAS Direct Submission.

[1]M.H., A.C., and Z.Z. contributed equally to this work.

[2]To whom correspondence may be addressed. Email: aude.bernheim@pasteur.fr or pateld@mskcc.org.

This article contains supporting information online at https://www.pnas.org/lookup/suppl/doi:10.1073/pnas.2519643122/-/DCSupplemental.

Published November 18, 2025.

Results

Short Lamassus Are Highly Diverse Antiphage Systems with Coevolving Components. We first set out to comprehensively identify Lamassu systems in prokaryotic genomes. We assembled a dataset incorporating experimentally validated Lamassu systems from previous studies (8, 9, 12, 20), as well as predicted systems (11, 20). We identified Lamassu systems through an iterative approach, aiming to uncover the complete diversity of the system in public genomes (*Materials and Methods*, Fig. 1*A*). This approach yielded 3,829 Lamassu systems across 22,920 complete genomes from the RefSeq database, 12% (n = 430) more than previous detections on the same genomes. We found Lamassu systems in 3,544 genomes, including seven archaeal genomes (*Euryarchaeota* and *Thaumarchaeota*) indicating that 15.4% of prokaryotic genomes encode at least one Lamassu system, with some species such as *Photorhabdus laumondii* encoding up to five distinct Lamassu systems.

To gain a deeper understanding of the diversity of the family, we focused on LmuB, as it is conserved across all Lamassu systems. Phylogenetic analysis of LmuB uncovered two major clades, which were mainly distinguished by the length of coiled-coil regions: short LmuB proteins averaged 600 amino acids (total protein length), whereas long LmuB proteins measured ~800 amino acids (Fig. 1 *B* and *C*). 3D structure comparison of LmuB AlphaFold 3 (AF3) models also corroborated this dichotomy (*SI Appendix*, Fig. S1), and both types were associated with defensive environments, with defense scores of 27% (long) and 25% (short; see *Materials and Methods*). Long and short Lamassus differed in abundance, with 3,293 systems classified as short and 536 as long. Previous experimentally validated systems belong to the short clade (8–10), except one system from *Bacillus cereus* B4077, encoding Hydrolase-Protease effectors (20). Long and short LmuB were associated with different LmuA and LmuC gene families, suggesting that the dichotomy extends to LmuB's partners. Notably, LmuC and the CTD of LmuA that are associated with long LmuB were also longer than the ones associated with short LmuB, which was also evident from a 3D structure comparison (*SI Appendix* Fig. S1). These findings further underscore the presence of two distinct Lamassu families

We next examined the conservation and diversity of Lamassu features between long and short Lamassus components. LmuB from both types displayed a prototypical Walker A (Gx4GK[S/T]), but their Walker B motifs differed: long Lamassu had the consensus sequence hhhhDD, whereas short Lamassu featured hhhh-D[Q/S/T/E] (*SI Appendix*, Fig. S2). This variation may modulate the activity of the ATPase, for example E→Q mutants can result in efficient ATP binding but decreased ATP hydrolysis (21). The putative dimerization module of LmuB, the CxxC zinc-hook, exhibited extensive variation across the LmuB phylogeny. Long LmuB consistently encoded a single copy, while short LmuB encoded either zero, one, or two copies (Figs. 1 *A* and *B* and *SI Appendix*, Fig. S2). This suggests that short LmuB may employ divergent dimerization strategies compared to other SMC-like proteins.

In terms of effectors, Lamassu is exceptionally diverse, with at least nine distinct domains found in the N-terminal region of LmuA (Fig. 1*A*). The repertoire of effectors is bigger in LmuAs associated with short (n = 8) Lamassus when compared to long ones (n = 4). Some effectors are shared between the two types, including Cap4,

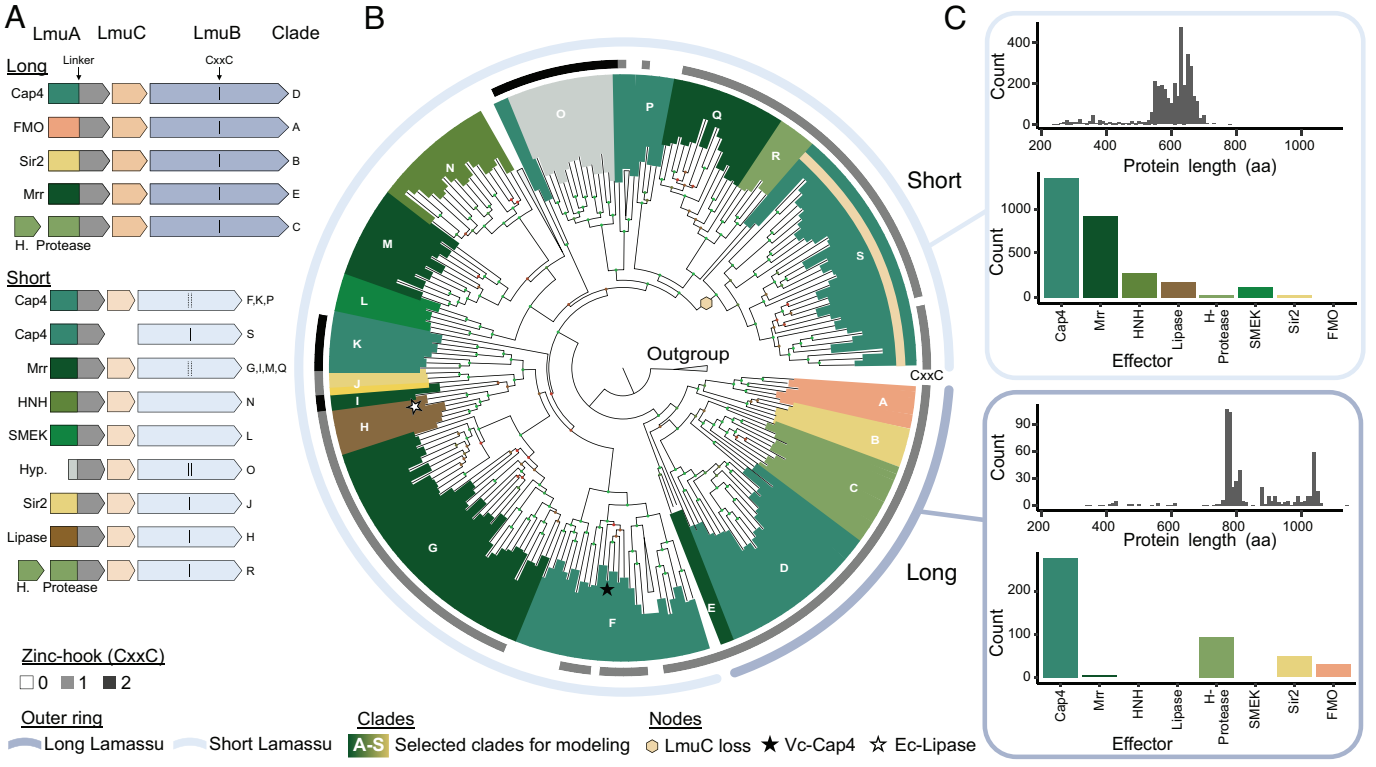

Fig. 1. Short Lamassus are highly diverse and modular antiphage systems. (*A*) Types and subtypes of Lamassu identified in DefenseFinder. Representative operons illustrate modularity, highlighting LmuA effector domains (colored) and C-terminal domain (CTD) (dark gray). LmuC variants are shown in beige. LmuB variants feature zinc-hook motifs (CxxC); solid black lines indicate all members of the represented subtype have the motif, while dotted lines indicate presence of the CxxC only in some members. Clade labels correspond to those in panel *B*. (*B*) Phylogeny of LmuB proteins based on whole-protein alignments of representative sequences clustered at 40% identity and coverage (n = 270). Inner colors at branch tips correspond to LmuA effector domains as in panel *A*. The beige hexagon indicates hypothesized loss of LmuC in clade S; the black star marks the studied Lamassu Vc-Cap4 Lamassu (ddmABC), and white star the Lamassu Ec-Lipase system. The first ring indicates zinc-hook motif copy number which is either none (white), one (gray), or two (black); the outer ring distinguishes short (light blue) from long (dark blue) types. (*C*) Distribution of LmuB protein lengths and LmuA effector domains across Lamassu systems.

Mrr, Protease (accompanied by a hydrolase-like protein), and Sir2, while others are type-specific like the flavin-containing monooxygenase (FMO) for long Lamassu, and HNH, SMEK, and Lipase for short Lamassu. To determine whether the shared effector domains—Cap4, Mrr, Protease, and Sir2—arose from the same ancestral domain between long and short Lamassu systems, we analyzed their phylogenies within the broader context of their homologs (*SI Appendix*, Fig. S3). In all cases except for Sir2, effectors from long and short Lamassu formed distinct but closely related sister clades, suggesting either vertical inheritance or exchange between Lamassu systems. In contrast, Sir2 domains exhibited a markedly divergent evolutionary trajectory, with long- and short-associated Sir2 homologs branching apart in the phylogeny. This pattern suggests that while Cap4, Mrr, and Protease were likely inherited or transferred between Lamassu systems, Sir2 effectors originated from independent acquisition events. Finally, the composition of Lamassu varied significantly in the short Lamassu families. We observed loss of LmuC in a specific clade (Clade S, Fig. 1*B*), an architecture previously called type I Lamassu shown to be antiphage (12). We also identified a clade of short Lamassu which seemingly lost the effector domain of LmuA (Clade O, Fig. 1*B*). This clade contained more than 100 sequences across more than 40 species and was associated with nontruncated LmuC and LmuB genes.

Overall, short Lamassus are more diverse in terms of sequence, effectors and more abundant in bacterial genomes. We therefore focused on short Lamassus, selecting the experimentally validated Cap4 system from *V. cholerae* (Lamassu Vc-Cap4) for structural elucidation. This system encodes a LmuB with no zinc-hook, a LmuC, and a Cap4-like LmuA effector predicted to degrade double-stranded DNA (Fig. 2*A*).

Apo-LmuABC Complex Reveals an Unanticipated Core Topology that Sequesters Monomeric LmuA. Our initial attempts to purify LmuABC from the full operon (Fig. 2*A*) were unsuccessful, prompting us to express its components separately. We were able to coexpress and purify the LmuBC subcomplex; however separately purified LmuA exhibited severe aggregation upon size-exclusion chromatography (SEC) analysis (*SI Appendix*, Fig. S4*A*). For

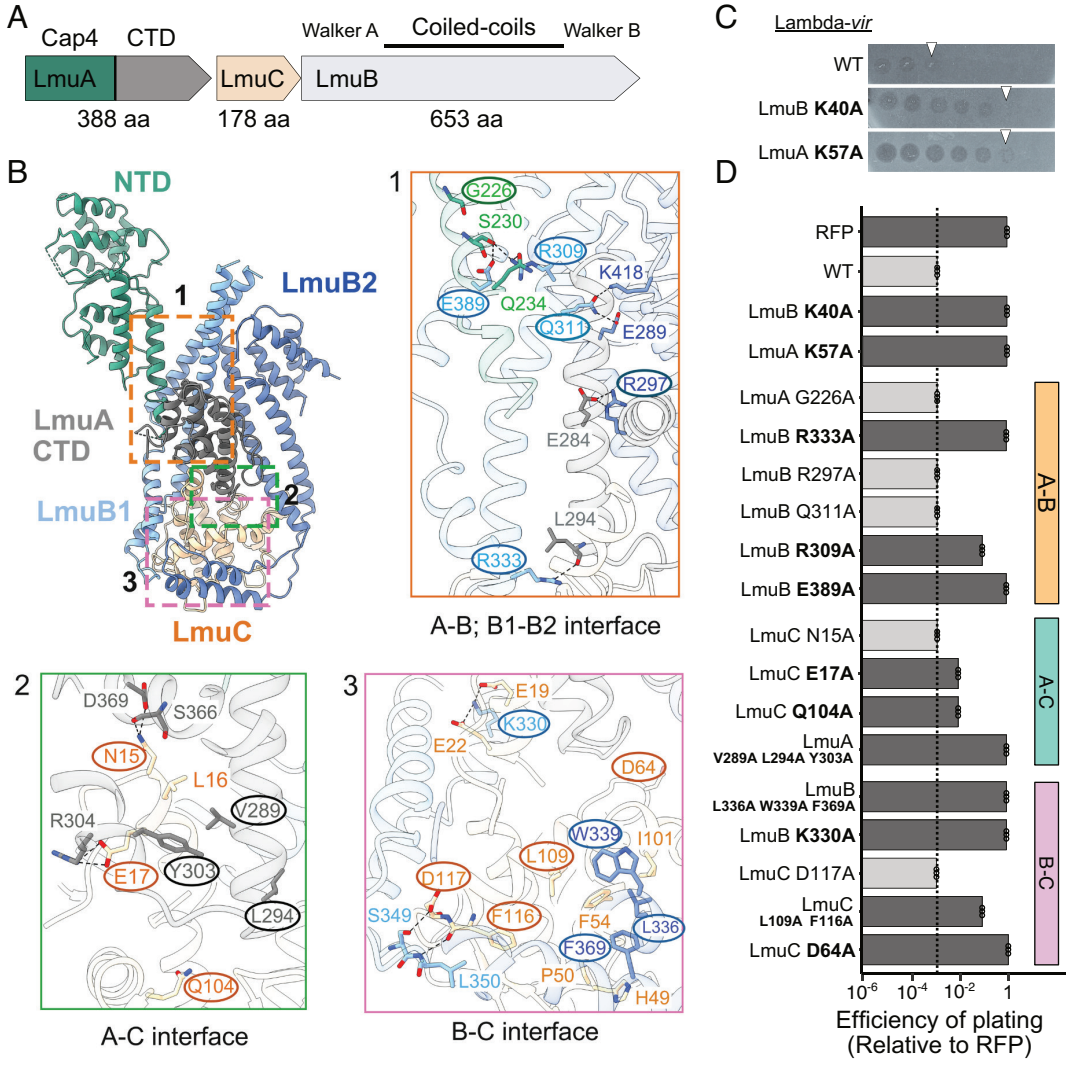

Fig. 2. Apo-state of $LmuA_1B_2C_1$. (*A*) Operon schematic of Lamassu Vc-Cap4 using color coding from Fig. 1*A*. (*B*) Cryo-EM structure of the apo-$LmuA_1B_2C_1$ shows core interactions involving LmuB (light/dark blue), LmuC (beige), and LmuA CTD (dark gray), with minimal interactions involving LmuA's N-terminal Cap4 domain (green). *Insets* (1 to 3) illustrate detailed hydrogen bonding and hydrophobic interactions, also shown in *SI Appendix*, Fig. S7. Encircled residues were selected for in vivo mutant design and efficiency of plating (EOP) assays. (*C*) Representative plaque assays of WT Lamassu Vc-Cap4 and catalytic mutants (LmuB K40A and LmuA K57A) against Lambda-vir phage. (*D*) Plaque assays probing structural interfaces. EOP indicates the relative plaque-forming efficiency of the tested phage on the bacterial strain compared to the negative control strain (RFP). An EOP of 1 indicates a completely nonfunctional defense system (no restriction of phage replication). Mutants impairing activity (dark gray bars) are highlighted in bold; WT activity level indicated by dotted line. Data points represent three biological replicates. Targeted interfaces (*A–B*, *A–C*, and *B–C*) are shown on the right.

coexpressed LmuBC, SEC yielded a well-defined peak containing both LmuB and LmuC, which migrated at their expected sizes on SDS-PAGE (*SI Appendix*, Fig. S4*B*). Notably, this copurification of tagged LmuB coeluting with LmuC, indicated a stable interaction between the two proteins. To reconstitute the full LmuABC complex, we mixed purified LmuA and LmuBC from their respective heparin columns followed by running it on a SEC column (*SI Appendix,* Fig. S4*C*) and native mass spectrometry (nMS) (*SI Appendix*, Fig. S5*A*) confirmed the formation of apo-$LmuA_1B_2C_1$. Addition of 20-bp DNA resulted in formation of an $LmuA_1B_2C_1$-dsDNA complex (*SI Appendix*, Fig. S5*B*).

Cryo-EM analysis of the LmuABC complex in the presence of a 20-bp dsDNA substrate revealed the presence of both the apo and dsDNA-bound states. To confirm the identity of the apo complex, we independently obtained the structure of the LmuABC in the absence of dsDNA, yielding an identical assembly. As both apo-LmuABC structures showed identical conformations and assemblies, with the apo- one obtained from the complex with dsDNA having higher resolution (3.21 Å), we decided to focus on this structure for further analysis (Fig. 2 *A* and *B*; work-flows in *SI Appendix*, Fig. S6; data statistics in *SI Appendix*, Table S1). This apo-LmuABC structure revealed a single full-length LmuA, the head-distal coiled coils of two LmuB subunits adopting different interconnected folds, and one full-length LmuC, with these interacting segments forming the core of the LmuABC. The kinked head-distal coiled coils of $LmuB_1$ (in light blue, single kink) and $LmuB_2$ (in dark blue, a pair of kinks) converge at the distal segment, interacting with the CTD of LmuA (in gray) and the globular fold of centrally positioned LmuC (in beige). Notably, LmuC directly associates with these coiled coils, forming an integral part of the complex (Fig. 2*B* and *SI Appendix*, Fig. S7). The Cap4 N-terminal domain of LmuA (in green) exhibits minimal interactions with the other subunits. A key feature of the apo-state is the monomeric nature of LmuA, with its NTD in an extended conformation relative to its CTD, with the latter stabilized through interactions with its partner subunits (Fig. 2*B* and *SI Appendix*, Fig. S7). Since LmuC was previously predicted to function as a kleisin (11), its placement within the LmuABC core was unexpected, given that in conventional SMC complexes, kleisins typically associate asymmetrically with the head-proximal coiled coils and head domains (15).

To investigate the functional relevance of the apo-LmuABC structure, we performed mutational analysis and assessed Lamassu's antiphage activity in vivo (Fig. 2 *C* and *D*). We expressed the natural operon of Lamassu Vc-Cap4 under an inducible promoter on a medium-copy plasmid in *Escherichia coli*. In this setup, at low induction levels of 0.0004% L-Arabinose, Lamassu is strongly antiviral against Lambda-vir, but not toxic (Fig. 2*C* and *SI Appendix*, Fig. S8). In contrast, the Lamassu Vc-Cap4 system is toxic at high inducer concentration, except for LmuA K57A mutant, suggesting autoactivation at high intracellular concentration (*SI Appendix*, Fig. S8). The $LmuB_1$-LmuC (B_1–C) interface was targeted through alanine substitutions at K330A in $LmuB_1$, and D64 and D117 in LmuC (Fig. 2*B*). Strikingly, all mutations except D117A abolished antiphage activity, highlighting the critical role of the B_1–C interface in system function (Fig. 2*D*). This was further supported by two mutants disrupting hydrophobic patches L336A+W339A+F369A in $LmuB_2$ and L109A+F116A in LmuC (Fig. 2*B*), both of which completely abolished activity (Fig. 2*D*). Similarly, targeting the LmuA-$LmuB_1$ (A–B_1) interface—R333A, R309A, and E389A in $LmuB_1$ (Fig. 2*B*)—abolished antiphage activity (Fig. 2*D*), reinforcing the functional significance of these interactions. For the LmuA-LmuC (A-C) interface, alanine mutations were introduced at N15, E17, and Q104 in LmuC, as well as V289+L294+Y303 in LmuA (Fig. 2*B*). Among these, E17A and Q104A displayed an intermediate phenotype, the hydrophobic patch mutant (V289A+L294A+Y303A in LmuA) completely disrupted function, whereas N15A had no significant effect (Fig. 2*D*). Together, these findings establish the structural and functional importance of the B–C, A–B, and A–C interfaces in LmuABC-mediated antiphage defense, while also revealing a previously uncharacterized role for LmuC in the complex architecture.

Lamassu Senses DNA Ends In Vitro. To understand how Lamassu effector activity is triggered, we next focused on the dsDNA-bound structure. We determined a 2.93 Å resolution structure, which revealed a complex composed of $LmuA_1B_2C_1$ bound to a single 20-bp dsDNA molecule (Fig. 3*A* work-flow in *SI Appendix*, Fig. S6, density maps in *SI Appendix*, Fig. S9 and data statistics in *SI Appendix* Table S1). The structure shows that Lamassu specifically binds dsDNA ends (Fig. 3*B*), while also making additional contacts along the duplex (Fig. 3*C*). Structural alignment of the apo- and dsDNA-bound complexes resulted in an RMSD of 1.0 Å (*SI Appendix*, Fig. S10), indicating that while the global architecture is largely maintained, DNA binding induced small localized rearrangements, particularly in $LmuB_1$, where additional density allowed tracing of its head domain and the head-proximal coiled coil that were too flexible to be resolved in the apo-state. By contrast, the head domain and head-proximal coiled coil of $LmuB_2$ remain flexible, suggesting asymmetry in DNA engagement.

The dsDNA duplex is captured within the LmuABC–DNA complex in a conformation that obstructs it from threading through the LmuB dimer, which is generally the case with SMC–dsDNA interactions. Further inspection of the LmuABC–DNA structure revealed that the phosphate at the 5′-DNA terminus inserts into the basic pocket formed by $LmuB_1$ residues K55, R152, K174, and H180 (Fig. 3*B*). Here, the side chains of R152 and H180 interact with the oxygen atom of the terminal phosphate, while the main chain of H180 also contributes to the interaction. These elements appear to define a DNA end-binding pocket rather than a DNA-threading mechanism. The DNA duplex segment interacts primarily with the head domain of $LmuB_1$ and the globular domain of LmuC. Key residues from the DNA contacting interface include K55, R152, K174, H180 (Fig. 3*B*), L92, T94, N96, N117, K122, and K147 (Fig. 3*C*) in $LmuB_1$, as well as S72, K73, K82, and N141 in LmuC (Fig. 3*C*). Unlike canonical SMC complexes, where the head-proximal coiled coils engage dsDNA, the LmuABC complex replaces these interactions with LmuC–DNA contacts. Certain residues within LmuA (38 to 78 and 117 to 153 at the N terminus, and 245 to 256 in the linker between the N and C termini), together with short segments of the dsDNA backbone, could not be modeled due to the poor density likely arising due to lack of interactions with other regions and a slight orientation-bias within the DNA-bound LmuABC density map.

To test the functional importance of distal contacts beyond the DNA end- and duplex-binding sites (*SI Appendix*, Fig. S11*A*), we generated $LmuB_1$ K183A and LmuC K87A and K149A mutants and tested their antiphage activity using a Lambda-vir plaque assay. $LmuB_1$ K183A and LmuC K149A mutants abolished defense, while LmuC K87A mutant caused a partial reduction in activity (*SI Appendix*, Fig. S11*B*). These results indicate that proper DNA recognition is not solely mediated by LmuB but also requires LmuC, while distal amino acid mutants also contribute to antiphage defense. Given its direct role in stabilizing dsDNA binding at the complex interface, LmuC likely functions as a specificity factor, ensuring that LmuABC engages dsDNA ends in a regulated manner. Accordingly, AF3 models of LmuC across the

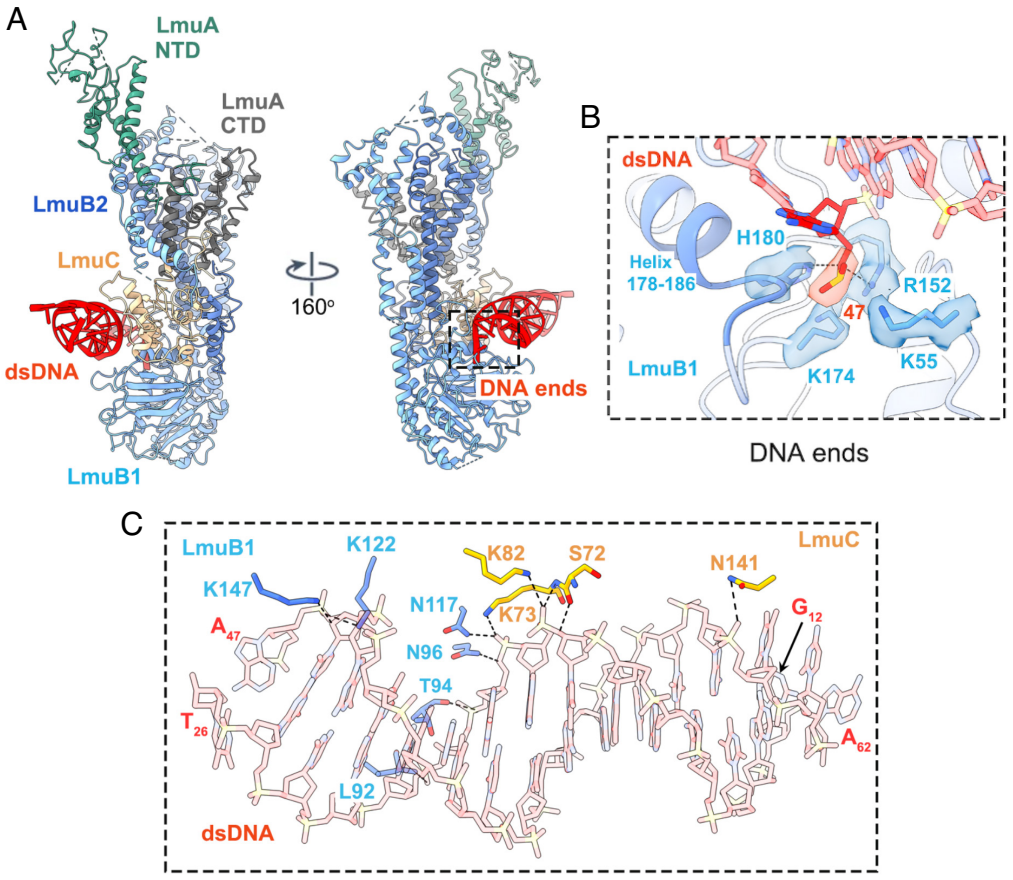

Fig. 3. Lamassu binds dsDNA ends in vitro. (*A*) Cryo-EM structure of $LmuA_1(K57A)B_2C_1$ bound to 20-bp dsDNA showcasing LmuB (light/dark blue), LmuC (beige), and LmuA CTD (dark gray), LmuA's N-terminal Cap4 domain (green) with dsDNA (red) in two views with the 160° rotated view highlighting the dsDNA end. (*B*) Key residues of $LmuB_1$ (light blue) and LmuC (beige) interacting with the terminal phosphate at the 5′-DNA terminus. The density for the 5′-phosphate is shown in red and for the $LmuB_1$ side chains (K55, R152, K174, and H180) are shown in light blue. (*C*) Key residues of $LmuB_1$ (blue) and LmuC (beige) interacting with dsDNA backbone.

phylogeny of short Lamassu repeatedly hit HTH (helix–turn–helix)-type transcription factors with DALI and FoldSeek. These findings suggest that LmuABC binds dsDNA in a structurally constrained mode, with localized rearrangements supporting selective recognition of DNA ends.

Lamassu Senses Phage Origins of Replication In Vivo. Having shown that Lamassu specifically binds dsDNA, particularly sensing the end, we hypothesized that it might recognize phage-specific DNA substrates during infection. To identify such substrates in vivo, we performed chromatin immunoprecipitation followed by sequencing (ChIP-seq) of Lamassu during infection with phage Lambda-vir (Fig. 4*A* and *SI Appendix*, Fig. S12*A*). We expressed the catalytically inactive LmuABC harboring a K57A LmuA mutation under the arabinose-inducible pBAD promoter, and then performed ChIP-seq of FLAG-tagged LmuB at 5, 10, and 20 min postinfection, with the goal of enriching specific substrates recognized by the Lamassu system during infection. The resulting data revealed a sharp and specific enrichment of LmuB around the four iterons—short direct DNA repeats—that comprise the Lambda origin of replication, located within the O gene (Fig. 4*A*). The O gene encodes the O protein, which initiates theta replication by binding its iterons and assembling a nucleoprotein complex that recruits the host replication machinery (22). LmuB enrichment around the iterons was detected specifically at 10 min postinfection, but not at 5 or 20 min, suggesting that Lamassu senses a transient replication intermediate substrate early during infection.

To test whether LmuB recognizes the DNA sequence of the O locus itself, independent of phage replication, we repeated ChIP-seq experiments on cells carrying a plasmid encoding only the enriched locus, but lacking Lambda phage. In this context, no enrichment was observed around the iterons (*SI Appendix*, Fig. S12*B*), indicating that Lamassu does not bind specifically to the sequence, but rather to a transient DNA structure during phage infection. This conclusion is further supported by the structure, which shows that LmuBC primarily contacts the DNA backbone rather than bases, consistent with the lack of sequence-specific interactions.

Lamassu Interfaces with DNA Are Conserved Across Diverse Lamassus. We aimed to determine whether the atypical organization and DNA binding observed in our structural work represents a recent innovation within the clade of short Lamassu Cap4 (Clade F, Fig. 1*B*), which have lost their Zn-hook. We tested this by experimentally validating a distant system from *E. coli*, referred to as Lamassu Ec-Lipase, which shares only 37.5% sequence similarity and belongs to short Lamassus encoding a lipase effector and a Zn-hook (Clade H, Fig. 1*B*). Similar to Lamassu Vc-Cap4, Lamassu Ec-Lipase was toxic at high inducer concentrations (*SI Appendix*, Fig. S9) but efficiently protected against Lambda-vir at lower, nontoxic levels (Fig. 4*B*); both the LmuB Walker A mutant (K42A) and the lipase catalytic mutant (H215A) impaired antiphage activity, but only the LmuA H215A mutant was nontoxic at high inducer levels, suggesting that LmuA autoactivates upon overexpression, independent of phage infection and ATP hydrolysis (Fig. 4*B* and *SI Appendix*, Fig. S9).

Assuming the atypical Lamassu Vc-Cap4 LmuABC molecular complex is conserved across short Lamassus, we extrapolated the interfaces identified in our structural work to the Lamassu Ec-Lipase system using AF3 multimer models with the stoichiometry determined earlier, comparing these to our Lamassu Vc-Cap4 structure and AF3 model (*SI Appendix*, Fig. S13). We identified

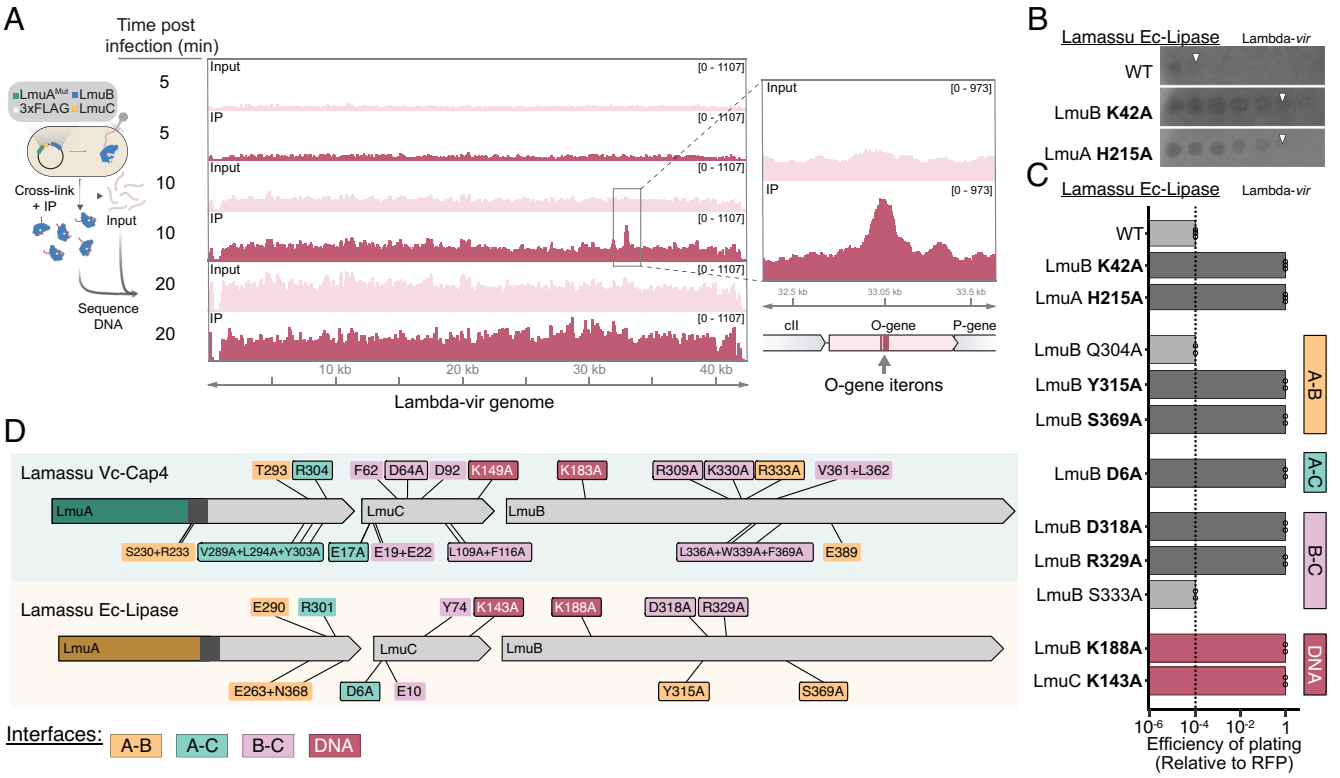

Fig. 4. Lamassu senses foreign origins of replication. Also, a comparative study of Lamassu Vc-Cap4 and Lamassu Ec-Lipase. (*A*) Schematic of ChIP-seq experimental workflow with FLAG-tagged LmuB, LmuC, and catalytically inactive LmuA (K57A Mut; *Left*), and representative ChIP-seq data (*Middle*) from experiments in which cells were infected with Lambda-vir. The genome-wide graphs show normalized read coverage mapped to the Lambda-vir genome at various time points after infection, with input and IP samples colored in pink and maroon, respectively. FLAG-tagged LmuB specifically enriched the Lambda-vir origin (O gene iterons) 10 min after infection, as shown in the magnified *Inset* at right; the genomic locus is annotated below the graph. (*B*) Validation of Lamassu Ec-Lipase antiphage activity by plaque assay. (*C*) Plaque assay of mutants of Lamassu Ec-Lipase. EOP indicates the relative plaque-forming efficiency of the tested phage on the bacterial strain compared to the negative control strain (RFP). An EOP of 1 indicates a completely nonfunctional defense system. Interfaces probed are indicated on the right; mutants with impaired activity (dark gray bars) are indicated in bold. WT activity level shown by dotted line. Data points represent three biological replicates. (*D*) Comparison of interface residues between; residues critical for defense verified by assays are indicated and shown in Figs. 2*D* and 3*G*.

four similar interchain interfaces (A–C, B–C, A–B, and DNA–Lamassu) and found nine analogous interactions: three pairs for A–B, three for B–C, one for A–C, and two involving dsDNA (Fig. 4 *B* and *C* and *SI Appendix*, Fig. S13). Alanine mutagenesis of Lamassu Ec-Lipase revealed that seven out of nine mutants abolished its activity, affecting at least one key interacting pair for each interface (Fig. 4 *C* and *D*). These results, in combination with our AF3 models, support that another member of the short Lamassu clade may form a similar molecular complex that interacts with dsDNA, with conserved interfaces essential for antiphage function. Hence, LmuABC assembles into an atypical SMC-like complex critical for activity across the diversity of short Lamassus.

Structural Basis of LmuA Activation. Since we observed that the LmuABC complex binds to dsDNA ends, and LmuB is a predicted ATPase, we wanted to understand how dsDNA and ATP play a role in Lamassu's activation.

The presence of a Walker A and Walker B domains in LmuB (*SI Appendix*, Fig. S2) suggest that ATP binding, and potentially hydrolysis, is required for Lamassu to function. Additionally, mutation of the catalytic Lysine (K40A) of the Walker A impairs defense against Lambda-vir (9, 20). We therefore tested the in vitro activity of LmuABC in the presence or absence of ATP. Our initial biochemical assays revealed nonspecific DNA degradation when high concentration of LmuABC (>20 nM) was added to pUC19 plasmid DNA without ATP (*SI Appendix*, Fig. S14*A*). To prevent cleavage from autoactivation, we decreased the LmuABC concentration until minimal DNA degradation could be observed. Increasing ATP concentrations correlated with enhanced DNA degradation, which was not the case for slowly hydrolyzable ATPγS (Fig. 5*A*). This suggests that ATP binding alone may be insufficient for activation and that ATP hydrolysis may contribute to LmuABC-mediated DNA processing.

Interestingly, DNA degradation was primarily observed at high ATP concentrations in the millimolar range, consistent with the intracellular abundance of ATP in bacteria. This suggests that LmuB, which contains a divergent Walker B motif (*SI Appendix*, Fig. S2), likely hydrolyzes ATP at a slow rate. Finally, comparison of degradation assays in the absence (Fig. 5*A*) and presence (Fig. 5*B*) of 20-bp dsDNA revealed that addition of this DNA substrate enhanced LmuABC-catalyzed pUC19 degradation. These results suggest that Lamassu's activity is likely to be regulated both by slow ATP hydrolysis and by interactions with dsDNA that participate in activating the nuclease activity of Cap4.

Our biochemical assays showed that Cap4 activity from LmuA required the presence of LmuABC and was not detectable with LmuA alone in solution (*SI Appendix*, Fig. S14*B*). We hypothesized that dsDNA end binding by LmuABC could activate LmuA, and sought to capture this active state by cryo-EM. Since the complex is biochemically active at high concentrations (>20 nM) of LmuABC, we reasoned that adding LmuBC to an excess of LmuA might stabilize active conformations suitable for cryo-EM. Analysis of this sample revealed a tetrameric architecture for LmuA effector—a dimer of dimers—present in both twofold asymmetric and

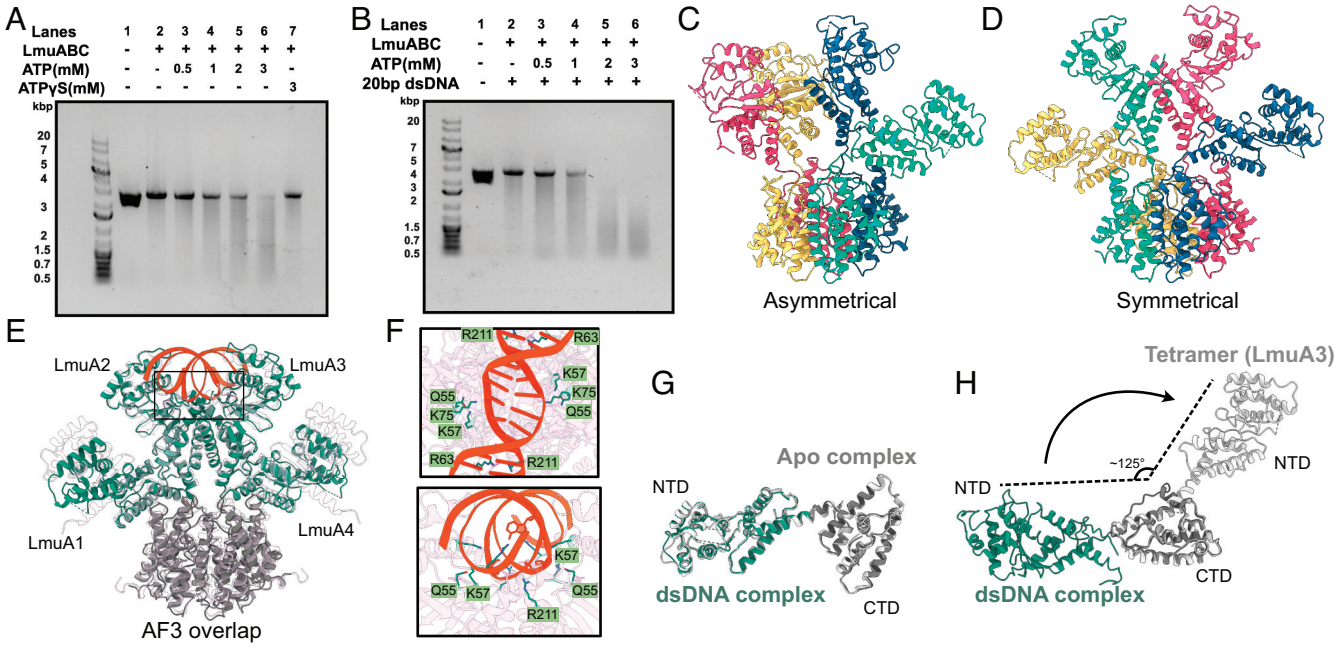

Fig. 5. Structural basis of LmuA activation. (*A*) ATP-dependent DNA degradation assay of LmuABC (20 nM) on pUC19 (40 nM); increasing ATP concentrations enhance DNA degradation activity of LmuABC. (*B*) Addition of 20-bp dsDNA (0.2 μM) enhances ATP-dependent DNA degradation activity of LmuABC. (*C* and *D*) Cryo-EM structures of asymmetric (*C*) and symmetric (*D*) LmuA tetramers, illustrating distinct bent and extended conformations stabilized by polar interfaces (shown in *SI Appendix*, Fig. S16 *B–D*). (*E*) AlphaFold3 (AF3) model (transparent pink) aligned with symmetric tetramer; dsDNA aligns precisely with predicted DNA cleavage site. (*F*) Close-up of residues involved in dsDNA binding and cleavage predicted by AF3, highlighting catalytic residues including K57. (*G*) LmuA shows no significant conformational changes between apo (in silver) and DNA-bound (in color) complexes (RMSD = 0.795 Å). (*H*) Dramatic conformational rearrangements occur at the linker region of LmuA between DNA-bound LmuABC (in color) and activated LmuA tetramer (in silver) structures following superpositioning of their CTDs.

symmetric conformations at similar abundance (Fig. 5 *C* and *D*, respectively; workflow in *SI Appendix*, Fig. S15 and Table S1). In both structures, tetramerization was mediated by interactions between the LmuA CTDs, which superimpose with an RMSD of 2.02 Å (*SI Appendix*, Fig. S16*A*). However, the Cap4 (NTDs) adopted different relative orientations: in the symmetric tetramer (Fig. 5*D*), a single predicted DNA degradation pocket was formed between the NTDs of LmuA2 and LmuA3 (Fig. 5*E*). Although the overall architecture could, in theory, support a second catalytic site between LmuA1 and LmuA2 (or alternately between LmuA3 and LmuA4), no such pocket was observed. This asymmetry in catalytic site formation may reflect a built-in regulatory mechanism that limits nuclease activity to a single cleavage site per tetramer, thereby raising the activation threshold and minimizing the risk of unintended autoimmunity. By contrast, the asymmetric tetramer (Fig. 5*C*) showed no predicted degradation pocket, suggesting that it may represent a partially activated or inactive state of LmuA.

AF3 modeling using four LmuA monomers and a dsDNA molecule recapitulated the symmetrical tetramer with the dsDNA bound between LmuA2 and LmuA3 (Fig. 5*E*). Hallmark residues comprising K57, which is essential for DNA degradation in vitro and phage protection in vivo, are aligned in close vicinity of the DNA, forming a DNA cleavage pocket, predicted to result in 1-nt 3′-overhang containing DNA duplexes as the cleavage product (Fig. 5*F*). We thus hypothesized that the LmuA tetramer is the end result of Lamassu activation. Notably, we could not observe significant conformational changes between LmuA in the dsDNA-bound vs. apo-LmuABC complex (Fig. 5*G*), but a significant rotation of the NTD Cap4 domain of LmuA in the active tetramer (Fig. 5*H*). Such a switch in conformation could be facilitated most likely by a complete disengagement of LmuA from the complex upon activation. This disengagement would facilitate the oligomerization into a tetramer, liberating essential residues from the CTD for the tetramerization that are occluded in the apo- or dsDNA-bound complexes and thus allowing the formation of the DNA degradation pocket in the NTD.

Lamassus Evolved from DNA Repair Complex SbcCD. Last, we investigated the likely evolutionary origins of Lamassu. SMC proteins are part of the P-loop NTPase superfamily and have diversified into two major families that can be distinguished by the presence of either a hinge (Hinge SMC) or zinc-hook (CxxC SMC) dimerization domain (11). Hinge-encoding SMCs, such as the SMC1–6 complexes in humans and MukB in *E. coli*, are primarily involved in the condensation and organization of chromosomal DNA (5, 23). CxxC SMCs such as RAD50 in humans and SbcC in *E. coli*, play a critical role in DNA repair, notably of double-strand breaks for RAD50-Mre11-Nsb1 (24) and of palindrome-like convergent replication fork intermediates for SbcCD (25). Thus, model CxxC SMCs are primarily implicated in processing noncanonical DNA ends to facilitate subsequent repair steps (26). Given the known biochemical activity of Lamassu, and the resemblance of LmuB to SbcC, we thus hypothesized that Lamassu might have emerged from SbcCD.

To test for this, we performed a sensitive search for all homologs of LmuB, resulting in n = 367,581 proteins grouped in 1,320 families (30% coverage and identity), including proteins from other defense systems such as Rloc (27) and DndD, the SMC-like ATPase of the Dnd system (28), and numerous SbcC homologs from the SbcD complex but no MukB homolog. This further supports that LmuB is a closer homolog of SMC-like proteins like SbcC than hinge-SMCs like MukB. The resulting phylogeny (*SI Appendix*, Fig. S17) also recovers the two clades of long and short LmuB. We thus subsampled diverse members of SbcC, long and short Lamassu and the distant SMC protein MukB, to root our phylogeny. This revealed that short Lamassu emerged from the clade of long Lamassu, and places long LmuB closer to SbcC (Fig. 6 *A* and *B*).

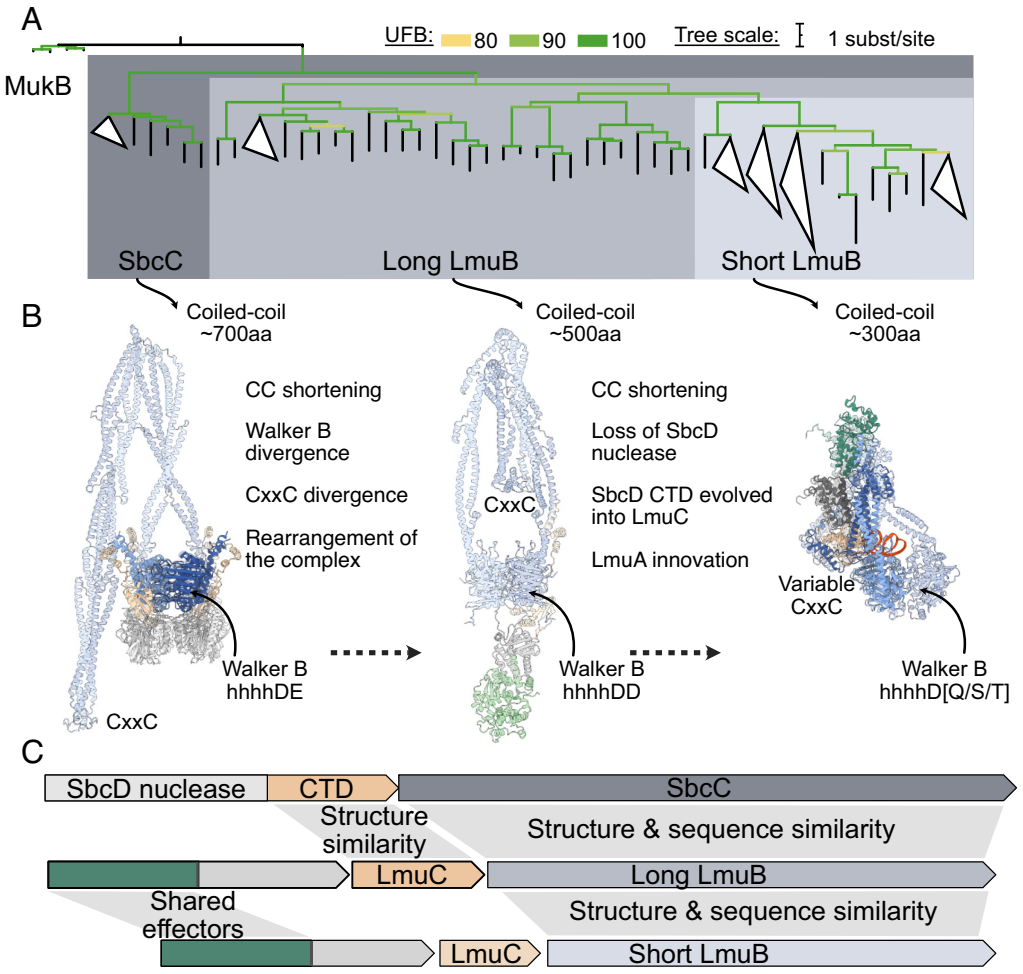

Fig. 6. Lamassu originated in SbcCD DNA repair complex. (*A*) LmuB homologs phylogenetic tree. Structure-guided alignment (MukB, SbcC, long LmuB, and short LmuB) of AF3 models used as seed alignment for Mafft. The tree was generated with IQ-tree2 using ModelFinder (Q.pfam+F + I + R7) and Ultra-fast Bootstraps (UFB). Nodes with UFB <80 are collapsed. (*B*) Evolutionary scenario of the emergence of Lamassu from the DNA repair system SbcCD. Experimentally solved structures shown in solid colors, AlphaFold3-modeled structures shown transparently. SbcCD partial complex is taken from PDB:6S7V, and the short Lamassu structures are taken from the cryo-EM structures of DNA-bound Lamassu Vc-Cap4. (*C*) Comparison of typical operons from SbcCD and Lamassu systems highlighting regions of similarity (gray shading).

We then studied the evolutionary links of LmuA and LmuC. For LmuC, a similar approach than for LmuB did not recover hits outside the Lamassu system, and overall, the size and sequence diversity of LmuC precluded us from relying on sequence alignment methods. Using structural modeling and structure-similarity search with DALI, we recovered weak but significant (Z = 2.6) hits from long LmuC to the C-terminal region of SbcD of *E. coli*. This partial match was further confirmed by using FoldSeek with different homologs of long LmuC (best hit probability $P = 1$, *SI Appendix*, Fig. S18). This striking observation suggests that long LmuC derived from SbcD. For LmuA short and long CTDs, we could not find similarity to other known domains. By contrast, many effectors like Cap4, Mrr, Protease, and Sir2 were also found in other defense systems, especially in the Avs family. Contrarily to LmuC, we could not find evidence of homology between LmuA and SbcCD.

Since long LmuB are close homologs of SbcC, and long LmuC are structurally similar to the SbcC-binding domain of SbcD (SbcD CTD), we hypothesized that Lamassu evolved from an ancient SbcCD system. The operon architecture of Lamassu is compatible with such a scenario, since the SbcD CTD lies just before SbcC, like LmuC before LmuB (Fig. 6*C*). We thus studied the conservation of the structure and molecular complex using structural model predictions. We generated 20 complexes in silico, including the SbcCD system of *E. coli*, long (n = 5) and short (n = 14) Lamassus (*SI Appendix*, Fig. S19). Complexes of long and short Lamassu's were more different between long and short than within, revealing only two major conserved architectures for both. Short LmuC were positioned similarly as in the Lamassu Vc-Cap4 complex, while long LmuC had an additional short domain separated by an unstructured linker (*SI Appendix*, Fig. S20). Similarly, short LmuA CTDs are predicted to form a conserved bundle of 5 to 6 helices, while long LmuA CTDs are predicted to form a more complex fold with helices, a conserved loop with negatively charged residues (predicted to bind LmuB), and a beta-sheet (*SI Appendix*, Fig. S20). In predicted complexes of long Lamassu, LmuA and LmuC were predicted to interact around the head domains of LmuB, reminiscent of the interaction between SbcC and SbcD (29). Strikingly, long LmuC were positioned similarly to the CTD of SbcD.

We propose an evolutionary model (Fig. 6*C* and *SI Appendix*, Fig. S20) for short Lamassu systems originating from an ancestral SbcCD complex with long coiled-coil domains, standard Walker motifs, and a single CxxC motif. Separation of SbcD's nuclease and CTD gave rise to LmuC and facilitated modular effector integration. Long Lamassu evolved through innovation of the LmuA CTD, retaining key motifs but shortening coiled-coil regions by ~200 amino acids. Short Lamassu further compacted from long Lamassu by additional coiled-coil reduction and significant divergence or innovation of the LmuA CTD and LmuC. These short variants display diversified Walker B motifs, greater effector diversity, and in some cases, loss of entire components (LmuC or effectors). This staged reduction underscores the gradual remodeling of a DNA-repair system into a defense complex.

Discussion

Our study reveals how Lamassu evolved from a DNA repair ancestor through major structural innovations that enabled its transition into an antiphage defense system. Based on our results, we propose that

as the LmuABC complex scans along dsDNA, it recognizes and binds free DNA ends or a related DNA substrate, with this binding event mediated by LmuB and LmuC. We postulate that ATP hydrolysis could promote the disruption of the LmuABC–DNA complex, facilitating LmuA release and tetramerization into an activated state. This active tetramer mediates widespread, nonspecific cleavage of cellular dsDNA. The need for disengagement of LmuA from the LmuABC–dsDNA complex for activation through changes in the oligomerization status could be a conserved mechanism across different Lamassu systems as several LmuA effectors are predicted to work in an oligomeric state. A schematic summarizing the sequestration, sensing, release, transition, and multimerization steps highlights a stepwise mechanism governing Lamassu-mediated antiphage defense (Fig. 7*A*). The factors that contribute to the mechanism triggering release of sequestered LmuA is not fully understood at this time and remains to be elucidated in the future.

Lamassu systems fall into two families—long and short forms—both built around a conserved SMC-like LmuB scaffold. While LmuA-NTDs are often homologous, the LmuA-CTDs and LmuC proteins differ markedly, reflecting dynamic modular evolution likely enabling adaptation to diverse phage threats. By solving the apo and dsDNA-bound structures of the short-form Lamassu Vc-Cap4, we reveal that it forms an atypical SMC-like complex, distinct from both canonical SMC proteins and RAD50/SbcC-like complexes. Unlike the expected rod-shaped dimer with partner proteins interacting around the head domains, LmuB adopts a kinked coiled-coil conformation, with its apex contacting LmuC and the CTD of LmuA. Both LmuB heads are not visible in the apo-LmuABC structure suggesting flexibility, with one of them ($LmuB_1$) becoming visible on binding to dsDNA. We suggest recognition of dsDNA ends leads to disengagement of the $LmuB_2$ head, while only the LmuC-bound head interacts with dsDNA. This disengaged state correlates with activation, likely driven by ATP hydrolysis, as ATPγS does not stimulate activation and the LmuB K40A mutant is inactive. Functional assays confirm that interactions between LmuB, LmuA, and LmuC subunits are essential for activity in vivo.

The Lamassu LmuABC complex recognizes dsDNA ends in vitro, a feature explained by structural elements in LmuB that block the duplex from threading through the complex (Fig. 7*A*). These results are in line with a recent preprint on a close homolog of Lamassu Vc-Cap4 (30). LmuC also interacts with the dsDNA, suggesting that it may function as a specificity factor that stabilizes the interaction between LmuB and DNA ends or a structurally related substrate. In uninfected cells, we suggest that LmuB could continuously scan genomic DNA in search of its cognate substrate. Upon infection with Lambda-*vir*, a strong and specific enrichment of LmuB at the phage's origin of replication is observed. This mode of replication continues until approximately 16 min postinfection (31), which aligns with our observation that LmuB is enriched at

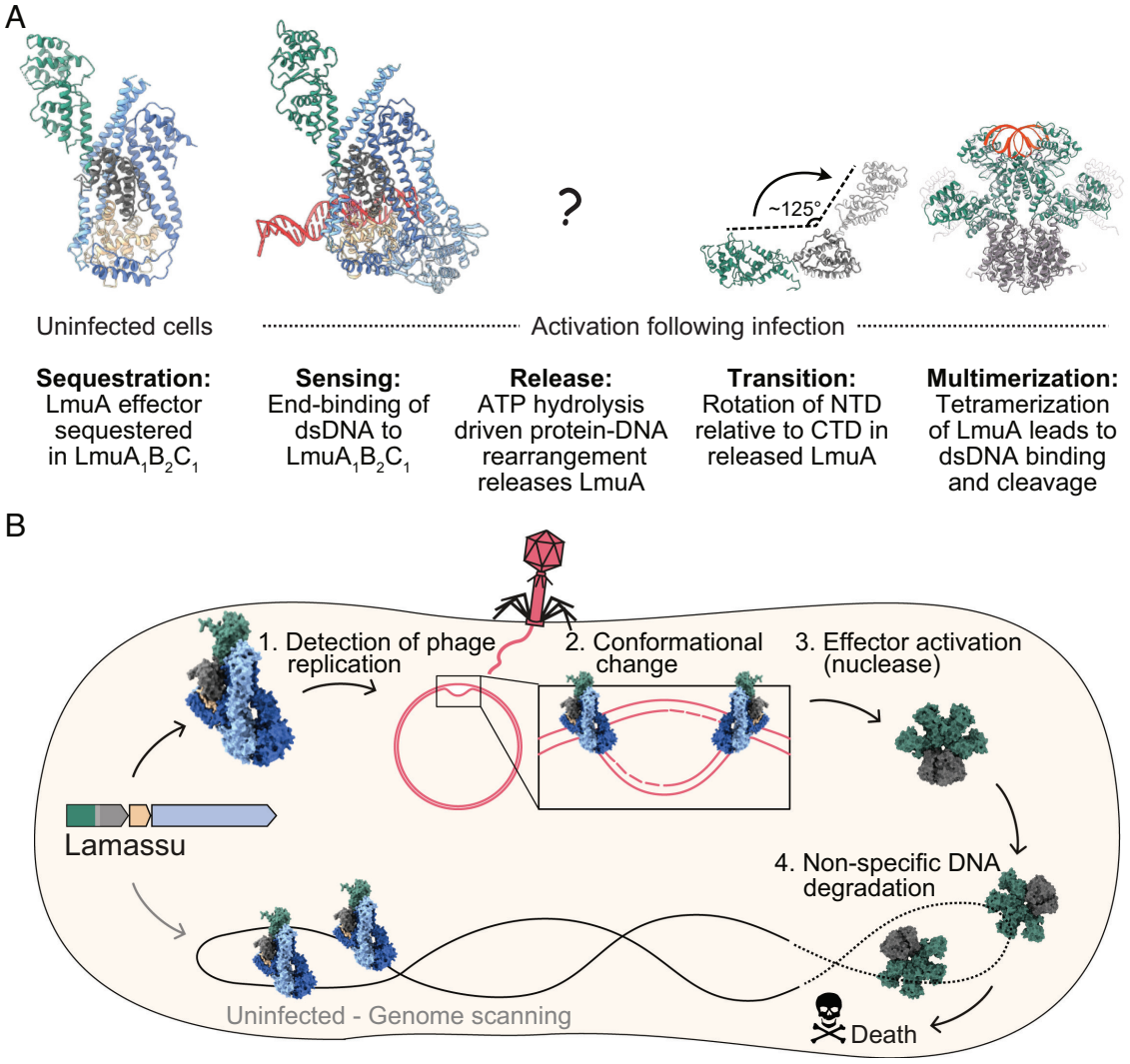

Fig. 7. Model for Lamassu-mediated antiviral defense. (*A*) Structural model of Lamassu activation. (*B*) Cellular model of Lamassu defense.

the origin at 10 min, but not at 20 min postinfection—when replication has transitioned to the rolling-circle mode (31).

We propose that Lamassu senses DNA structures specific to phage replication (Fig. 7*B*), which could derive from iterons. Supporting this, Lamassu is triggered by structured DNA motifs like palindromes and can be blocked by host or phage mutations affecting replication initiation (8). We observe that LmuB specifically binds iterons in the O gene in a replication context. Those iterons are predicted to form hairpin or cruciform structures (32) and are bound to the O protein. Formation of the O-iterons complex leads to DNA bending that facilitates the unwinding of an adjacent AT-rich sequence (33–35), allowing bidirectional theta replication to proceed (22, 33, 34). These findings suggest that Lamassu detects unstable DNA intermediates—such as open origins or stalled forks—associated with phage replication, allowing it to distinguish foreign from host DNA and act as a sensor of "reckless" replication events. In this way, Lamassu may have evolved not simply to recognize replication per se, but to discriminate the timing, coordination, and molecular signatures that distinguish phage replication from its host counterpart.

Beyond DNA recognition, our study reveals a major conformational transition in LmuA, linked to activation. In $LmuA_1B_2C_1$, LmuA is monomeric and extended, stabilized by interactions with LmuB and LmuC. However, when excess LmuA is added to LmuBC, forming an activated complex in biochemical assays, LmuA assembles into a tetrameric state observed by cryo-EM, in line with recent observations in a related system (30). This tetramer stabilizes two Cap4 NTDs, a distant relative of the PD-(D/E)xK nuclease family, also associated to CBASS and Avs systems (36, 37), into a single active cleavage pocket, while the remaining two Cap4 NTDs do not engage in cleavage. ATP hydrolysis most likely provides the energy to trigger the conformational changes necessary for disengagement of LmuA from LmuABC and thus LmuA activation by oligomerization via its CTD. Our observation of an asymmetrical tetramer that does not form an active degradation pocket suggests the existence of a potentially autoinhibited or transition state. Such an arrangement could provide a safeguard against accidental activation by requiring a specific structural transition, LmuA activity would remain suppressed when present as free subunits, ensuring that cleavage is triggered only once LmuBC has sensed the phage DNA and initiated the activation cascade. In this view, the "non-catalytic" LmuA subunits may function as regulatory elements, stabilizing the tetramer while preventing premature nuclease activity, and thereby linking effector activation tightly to upstream detection. Consistently, the requirement for millimolar ATP concentrations suggests that ATP hydrolysis by LmuB itself may act as an additional safeguard against autoactivation.

The evolutionary trajectory of the Lamassu system suggests it originated from the DNA repair complex SbcCD, with long Lamassu being an intermediate between this ancestral repair machinery and the compacted short Lamassu variants. This transition likely involved repurposing SbcCD's "detect-and-repair" function into a "detect-and-kill" strategy, shifting from site-specific degradation of DNA to an abortive infection system relying on promiscuous effectors. Unlike SbcCD, which precisely positions SbcD for cleavage, Lamassu activates diverse effectors, including nucleases like LmuA Cap4 and enzymes such as Sir2, proteases, and FMO, broadening its defensive potential. Beyond structural and sequence similarities between SbcCD and Lamassu is their shared substrate recognition: SbcCD detects and processes palindromic DNA, excising these sequences and resolving replication fork structures (25), while the human orthologs Rad50-Mre11 detects dsDNA ends (26). The ability to recognize structural DNA motifs rather than specific sequences suggests that Lamassu evolved by co-opting an ancestral DNA surveillance function, transforming it from a genome maintenance tool into a broad-spectrum immune defense system.

Many proteins involved in antiphage defense systems contain domains that are also found in DNA repair proteins, highlighting a deep evolutionary link between genome maintenance and immunity (38). For instance, the RecB-like nuclease of RecBCD is co-opted in Cas4 for CRISPR spacer acquisition (39), and RecBCD itself degrades foreign DNA (40), countered by phage-encoded inhibitors like Lambda Gam. Other repair domains—Toprim, Ku, and helicase motifs from proteins like RecQ—also appear in defense systems (20, 41, 42), highlighting a recurrent evolutionary link between genome maintenance and immunity. This evolutionary theme extends beyond bacteria, as eukaryotic cells also exploit DNA repair complexes for antiviral defense. For example, RAD50, part of the MRN complex, partners with CARD9 to sense cytoplasmic viral DNA (43) while the Smc5/6 complex restricts viruses but is targeted for degradation by Hepatitis B virus (44, 45). These examples illustrate that DNA repair enzymes have repeatedly served as a reservoir for the evolution of immune defenses, reflecting a broader evolutionary strategy where existing molecular machineries have been repurposed for defense against pathogens.

Materials and Methods

Additional materials available in *SI Appendix*.

Genome Database. We used the RefSeq complete genomes database from the National Center for Biotechnology Information (NCBI), downloaded in July 2022, composed of 22,920 prokaryotic genomes. Genomes were formatted in the gembase format with PaNaCota v1.3 (46) and annotated with DefenseFinder v1.09 (47–49) and DefenseFinder models v1.1.0 (47–49). The list of genomes is available as *SI Appendix*, Table S2.

Lamassu Detection. To comprehensively identify Lamassu systems, we retrieved candidate systems using an iterative search strategy that included experimentally validated systems from Doron et al. (12), Millman et al. (20), the ddmABC system (8–10), and predicted homologs from Krishnan et al. (11). Each system was assigned to a clade based on LmuB sequence clustering [MMseqs2 (50), 40% identity and coverage] followed by clade selection on the phylogeny (*SI Appendix*, Fig. S15), LmuA effector type, and operon architecture. Zinc-hook motifs and Walker boxes were annotated using regular expressions and HHPred, and manually validated. The list of detected systems is available as *SI Appendix*, Table S3. Additional phylogenetic methods are described in *SI Appendix*, *Phylogenetic Analysis of Lamassu Components*.

Strains and Plasmids. Strains and plasmids used in this study are available as *SI Appendix*, Table S4. Mutagenesis protocols are provided in *SI Appendix*, *Mutagenesis* and mutagenic primers listed as *SI Appendix*, Table S5.

Protein Expression and Purification. Codon-optimized genes encoding LmuA, LmuB, and LmuC from *V. cholerae* El Tor were synthesized (IDT) and cloned into pRSF-Duet vectors. LmuB (UniProt: A0A060KSR0) was His-SUMO tagged and coexpressed with untagged LmuC (UniProt: B7X6T6) in *E. coli* BL21(DE3). Expression was induced at OD_{600} = 0.4 to 0.6 with 0.4 mM IPTG at 16 °C overnight. Proteins were purified by Ni-NTA affinity, Ulp1 digestion, heparin chromatography, and SEC on a Superdex 200 Increase 10/300 GL column. Final buffer for SEC was 25 mM HEPES pH 7.5, 250 mM NaCl, and 5 mM DTT. LmuA (UniProt: A0A060KT36) was expressed and purified similarly, with His-SUMO tagging. A K57A mutant was generated by PCR mutagenesis. Both wild-type and mutant LmuA exhibited aggregation during SEC and were therefore used after the heparin step for reconstitution of the LmuABC complex. Detailed protocols including buffers and chromatography conditions are provided in *SI Appendix*, *Protein Expression and Purification*.

Phage Plaque Assays. Phage plaque assays were conducted using LB-Lennox agar supplemented with 0.7% agar, 100 μM carbenicillin (Carb), 5 mM $MgSO_4$, and 5 mM $CaCl_2$. Varying concentrations of L-arabinose were included to induce the expression of the systems. To prepare the bacterial lawn, 300 μL of overnight bacterial cultures grown in LB-Miller medium supplemented with 100 μM carbenicillin and 0.02% glucose were mixed with 20 mL of LB-Lennox agar. The mixture was poured onto plates and allowed to solidify for 30 min at room temperature.

Phage spotting was performed by applying 5 μL of serially diluted phage suspensions in LB-Lennox medium onto the solidified bacterial lawn. Plates were incubated at 37 °C for 16 to 18 h before imaging the resulting plaques. The last phage dilution showing lysis on the plate was reported and compared with the associated negative control (pBAD-RFP). EOP is the ratio of last dilution showing lysis of the experiment over the control. An EOP of 1 indicates that there is no protection.

Structural Modeling. We used AlphaFold3 (51) with a fixed stoichiometry deduced from the nMS data and cryo-EM structures for comparative structural studies. 1 LmuA, 1 LmuC, 2 LmuB with 2 Mg^{2+}, 1 Zn^{2+}, 2 ADP, and 72-bp dsDNA. The best ranking models were evaluated and visualized with ChimeraX v1.9 (52). We used the match command to compute the RMSD between models. We used the matchmaker function with the secondary structure weighting set to 1 to superimpose distant domains such as LmuA CTD and whole-proteins such as LmuB and LmuC. We used the DALI webserver (53) to compute pairwise similarity as a Z-score between all models.

Cryo-EM Sample Preparation, Data Collection, and Data Processing.
LmuA tetramer. A sample of LmuA and LmuBC (5:1 molar ratio; 0.3 mg/mL) was prepared from the heparin (LmuA) and SEC (LmuBC) purification steps, with no cross-linking. 4 μL of sample was applied to graphene oxide grids (R1.2/1.3, 400 mesh), blotted for 1 s (blot force 0) at 4 °C and 100% humidity, and vitrified using a Vitrobot Mark IV (FEI). Data collection was done at MSKCC on a Titan Krios G2 (FEI, 300 kV) with a K3 direct detector, using SerialEM. 7,002 movies were acquired in superresolution mode (total dose: 53 $e^-/Å^2$, defocus: –0.8 to –2.2 μm, pixel size: 1.064 Å). Processing in cryoSPARC yielded two LmuA tetramer conformations at 3.13 Å (193,953 particles, symmetric) and 3.4 Å (197,714 particles, asymmetric).
Apo-LmuABC. LmuA and LmuBC were mixed (5:1), incubated overnight at 4 °C, crosslinked with 0.025% glutaraldehyde (30 min on ice), and quenched with 0.1 M Tris pH 8.0. 4 μL of the sample (in 1 mM PPi) was applied to glow-discharged UltrAuFoil 300 mesh R1.2/1.3 grids, blot time 2 s, wait time 10 s. Data were collected on a Titan Krios G4 (300 kV, Falcon 4i detector) with EPU software (18,780 movies, pixel size: 0.725 Å, dose: 29.28 $e^-/Å^2$, defocus: –0.8 to –2.2 μm). Final map resolution for apo-$LmuA_1B_2C_1$ was 3.28 Å (127,014 particles).
LmuA(K57A)BC–DNA complex. The LmuA K57A mutant was incubated with LmuBC (1:2 molar ratio), then mixed with a 20-bp dsDNA duplex (sequence: GTGATAGTTAGAAACGTAAT and complement) at a 1:4 protein:DNA ratio, incubated 1 h on ice, crosslinked with 0.025% glutaraldehyde, and applied to UltrAuFoil R1.2/1.3 300 mesh grids. Same blotting conditions were used. Data were collected on the Titan Krios G4 (Falcon 4i), with 11,653 movies. CryoSPARC processing yielded: 3.21 Å for apo $LmuA(K57A)_1B_2C_1$ (135,737 particles); 2.93 Å for monomeric DNA-bound $LmuA(K57A)_1B_2C_1$ (190,887 particles); 11.2 Å for dimeric DNA-bound $LmuA(K57A)_1B_2C_1$ (2,690 particles).

Model building was performed using AlphaFold3 predictions, Coot, and Phenix real-space refinement. For the DNA-bound structure, the model building was based on the map generated using EM-ready, then validated against the raw map, further adjusted in Coot, and refined through additional cycles of real-space refinement in PHENIX using the raw map. For all maps, the resolution estimates were based on a Fourier shell correlation (FSC) cutoff of 0.143. Figures were prepared using UCSF Chimera (54) and USCF ChimeraX (52) with the final image layout created in Adobe Illustrator. Full procedures, including map filtering and figure generation, are described in *SI Appendix, Cryo-EM Sample Preparation and Processing*.

Chromatin immunoprecipitation sequencing (ChIP-Seq). Chromatin immunoprecipitation was performed using LmuB tagged with 3× Flag and catalytically inactive LmuA(K57A) to prevent DNA degradation. The full Lamassu operon was expressed under arabinose induction (0.02% L-arabinose), with or without an additional plasmid encoding a DNA target (pTarget). Where indicated, lambda phage infection was performed at MOI 5, and samples were collected at 5, 10, and 20 min postinfection (or at OD_{600} = 0.6 for plasmid-only experiments). Sequencing was performed on Illumina MiniSeq and NextSeq platforms, generating >2 million reads per sample.

Additional details–including buffer compositions, culture preparation, wash steps, and sonication settings–are described in *SI Appendix, ChIP-Seq Protocol*.

In Vitro DNA Degradation Experiments. DNA cleavage was assayed using 40 nM pUC19 plasmid DNA (Thermo Scientific) and LmuABC complex in reaction buffer (50 mM HEPES pH 7.5, 50 mM KCl, 1 mM DTT, 10 mM $MgCl_2$). To assess ATP dependence (Fig. 4*A*), reactions were supplemented with ATP (0.5 to 3 mM) or 3 mM ATPγS. Reactions were incubated at 37 °C for 1 h, stopped with 1 mg/mL proteinase K (Thermo Scientific) at 55 °C for 10 min, and loaded on 1% SYBR Safe-stained agarose gels (Thermo Fisher), with 1× purple loading dye (NEB). For short DNA duplex substrates (Fig. 4*B*), 0.2 μM 20-bp dsDNA was preincubated with LmuABC under identical conditions. Cleavage efficiency was also tested with increasing LmuABC concentrations (20, 40, 80, and 160 nM) (*SI Appendix*, Fig. S14*A*), and with individual protein components and mutants to assess functional requirements (*SI Appendix*, Fig. S14*B*). These included reactions with LmuBC alone (100 nM), LmuA (100 nM, heparin fraction), combinations of wild-type or K57A mutant LmuA with LmuBC, with or without 1 mM ATP. Full details are provided in *SI Appendix, DNA Degradation Assay Protocols*.

Native Mass Spectrometry (nMS) Analysis of Complexes. LmuABC and LmuABC–DNA complexes were buffer-exchanged into 300 mM ammonium acetate (pH 7.5, 0.01% Tween-20) using 7 kDa MWCO Zeba microspin columns (Thermo Scientific). Samples (2 to 3 μL) were electrosprayed from in-house gold-coated quartz capillaries into an Exactive Plus EMR mass spectrometer (Thermo Fisher Scientific) using static nano-ESI. Typical parameters included: spray voltage 1.20 to 1.22 kV, capillary temperature 150 °C, in-source dissociation 10 V, S-lens RF level 200, resolving power 8,750 or 17,500 at m/z 200, AGC target 1e6, max injection time 200 ms, and ≥100 scans. HCD energy was set to 150–200 V; ultrahigh vacuum pressure ranged from 3.4–4.8 × 10^{-10} mbar. Mass calibration was done using cesium iodide. Raw spectra were visualized with Thermo Xcalibur Qual Browser (v4.2.47), and deconvoluted using UniDec v4.2.0 with Gaussian peak shape and beta = 10 to 20. Expected monomeric masses were LmuA: 44,615.3 Da, LmuB: 74,619.0 Da, LmuC: 20,286.8 Da, and 20-mer dsDNA: 12,230.1 Da. Observed deviations were between 0.004 and 0.03%.

Data, Materials, and Software Availability. All standardized datasets are available as follows: cryo-EM maps have been deposited in the Electron Microscopy Data Bank (EMDB) under the accession codes EMDB: EMD-49911 (LmuA tetramer_conformation 1_symmetric) (55), EMD-49915 (LmuA_tetramer conformation 2_asymmetric) (56), EMD-49922 (LmuABC_apo) (57), EMD-49934 (LmuABC–DNA) (58). The corresponding atomic coordinates of the cryo-EM structures have been deposited in the Protein Data Bank (PDB) under the accession codes PDB: 9NXX (LmuA tetramer_conformation 1_symmetric) (59), PDB: 9NY1 (LmuA_tetramer conformation 2_asymmetric) (60), PDB: 9NY5 (LmuABC_apo) (61), and PDB: 9NYG (LmuABC–DNA) (62). Alignments, phylogenetic trees, DefenseFinder models and profiles, AF3 models, Lambda-vir genome and vector sequences are available in Dryad at https://doi.org/10.5281/zenodo.15120680 (63). All other data are included in the manuscript and/or supporting information.

ACKNOWLEDGMENTS. The Molecular Diversity of Microbes Lab at Institut Pasteur is grateful to its members and Enzo Poirier for their useful comments on earlier versions of the manuscript. We thank F.T. Hoffmann and G.D. Lampe for guidance and technical support on ChIP-seq experiments and data analysis. Several bioinformatic analyses were performed on the Core Cluster of the Institut Français de Bioinformatique(ANR-11-INBS-0013). M.H., R.L., M.A.A., and A.B. are supported by European Research Council Starting Grant (PECAN 101040529) and core funding from the Pasteur Institute. M.H. is funded by INSERM "Impact Santé" supported by Agence Nationale de la Recherche under France 2030 (ANR-24-RRII-0005, ANR-10-IDEX-0001-02 PSL), funding from Institut Curie. A.H.d.V. received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie Actions grant agreement No 101151697 and Human Frontier Science Program Long-Term Fellowship LT000454/2021-L. D.J.P. is supported by grants from the NIH (GM129430, AI141507, and GM145888), a Specialized Center of Research grant from the Leukemia and Lymphoma Foundation, the Maloris Foundation and Memorial Sloan-Kettering Core grant (P30-CA008748). Some of this work was performed at the National Center for CryoEM Access and Training and the Simons Electron Microscopy Center located at the New York Structural Biology Center, supported by the NIH Common Fund Transformative

High Resolution Cryo-Electron Microscopy program (U24 GM129539) and by grants from the Simons Foundation (SF349247) and NY State Assembly. We thank the Simons Electron Microscopy Center staff for help with data acquisition at their site. P.D.B.O. is supported by funding from NIH P41 GM109824 and P41 GM103314 to B.T.C. and Stavros Niarchos Foundation (SNF) as part of its grant to the SNF Institute for Global Infectious Disease Research at The Rockefeller University and funding from NIH P41 GM109824 and P41 GM103314 to B.T.C.

Author affiliations: [a]Institut Pasteur, Université Paris-Cité, CNRS UMR2535, Department of Genomes and Genetics, Molecular Diversity of Microbes, Paris 75015, France; [b]Institut Curie, Paris Sciences et Lettres Research University, Institut national de la santé et de la recherche médicale, Unit Immunity and Cancer U932, Innate Immunity in Physiology and Cancer Team, Paris 75005, France; [c]Structural Biology Program, Memorial Sloan-Kettering Cancer Center, New York, NY 10065; [d]Department of Biochemistry and Molecular Biophysics, Columbia University, New York, NY 10065; [e]Laboratory of Mass Spectrometry and Gaseous Ion Chemistry, Rockefeller University, New York, NY 10065; and [f]HHMI, Columbia University, New York, NY 10065

1. E. V. Koonin, K. S. Makarova, Y. I. Wolf, M. Krupovic, Evolutionary entanglement of mobile genetic elements and host defence systems: Guns for hire. *Nat. Rev. Genet.* **21**, 119–131 (2020).
2. H. Georjon, A. Bernheim, The highly diverse antiphage defence systems of bacteria. *Nat. Rev. Microbiol.* **21**, 686–700 (2023).
3. S. Zilberzwige-Tal *et al.*, Reprogrammable RNA-targeting CRISPR systems evolved from RNA toxin-antitoxins. *Cell* **188**, 1925–1940.e20 (2025).
4. H. Vaysset, C. Meers, J. Cury, A. Bernheim, S. H. Sternberg, Evolutionary origins of archaeal and eukaryotic RNA-guided RNA modification in bacterial IS110 transposons. *Nat. Microbiol.* **10**, 20–27 (2025).
5. F. Bürmann, J. Löwe, Structural biology of SMC complexes across the tree of life. *Curr. Opin. Struct. Biol.* **80**, 102598 (2023).
6. J. A. Raaijmakers *et al.*, SRBD1, a highly conserved gene required for chromosome individualization. *Cell Rep.* **44**, 115443 (2025).
7. B.-G. Lee, J. Rhodes, J. Löwe, Clamping of DNA shuts the condensin neck gate. *Proc. Natl. Acad. Sci. U.S.A.* **119**, e2120006119 (2022).
8. W. P. Robins, B. T. Meader, J. Toska, J. J. Mekalanos, DdmABC-dependent death triggered by viral palindromic DNA sequences. *Cell Rep.* **43**, 114450 (2024).
9. M. Jaskólska, D. W. Adams, M. Blokesch, Two defence systems eliminate plasmids from 7th pandemic *Vibrio cholerae*. *Nature* **604**, 323–329 (2022).
10. B. J. O'Hara, M. Alam, W.-L. Ng, The *Vibrio cholerae* seventh pandemic islands act in tandem to defend against a circulating phage. *PLoS Genet.* **18**, e1010250 (2022).
11. A. Krishnan, A. M. Burroughs, L. M. Iyer, L. Aravind, Comprehensive classification of ABC ATPases and their functional radiation in nucleoprotein dynamics and biological conflict systems. *Nucleic Acids Res.* **48**, 10045–10075 (2020).
12. S. Doron *et al.*, Systematic discovery of antiphage defense systems in the microbial pangenome. *Science* **359**, eaar4120 (2018).
13. Y. Kim, Z. Shi, H. Zhang, I. J. Finkelstein, H. Yu, Human cohesin compacts DNA by loop extrusion. *Science* **366**, 1345–1349 (2019).
14. T. L. Higashi *et al.*, A structure-based mechanism for DNA entry into the cohesin ring. *Mol. Cell.* **79**, 917–933.e9 (2020).
15. N. L. Rittenhouse, J. M. Dowen, Cohesin regulation and roles in chromosome structure and function. *Curr. Opin. Genet. Dev.* **85**, 102159 (2024).
16. Y. Yu *et al.*, Cryo-EM structure of DNA-bound Smc5/6 reveals DNA clamping enabled by multi-subunit conformational changes. *Proc. Natl. Acad. Sci. U.S.A.* **119**, e2202799119 (2022).
17. V. V. Rybenkov, V. Herrera, Z. M. Petrushenko, H. Zhao, MukBEF, a chromosomal organizer. *J. Mol. Microbiol. Biotechnol.* **24**, 371–383 (2014).
18. F. Bürmann, L. F. H. Funke, J. W. Chin, J. Löwe, Cryo-EM structure of MukBEF reveals DNA loop entrapment at chromosomal unloading sites. *Mol. Cell.* **81**, 4891–4906.e8 (2021).
19. B. Pradhan *et al.*, Loop-extrusion-mediated plasmid DNA cleavage by the bacterial SMC Wadjet complex. *Mol. Cell.* **85**, 107–116.e5 (2025).
20. A. Millman *et al.*, An expanded arsenal of immune systems that protect bacteria from phages. *Cell Host Microbe* **30**, 1556–1569.e5 (2022).
21. P. I. Hanson, S. W. Whiteheart, AAA+ proteins: Have engine, will work. *Nat. Rev. Mol. Cell Biol.* **6**, 519–529 (2005).
22. S. R. Casjens, R. W. Hendrix, Bacteriophage lambda: Early pioneer and still relevant. *Virology* **479–480**, 310–330 (2015).
23. H. W. Liu, F. Roisné-Hamelin, S. Gruber, Smc-based immunity against extrachromosomal DNA elements. *Biochem. Soc. Trans.* **51**, 1571–1583 (2023).
24. E. Kinoshita, E. van der Linden, H. Sanchez, C. Wyman, RAD50, an SMC family member with multiple roles in DNA break repair: How does ATP affect function? Chromosome Res. *Int. J. Mol. Supramol. Evol. Asp. Chromosome Biol.* **17**, 277–288 (2009).
25. B. M. Wendel, J. M. Cole, C. T. Courcelle, J. Courcelle, SbcC-SbcD and ExoI process convergent forks to complete chromosome replication. *Proc. Natl. Acad. Sci. U.S.A.* **115**, 349–354 (2018).
26. K.-P. Hopfner, Mre11-Rad50: The DNA end game. *Biochem. Soc. Trans.* **51**, 527–538 (2023).
27. E. Davidov, G. Kaufmann, RloC: A wobble nucleotide-excising and zinc-responsive bacterial tRNase. *Mol. Microbiol.* **69**, 1560–1574 (2008).
28. S. Rakesh, L. Aravind, A. Krishnan, Reappraisal of the DNA phosphorothioate modification machinery: Uncovering neglected functional modalities and identification of new counter-invader defense systems. *Nucleic Acids Res.* **52**, 1005–1026 (2024).
29. L. Käshammer *et al.*, Mechanism of DNA end sensing and processing by the Mre11-Rad50 complex. *Mol. Cell* **76**, 382–394.e6 (2019).
30. Y. Li *et al.*, Structure and activation mechanism of a Lamassu phage defence system. bioRxiv [Preprint] (2025). https://doi.org/10.1101/2025.03.14.643221 (Accessed 31 March 2025).
31. M. Better, D. Freifelder, Studies on the replication of *Escherichia coli* phage λ. DNA. I. The kinetics of DNA replication and requirements for the generation of rolling circles. *Virology* **126**, 168–182 (1983).
32. K. Kozlowska *et al.*, Formation of complexes between O proteins and replication origin regions of Shiga toxin-converting bacteriophages. *Front. Mol. Biosci.* **7**, 207 (2020).
33. M. Dodson *et al.*, Specialized nucleoprotein structures at the origin of replication of bacteriophage lambda: Localized unwinding of duplex DNA by a six-protein reaction. *Proc. Natl. Acad. Sci. U.S.A.* **83**, 7638–7642 (1986).
34. M. Zylicz, K. Liberek, A. Wawrzynow, C. Georgopoulos, Formation of the preprimosome protects lambda O from RNA transcription-dependent proteolysis by ClpP/ClpX. *Proc. Natl. Acad. Sci. U. S. A.* **95**, 15259–15263 (1998).
35. D. Bikard, C. Loot, Z. Baharoglu, D. Mazel, Folded DNA in action: Hairpin formation and biological functions in prokaryotes. *Microbiol. Mol. Biol. Rev.* **74**, 570–588 (2010).
36. B. Lowey *et al.*, CBASS immunity uses CARF-related effectors to sense 3′-5′– and 2′-5′-linked cyclic oligonucleotide signals and protect bacteria from phage infection. *Cell* **182**, 38–49.e17 (2020).
37. L. A. Gao *et al.*, Prokaryotic innate immunity through pattern recognition of conserved viral proteins. *Science* **377**, eabm4096 (2022).
38. E. V. Koonin, M. Krupovic, Evolution of adaptive immunity from transposable elements combined with innate immune systems. *Nat. Rev. Genet.* **16**, 184–192 (2015).
39. S. Hudaiberdiev *et al.*, Phylogenomics of Cas4 family nucleases. *BMC Evol. Biol.* **17**, 232 (2017).
40. M. Wilkinson *et al.*, Structures of RecBCD in complex with phage-encoded inhibitor proteins reveal distinctive strategies for evasion of a bacterial immunity hub. *eLife* **11**, e83409 (2022).
41. C. J. Schiltz, A. Lee, E. A. Partlow, C. J. Hosford, J. S. Chappie, Structural characterization of class 2 OLD family nucleases supports a two-metal catalysis mechanism for cleavage. *Nucleic Acids Res.* **47**, 9448–9463 (2019).
42. A. Bernheim *et al.*, Inhibition of NHEJ repair by type II-A CRISPR-Cas systems in bacteria. *Nat. Commun.* **8**, 2094 (2017).
43. S. Roth *et al.*, Rad50-CARD9 interactions link cytosolic DNA sensing to IL-1β production. *Nat. Immunol.* **15**, 538–545 (2014).
44. X. P. Peng, X. Zhao, The multi-functional Smc5/6 complex in genome protection and disease. *Nat. Struct. Mol. Biol.* **30**, 724–734 (2023).
45. C. M. Murphy *et al.*, Hepatitis B virus X protein promotes degradation of SMC5/6 to enhance HBV replication. *Cell Rep.* **16**, 2846–2854 (2016).
46. A. Perrin, E. P. C. Rocha, PanACoTA: A modular tool for massive microbial comparative genomics. *NAR Genomics Bioinf.* **3**, lqaa106 (2021).
47. F. Tesson *et al.*, Systematic and quantitative view of the antiviral arsenal of prokaryotes. *Nat. Commun.* **13**, 2561 (2022).
48. B. Néron *et al.*, MacSyFinder v2: Improved modelling and search engine to identify molecular systems in genomes. *Peer Community J.* **3**, e28 (2023).
49. F. Tesson *et al.*, A comprehensive resource for exploring antiphage defense: DefenseFinder webservice. Wiki and databases. *Peer Community J.* **4**, e91 (2024).
50. M. Steinegger, J. Söding, MMseqs2 enables sensitive protein sequence searching for the analysis of massive data sets. *Nat. Biotechnol.* **35**, 1026–1028 (2017).
51. J. Abramson *et al.*, Accurate structure prediction of biomolecular interactions with AlphaFold 3. *Nature* **630**, 493–500 (2024).
52. E. F. Pettersen *et al.*, UCSF ChimeraX: Structure visualization for researchers, educators, and developers. *Protein Sci. Publ. Protein Soc.* **30**, 70–82 (2021).
53. L. Holm, A. Laiho, P. Törönen, M. Salgado, DALI shines a light on remote homologs: One hundred discoveries. *Protein Sci.* **32**, e4519 (2023).
54. E. F. Pettersen *et al.*, UCSF Chimera–A visualization system for exploratory research and analysis. *J. Comput. Chem.* **25**, 1605–1612 (2004).
55. A. Chakravarti *et al.*, EMD-49911, Cryo-EM structure of LmuA_tetramer symmetric conformation. Electron Microscopy Data Bank. https://www.ebi.ac.uk/emdb/EMD-49911. Deposited 26 March 2025.
56. A. Chakravarti *et al.*, EMD-49915, Cryo-EM structure of LmuA_tetramer asymmetric conformation. Electron Microscopy Data Bank. https://www.ebi.ac.uk/emdb/EMD-49915. Deposited 26 March 2025.
57. A. Chakravarti *et al.*, EMD-49922, Cryo-EM structure of LmuABC-apo structure. Electron Microscopy Data Bank. https://www.ebi.ac.uk/emdb/EMD-49922. Deposited 26 March 2025.
58. A. Chakravarti *et al.*, EMD-49934, Cryo-EM structure of LmuABC-DNA structure. Electron Microscopy Data Bank. https://www.ebi.ac.uk/emdb/EMD-49934. Deposited 27 March 2025.
59. A. Chakravarti *et al.*, 9NXX, Cryo-EM structure of LmuA_tetramer symmetric conformation. Protein Data Bank. https://www.rcsb.org/structure/unreleased/9NXX. Deposited 26 March 2025.
60. A. Chakravarti *et al.*, 9NY1, Cryo-EM structure of LmuA_tetramer asymmetric conformation. Protein Data Bank. https://www.rcsb.org/structure/unreleased/9NY1. Deposited 26 March 2025.
61. A. Chakravarti *et al.*, 9NY5, Cryo-EM structure of LmuABC-apo structure. Protein Data Bank. https://www.rcsb.org/structure/unreleased/9NY5. Deposited 26 March 2025.
62. A. Chakravarti *et al.*, 9NYG, Cryo-EM structure of LmuABC-DNA structure. Protein Data Bank. https://www.rcsb.org/structure/unreleased/9NYG. Deposited 27 March 2025.
63. M. Haudiquet *et al.*, Alignments, phylogenetic trees, DefenseFinder models and profiles, AF3 models, Lambda-vir genome and vector sequence. Zenodo. https://doi.org/10.5281/zenodo.15120680. Deposited 4 November 2025.