

1 **Title**

2 Co-option and Detoxification of a Phage Lysin for Housekeeping Function

3

4 **Authors**

5 Amelia M. Randich, Indiana University, Bloomington, IN USA

6 David T. Kysela, Bloomington, IN USA

7 Cécile Morlot, Institut de Biologie Structurale (IBS), Université Grenoble Alpes, CEA, CNRS,

8 France

9 Yves V. Brun, Indiana University, Bloomington, IN USA

10

11 Correspondence: Y.V.B. ybrun@indiana.edu

12

13 **Summary**

14

15 Temperate phages constitute a potentially beneficial genetic reservoir for bacterial innovation  
16 despite being selfish entities encoding an infection cycle inherently at odds with bacterial fitness.  
17 These phages integrate their genomes into the bacterial host during infection, donating new, but  
18 deleterious, genetic material: the phage genome encodes toxic genes, such as lysins, that kill the  
19 bacterium during the phage infection cycle. Remarkably, some bacteria have exploited the  
20 destructive properties of phage genes for their own benefit by co-opting them as toxins for  
21 functions related to bacterial warfare, virulence, and secretion. However, do toxic phage genes  
22 ever become raw material for functional innovation? Here we report on a toxic phage gene  
23 whose product has lost its toxicity and has become a domain of a core cellular factor, SpmX,  
24 throughout the bacterial order Caulobacterales. Using a combination of phylogenetics,  
25 bioinformatics, structural biology, cell biology, and biochemistry, we have investigated the  
26 origin and function of SpmX and determined that its occurrence is the result of the detoxification  
27 of a phage peptidoglycan hydrolase gene. We show that the retained, attenuated activity of the  
28 phage-derived domain plays an important role in proper cell morphology and developmental  
29 regulation in representatives of this large bacterial clade. To our knowledge, this is the first  
30 observation of phage gene domestication in which a toxic phage gene has been co-opted for a  
31 housekeeping function.

32

33 **Keywords**

34 Lysozyme, GH24, Prophage domestication, Bacterial evolution, Alphaproteobacteria,

35 *Caulobacter*, *Asticcacaulis*

36

37 **Introduction**

38

39 Understanding how new genes arise is key to studying the forces that drive diversity and  
40 evolution. Although horizontal gene transfer (HGT) is widely regarded as an important  
41 mechanism for the exchange of existing genes among bacteria, mobile genetic elements can  
42 transfer exogenous genetic material that gives rise to novel genes. These new genes can provide  
43 the basis for evolving new traits and propelling evolutionary transitions [1,2]. Temperate  
44 bacteriophages mediate genetic transfer by integrating their genomes into bacterial hosts [3–6].  
45 These integrated tracts of genes, called prophages, remain dormant until induced by various  
46 signals to produce phage particles and phage-encoded proteins that lyse the cell. In many cases,

47 prophages contain genes that benefit the host, promoting retention of prophages in many  
48 bacterial lineages, even after mutations have inactivated the prophages and they no longer are  
49 capable of producing phage [7–9]. Accumulation of host-specific beneficial mutations in  
50 prophages has been referred to as “domestication.” Many so-called domesticated segments of  
51 inactivated prophages unexpectedly contain lytic and phage structural genes, which would  
52 intuitively be of little use or even detrimental to the bacterial host [7]. These genes are often used  
53 by the bacteria as weapons against competing bacteria and eukaryotic hosts [10–14]. In contrast,  
54 we have identified an instance in which a toxic phage gene has not been repurposed as a weapon,  
55 but has rather been incorporated into a bacterial housekeeping factor, *spmX*. Here, we report that  
56 SpmX resulted from an ancient domestication event at the root of the alphaproteobacterial order  
57 Caulobacterales, in which co-option and detoxification of a toxic phage gene gave rise to a novel  
58 bacterial gene with roles in developmental regulation and morphogenesis.

59 SpmX was first identified as a developmental regulator in the model organism  
60 *Caulobacter crescentus* [15]. Like most members of Caulobacterales, stalked *C. crescentus* cells  
61 divide asymmetrically to produce a “mother” cell with a polar stalk and a motile “daughter” or  
62 “swarmer” cell with a polar flagellum. The *Caulobacter* developmental cycle depends on strict  
63 coordination of cell growth, chromosome replication and segregation, and division by various  
64 regulatory proteins that differ in localization and timing [16]. This network depends on  
65 regulatory phospho-signaling factors localized and regulated by polar scaffolds. SpmX is one  
66 scaffold that localizes at the old pole during the swarmer-to-stalked cell transition and recruits  
67 and potentially activates the histidine kinase DivJ [15]. Intriguingly, SpmX is required for stalk  
68 synthesis initiation and elongation in the closely related *Asticcacaulis* species *A. excentricus* and  
69 *A. biprosthecum* [17]. Therefore, this gene appears to have evolved multiple roles within this  
70 family of dimorphic, stalked bacteria.

71 Perplexingly, SpmX contains an N-terminal putative phage muramidase domain that is  
72 generally toxic to bacteria. This enzymatic domain is typically used by phage to lyse bacteria and  
73 release infectious phage particles during the lytic cycle. As a part of SpmX, this domain is  
74 critical for SpmX’s role in both developmental regulation and stalk biogenesis. In particular, the  
75 muramidase domain is necessary for proper localization of SpmX in both *C. crescentus* [15,18]  
76 and the *Asticcacaulis* genus [17]. Various studies have shown that SpmX localizes with the polar  
77 scaffold PopZ in *C. crescentus* [19,20], and this interaction occurs entirely through the  
78 muramidase domain [18]. A previous study was unable to measure enzymatic activity from  
79 purified *C. crescentus* SpmX muramidase domain, and consequently concluded that the domain  
80 was repurposed for protein interactions and oligomeric assembly [18]. However, given the  
81 remarkable sequence similarity of the SpmX muramidase domain to functional phage lysozymes,  
82 including the canonical catalytic glutamate, it seems unlikely that its enzymatic activity is  
83 entirely lost. Why would this domain be so highly conserved if its new function were merely for  
84 non-essential protein-protein interactions?

85 To better characterize the SpmX muramidase domain and the constraints underlying its  
86 conservation, we performed an in-depth bioinformatics study of more than 60 available SpmX  
87 genes together with structural determination, biochemical analysis, and comparative cell biology  
88 between *Caulobacter* and *Asticcacaulis*. We show that SpmX arose prior to the diversification of  
89 Caulobacterales, a large class of stalked bacteria. We establish that the SpmX muramidase  
90 domain is a close relative of GH24 (T4-like) autolysin/endolysins that have been laterally  
91 exchanged throughout Proteobacteria via prophages. We find that the SpmX domain is still an  
92 active muramidase, albeit with attenuated ancestral phage activity, consistent with its remodeled

93 active cleft. Finally, we demonstrate that the enzymatic activity is necessary for SpmX function  
94 in three representative species. We conclude that, close to the time of the genesis of the full-  
95 length *spmX* gene, this co-opted domain accumulated mutations that attenuated its hydrolytic  
96 activity on peptidoglycan and detoxified it for bacterial use. To our knowledge, this is the first  
97 case of phage gene domestication in which a toxic phage gene has been incorporated into a new  
98 bacterial gene with housekeeping function.

99

## 100 Results

101

102 **The SpmX muramidase domain was co-opted from prophage in an early Caulobacterales**  
103 **ancestor.** Our first step in characterizing the muramidase domain of SpmX was to determine the  
104 prevalence of SpmX and its homologues in the bacterial domain. Simple pBLAST analysis  
105 revealed that SpmX, as defined by its three-part architecture with an N-terminal muramidase  
106 domain, a charged and proline-rich intermediate domain, and two C-terminal transmembrane  
107 (TM) segments (**Fig. 1A**), is taxonomically constrained to Caulobacterales and one member of  
108 its sister taxa, Parvularculales. It is conserved as a single-copy gene in all sequenced members  
109 (**Table S1**). In all 69 identified *spmX* orthologues, the muramidase domains exhibit high amino  
110 acid sequence conservation (**Fig. S1**), the TM segments moderate sequence conservation, and the  
111 intermediate domains high variability in length and composition among genera (**Fig. 1A**). Apart  
112 from these SpmX orthologues, BLAST searches using SpmX only returned hits for the  
113 muramidase domain. These hits came from Gram-negative bacterial genomes that span the entire  
114 bacterial domain, particularly proteobacteria, and from viral genomes. Most of these bacterial  
115 genes are likely to be in prophage regions, as evidenced by their position in long runs of  
116 predicted prophage genes. We did not detect sequences homologous to SpmX TMs in our search,  
117 although we occasionally detected homologous phage muramidase domains fused to other, non-  
118 homologous TM segments.

119 Consistent with finding close relatives of SpmX muramidase in prophages, NCBI's  
120 Conserved Domain Database (CDD) tool [21,22] clustered SpmX muramidase with glycoside  
121 hydroloase (GH) 24 lysozymes in the autolysin/endolysin class (**Fig. S2**). These enzymes are  
122 closely related to classical phage T4 lysozyme-like (T4L-like) peptidoglycan hydrolases. Phages  
123 generally use lysozymes of this class to cleave peptidoglycan and lyse cells during the lytic  
124 cycle. These lysozymes are distinct from the other well-known phage Lambda lysozyme-like  
125 lytic transglycosylases, which are related to known housekeeping bacterial hydrolases with roles  
126 in cell growth and division. Lytic transglycosylases are also assigned to the GH24 group but  
127 share no sequence similarity with and catalyze a different peptidoglycan cleaving reaction than  
128 T4L-like muramidases [23–25]. Thus although core bacterial genomes encode peptidoglycan  
129 hydrolases, the muramidase domain of SpmX is most closely related to peptidoglycan hydrolases  
130 encoded by prophages.

131 Unlike its close relatives that have been transferred horizontally through the bacterial  
132 domain via prophage, the SpmX muramidase domain coding region has been inherited vertically  
133 as part of the *spmX* gene in Caulobacterales. The SpmX gene tree mirrors the phylogeny of  
134 Caulobacterales from concatenated gene alignments (**Fig. 1B**). None of the SpmX genes appear  
135 in tracks of prophage genes, and the genomic context of *spmX* appears to be well maintained in  
136 members of Caulobacterales, with the gene occurring between a putative Mg<sup>2+</sup> transporter and a  
137 putative isovaleryl-CoA dehydrogenase in nearly all species. Together, these findings suggest  
138 that SpmX muramidase is a peptidoglycan hydrolase derived from prophage that has become a

139 domain in a bacterial protein that is no longer within prophage and is under direct cellular  
140 control. It likely fused with the intermediate and TM domains in a common ancestor of  
141 Parvularculales and Caulobacterales. The vertical transmission of SpmX and strong sequence  
142 conservation of the muramidase domain suggests an important function in members of  
143 Caulobacterales.

144  
145 **The SpmX muramidase domain retains the canonical GH24 motif but contains mutations**  
146 **in the catalytic cleft known to inactivate phage lysozymes.** To determine if critical enzymatic  
147 residues in SpmX muramidase were conserved, we compared the amino acid sequences of SpmX  
148 and other lysozymes. By definition, lysozymes catalyze the hydrolysis of  $\beta$ 1,4-linked glycosidic  
149 bonds in peptidoglycan and chitin [25]. This superfamily includes at least seven distinct groups  
150 that are unrelated by sequence similarity but share a common fold in which the catalytic Glu and  
151 the beta-hairpin motif in the N-terminal lobe form the catalytic cleft by packing against the C-  
152 terminal lobe (**Fig. 2A**) [26]. This beta-hairpin, or GH motif, contains family-specific residues  
153 critical for enzyme activity in all members of the lysozyme superfamily [26].

154 We compared the GH motif of SpmX to lysozymes from the T4L and endolysin/autolysin  
155 classes with varying degrees of relatedness. **Figure 2B** shows the amino acid conservation in the  
156 GH motif of T4L-like, autolysin/endolysin, closely related non-SpmX muramidase, and SpmX  
157 muramidase protein sequences. Because the autolysin/endolysin class and the closest non-SpmX  
158 relatives are likely to be active phage enzymes, highly conserved residues shared by these groups  
159 with T4L delineate positions that are evolutionarily constrained for phage lysozyme activity and  
160 stability. For example, D10 is not conserved outside of T4L-like enzymes because the  
161 autolysin/endolysin class does not have a salt bridge between D10 and the C-terminal lobe [27].  
162 On the other hand, all of the putative phage sequences (**Fig. 2B(i-iii)**) conserve the “catalytic  
163 triad” of T4 lysozymes: the catalytic residue E11 and active site residues D20 and T26. While  
164 the exact roles of D20 and T26 are not clear, they are posited as being critical for effective  
165 catalysis [26–29]. Position D20 is very sensitive to mutation, with only substitutions D20C/A  
166 retaining the hydrolytic activity of T4L or P22 phage lysozymes [30]; these substitutions are  
167 tellingly well represented amongst the putative phage sequences. Remarkably, SpmX  
168 muramidase domains demonstrate strong conservation of residues required for T4L stability and  
169 creating the GH fold, but low conservation of residues associated with catalysis, with the  
170 exception of the main catalytic residue, E11 (**Fig. 2B(iv)**). The majority of SpmX genes contain  
171 the mutation D20L/R, both of which reduced T4L activity to less than 3% of WT in previous  
172 studies [27] and which are distinctly unrepresented in the other phage muramidases. Moreover,  
173 the T26 position no longer appears to be under constraint in SpmX. The conservation of the GH  
174 motif coupled with the apparent inactivation of the catalytic triad across all SpmX genes suggests  
175 that muramidase function may not remain fully active.

176  
177 **The SpmX muramidase domain has a wider, more dynamic catalytic cleft than related**  
178 **phage lysins.** Obtaining the structure of the SpmX muramidase domain (residues 1-150) from  
179 *Asticcacaulis excentricus* (SpmX-Mur-*Ae*) allowed us to directly visualize the effect of the D20L  
180 and T26X mutations on the catalytic cleft. We solved the X-ray crystal structure by molecular  
181 replacement using the atomic coordinates of phage P22 lysozyme (PDB 2ANX) to 1.9 Å  
182 resolution ( $R_{\text{cryst}}$  of 21.1%,  $R_{\text{free}}$  of 25.5%,) (**Table S2**). SpmX-Mur-*Ae* exhibited the expected  
183 structural similarity with the phage P22 lysozyme (root mean square deviation (rmsd) of 1.7 Å  
184 and 40% identity over 141 aligned C $\alpha$  atoms, Dali Z-score of 21.5). It therefore shares the

185 characteristic structure of T4 lysozymes, with the predicted catalytic glutamate occurring at the  
186 C-terminal end of the first alpha-helix and situated in the catalytic cleft formed between the N-  
187 and C-terminal lobes (**Fig. 3A**). The active conformation of the distantly related SAR endolysin  
188 protein R21 (PDB 3HDE) from bacteriophage P21 (**Fig. S2**) is included in the structural  
189 alignment in **Figure 3A** to emphasize the ways in which the SpmX muramidase domain deviates  
190 from phage lysozymes: SpmX-Mur-*Ae* contains an extended beta-hairpin in the C-terminal lobe  
191 and, more importantly, the canonical GH beta-hairpin in the N-terminal lobe of SpmX-Mur-*Ae*  
192 splays away from the catalytic cleft relative to the GH beta-hairpins of the phage lysozymes.  
193 There is additional evidence for the flexibility of the GH beta-hairpin in the various  
194 conformations observed in the crystallography data, as among the three molecules of SpmX-  
195 Mur-*Ae* in the asymmetric unit, this region exhibited the most conformational differences. The  
196 overlay of the three chains in **Figure 3B** illustrates how the orientation of the GH beta-hairpin is  
197 tilted by about 16° between chains A and B; moreover, chain A appears to have lost some of its  
198 beta-strand secondary structure. That SpmX-Mur-*Ae* is capable of both conformations suggests a  
199 heightened flexibility in this region compared to other T4L-like lysozymes, which may reduce  
200 the ability of the enzyme to coordinate the hydrolysis of peptidoglycan in the catalytic cleft.

201 Sequence comparisons of the GH motif in **Figure 2B** show that SpmX muramidase  
202 domains have lost a highly conserved tyrosine residue at position 18. Although T4L enzymatic  
203 activity is not sensitive to mutation at this position [27], it is invariant across all the phage  
204 lysozyme classes we analyzed. Visualization of Y18 in the P22 lysozyme structure (**Fig. 3C**)  
205 shows that it interacts with R14 at the base of the beta-hairpin, possibly a critical interaction for  
206 coordinating the beta-hairpin with the catalytic glutamate. In SpmX-Mur-*Ae*, Y18S still appears  
207 to make hydrogen-bonding contact with R14; however, most SpmX muramidase domains have  
208 non-polar residues at position 18 (**Fig. 2B(iv)**), which may reduce coordination between the  
209 catalytic glutamate and the GH beta hairpin. It has been previously shown that the Y18 position  
210 is a hot-spot for compensatory mutations that restore activity to inactive catalytic mutants [31],  
211 and it is intriguing to imagine that mutations at this position in SpmX muramidase are the result  
212 of a remodeled catalytic cleft.

213 **Figure 3** shows the catalytic clefts of both P22 lysozyme (**D**) and SpmX-Mur-*Ae* (**E**). In  
214 P22 lysozyme, the E11 carbonyl, D20 carboxyl, and T26 hydroxyl groups point into the aqueous  
215 catalytic cleft. In SpmX-Mur-*Ae*, the cleft is slightly reorganized, with the S-methyl thioether of  
216 T26M still within 20 Å of E11 and potentially capable of interacting with peptidoglycan. In  
217 about two thirds of the SpmX genes, position 26 has either a valine or an isoleucine, which have  
218 no polar moieties to contribute to the cleft (**Fig. 2B(iv)**). With the structural data from SpmX-  
219 Mur-*Ae*, we can infer that the SpmX muramidase domain has a remodeled catalytic cleft with a  
220 correctly positioned catalytic glutamate. However, the increased flexibility between the GH  
221 motif and the glutamate, as well as the loss of key coordinating residues might reduce, if not  
222 eliminate, hydrolytic activity of SpmX.

223  
224 **SpmX retains reduced hydrolytic activity on peptidoglycan.** To determine if the SpmX  
225 muramidase domain was capable of binding peptidoglycan, various constructs from *C.*  
226 *crescentus*, *A. excentricus*, and *A. biprosthicum* were purified and incubated with sacculi isolated  
227 from the three species. Both muramidase domains and entire soluble domains (muramidase with  
228 intermediate domain) from all three species, including the E11A and N105R mutants of SpmX  
229 from *C. crescentus* bound sacculi from all three species (**Fig. S3**). Despite previous reports of  
230 inactivity [18], we found that purified SpmX muramidase exhibited hydrolytic activity. We used

231 remazol brilliant blue (RBB) assays, which are standard for measuring lysozyme activity [32], to  
232 compare the activity of SpmX muramidase from *C. crescentus* (SpmX-Mur-*Cc*) to that of the  
233 P22 lysozyme (P22Lyso) and a P22Lyso mutant with a similarly inactivated catalytic cleft  
234 (P22Lyso-D20L) (**Fig. 4A**). The activity curves demonstrate that both SpmX-Mur-*Cc* and  
235 P22Lyso-D20L exhibit similarly attenuated hydrolytic activity in comparison to P22Lyso: both  
236 reached maximal levels of RBB release near enzyme concentrations of 15  $\mu$ M while P22Lyso  
237 reached the same levels near 5  $\mu$ M. Mutants in which the catalytic glutamate was replaced with  
238 alanine (SpmX-Mur-*Cc*-E11A and P22Lyso-E11A) did not exhibit activity (**Fig. S4A**). Restoring  
239 the ancestral D20 (SpmX-Mur-*Cc*-L20D) did not increase SpmX activity *in vitro* (**Fig. S4B**).  
240 These data indicate that the “inactivating” substitution D20L attenuates enzymatic activity  
241 whereas mutating the catalytic glutamate abolishes it altogether.

242 We were surprised by the activity of P22Lyso-D20L and SpmX-Mur given the reported  
243 inhibition of lytic activity of T4L with D20 mutations in phage plaque assays [27]. We were  
244 interested in the biological implications of this *in vitro* activity, and whether the attenuated  
245 mutant was active in the periplasmic environment. We therefore designed an experimental  
246 system to test the activity of P22Lyso and SpmX-Mur-*Cc* mutants in the periplasmic space of *E.*  
247 *coli*. In this system, we expressed the various muramidase constructs with N-terminal PelB  
248 leader sequences (pET22b) for periplasmic expression in Lemo21(DE3) cells, which allow  
249 tunable expression of toxic products (see Methods). Lemo21(DE3) cells carrying P22Lyso lysed  
250 even in the absence of induction (**Fig. 4B**). In contrast, Lemo21(D32) cells carrying P22Lyso-  
251 D20L lysed only after induction, and strains carrying SpmX-Mur-*Cc* and P22Lyso-E11A never  
252 lysed (**Fig. 4C** and **S4C**) despite equivalent periplasmic expression levels (**Fig. S4D**). Different  
253 growth conditions and media did not affect the viability of cells expressing SpmX-Mur-*Cc* (**Fig.**  
254 **S4C**). Moreover, SpmX-Mur-*Cc* was active on sacculi isolated from Lemo21(DE3), eliminating  
255 the possibility that it cannot cleave *E. coli* peptidoglycan (**Fig. S4E**). It remains unclear why the  
256 periplasmic activity of SpmX-Mur-*Cc* does not match its *in vitro* activity, as it appears to have a  
257 similar activity to P22Lyso-D20L and yet surpasses the periplasmic levels at which P22Lyso-  
258 D20L lyses *E. coli*. It is possible that SpmX-Mur has additional mutations that either make it  
259 difficult to fold correctly in the *E. coli* periplasm, or inactivate its activity in the periplasmic  
260 environment. However, the periplasmic expression tests in *E. coli* confirm that the D20L  
261 mutation attenuates the hydrolytic activity of P22 lysozyme. This tuning of enzymatic activity  
262 might have served as a critical detoxifying step in the co-option of the muramidase domain from  
263 phage. Moreover, such attenuation helps explain why the D20L substitution is absent from phage  
264 despite broad conservation among *spmX* alleles. Because SpmX has retained the ancestral  
265 catalytic glutamate and its modified catalytic cleft is capable of hydrolytic activity, we conclude  
266 that this attenuated activity is under purifying selection in SpmX and important for SpmX  
267 function.

268  
269 **Inactivating the muramidase domain impacts SpmX localization *in vivo*.** To investigate the  
270 role of the conserved catalytic glutamate in SpmX function *in vivo*, chromosomal E11A mutants  
271 (E19A in SpmX numbering) were made in *C. crescentus*, *A. excentricus*, and *A. biprosthicum*  
272 with eGFP gene fusions (**Fig. 5**). We determined the effects of the E11A mutation on cellular  
273 morphology, as the  $\Delta$ *spmX* strains in all three species have morphological phenotypes (**Fig.**  
274 **5ABCii**). In *C. crescentus*,  $\Delta$ *spmX* cells have a characteristic elongated morphology resulting  
275 from failed division cycles and often grow stalks prematurely from daughter cells that fail to  
276 divide completely (**Fig. 5Aii**) [15]. In *Asticaccaulis*,  $\Delta$ *spmX* cells lack stalks and do not appear to

277 have other developmental phenotypes (**Fig. 5BCii**) [17]. In *C. crescentus*, the E11A mutant  
278 population contained both WT-like cells and cells exhibiting the division defect, but with less  
279 severity than in  $\Delta spmX$  (**Fig. 5Aiii**). In both *Asticcacaulis* species, the E11A mutants still grew  
280 stalks (**Fig. 5BCiii**). Nevertheless, the *A. biprosthecum* E11A mutant exhibited a significant loss  
281 of bilateral stalks (3.5 fold reduction) and an increase in the frequency of cells with a single stalk  
282 (**Fig. S5D**).

283 We used GFP fusions of WT and mutant SpmX to monitor changes in SpmX  
284 localization. In all species, WT SpmX localizes at the future position of the stalk, either at the  
285 pole as in *C. crescentus*, or at sub-polar or bilateral positions in *Asticcacalis*, and is retained at  
286 this position (**Fig. 5ABCiii**). Both *C. crescentus* and *A. biprosthecum spmX* E11A mutants  
287 exhibited an increase in mislocalized SpmX throughout the cell body compared to WT (**Fig.**  
288 **5ABiv**). Quantification of the fluorescence data indicated that while the overall mean  
289 fluorescence of the cells were the same as WT, the SpmX foci were significantly less intense in  
290 the mutants (**Fig. S5AB**). We also observed a 3X increase in SpmX E11A in the stalks of *A.*  
291 *biprosthecum* compared to WT (**Fig. S5B**). Although no difference in focal fluorescence  
292 intensity was observed in *A. excentricus spmX* E11A mutant cells, more cells had a second  
293 SpmX focus at the tip of the stalk than WT cells (**Fig. 5iv, S5C**), indicating altered localization.  
294 Together these data show that the E11A mutation impacts SpmX localization in all three species.

295 N/Q105 has been shown to coordinate peptidoglycan in the active cleft of T4L [33] and  
296 the mutation Q105R abolished activity in T4 phage plaque assays [27]. The mutation N105R  
297 (N91R in SpmX numbering) in *C. crescentus* and *A. biprosthecum* resulted in similar  
298 delocalization phenotypes as E11A (**Fig. 5BCiiv**). Western blots of cells expressing WT SpmX-  
299 eGFP and SpmX mutants indicated that the delocalization was not due to clipping (**Fig. S5E**).  
300 Overall, these data show that eliminating the catalytic glutamate affects SpmX localization and  
301 function. Although it is not clear whether the E11A mutation specifically affects the potential  
302 hydrolytic activity or the peptidoglycan-binding ability, the similar phenotype from mutating a  
303 predicted peptidoglycan-interacting residue (N105R) underscores the importance of SpmX-  
304 peptidoglycan interactions. That the catalytic mutant has an intermediate morphological  
305 phenotype in *C. crescentus* and one *Asticcacaulis* species indicates that the muramidase domain  
306 may coordinate SpmX functions similarly in the two genera and that this function may rely on its  
307 enzymatic activity.

308  
309 **Replacing the muramidase domain, or removing it, interferes with native SpmX protein**  
310 **levels *in vivo*.** In an effort to understand the role of the highly conserved muramidase fold in  
311 SpmX function, we made chimeras wherein P22 lysozyme replaced the muramidase domain of  
312 SpmX in *C. crescentus* (P22Lyso-SpmX). While similarly made constructs with SpmX and the  
313 SpmX-E11A mutant exhibited the previously determined morphological and delocalization  
314 phenotypes (**Fig. 6Aii-iii**), replacing the SpmX muramidase domain with P22 lysozyme  
315 phenocopied the parent  $\Delta spmX$  strain. However, we were unable to detect any sfGFP-fusion  
316 products in these chimeras by Western blot (**Fig. 6B**). We confirmed that P22Lyso-SpmX had  
317 been correctly inserted in these strains by sequencing, suggesting that the chimeras were  
318 expressed but degraded quickly in *C. crescentus*. Therefore the phenotype of this chimera is due  
319 to loss of SpmX and not the addition of the P22 lysozyme domain. Inactivation of P22Lyso  
320 (E11A) did not change the outcome, suggesting that the toxicity of the phage muramidase was  
321 not driving degradation of SpmX. Deletion of the muramidase domain from the *spmX* locus in all  
322 three species also resulted in strains that failed to produce detectable amounts of  $\Delta mur$ -SpmX-

323 sfGFP by Western blot (**Fig. S5E**). These results suggest that the SpmX muramidase domain is  
324 necessary to produce and/or maintain WT levels of SpmX in all three species, and that phage  
325 muramidase P22Lyso, with high sequence similarity (51%) and structural homology (RMSD 1.7  
326 Å), is not sufficient to replace it.

327 Given that P22Lyso and SpmX muramidase are fairly distant relatives despite high  
328 sequence conservation, we tested the ability of other SpmX muramidase domains to replace that  
329 of *C. crescentus*. First, we attempted to restore the phage active site D20 (L28D in SpmX  
330 numbering) in the native SpmX gene, but this had no evident effect on SpmX localization or cell  
331 morphology (**Fig. 6iv**). Previously, the muramidase domains of *C. crescentus* and *Asticcacaulis*  
332 were shown to be interchangeable [17], so we extended the sequence distance to SpmX  
333 muramidases from the next closest relative *B. subvibrioides*, which has D20R in the catalytic  
334 cleft, and the most distant relative *P. bermudensis*, which shares the D20L mutation. We found  
335 that the muramidase domain from *B. subvibrioides* supported the WT phenotype in *C. crescentus*  
336 (**Fig. 6v**), but that the SpmX muramidase domain from *P. bermudensis* did not. We were  
337 surprised to see no evidence of delocalization in the *B. subvibrioides* SpmX chimera because the  
338 L20R mutant of SpmX in *C. crescentus* showed some delocalization (**Fig. 6vi**). This result  
339 suggests that the L20R mutation in the brevundimonads must coexist with other compensatory  
340 mutations. The SpmX muramidase domain from *P. bermudensis*, like P22Lyso, must be too  
341 distant from *C. crescentus* to support WT levels of SpmX protein. In combination with data from  
342 the P22Lyso chimeras, these data indicate that a T4L GH fold alone is not sufficient for SpmX  
343 function, and that the SpmX muramidase domain must contain mutations necessary for stable  
344 protein levels in Caulobacterales. It also suggests that this domain has additional constraints on it  
345 unrelated to potential peptidoglycan interactions.

## 346 347 **Discussion**

348  
349 Bacteriophages shape bacterial evolution in various ways: they increase bacterial diversity by  
350 selectively preying on species [2,34], drive horizontal gene transfer of phage and bacterial genes  
351 [8,35], and serve as reservoirs of raw material for new bacterial genes [36,37]. Phages are  
352 heralded as a major source of genetic material for novel gene emergence in bacteria [2,36,37],  
353 but, as we discuss later in this section, very few examples of novel gene emergence from  
354 prophage exist in the literature. We have investigated the origin and function of a taxonomically  
355 restricted gene from Caulobacterales, *spmX*, and determined that its occurrence is the result of  
356 the fusion and domestication of a peptidoglycan hydrolase gene commonly found in prophage.  
357 Despite the previously demonstrated role of SpmX as a scaffold in developmental regulation and  
358 morphology, the muramidase domain retains high sequence similarity to phage lysozymes,  
359 which are toxic to bacteria. The active cleft contains mutations that have attenuated the toxic  
360 activity of the domain, presumably making it available for genetic innovation and bacterial use.  
361 We show here that the domain remains enzymatically active on peptidoglycan and that reducing  
362 or eliminating this activity alters the function of the full-length protein *in vivo*. Thus, the SpmX  
363 gene represents a bacterial gene innovation, specific to the Caulobacterales order and with  
364 housekeeping function, that originally arose from a prophage gene with antibacterial activity.

365 Previously, it was suggested that the SpmX muramidase domain functions only in  
366 protein-protein and self-oligomerizing interactions in SpmX's role as a developmental regulator  
367 and scaffold in *C. crescentus* [18]. This conclusion was based on the lack of detectable activity  
368 from the purified domain and the inability of the catalytic E11R (E19R in SpmX numbering)



369 mutant to self-oligomerize. The conservation of the GH fold and the catalytic glutamate clearly  
370 suggests that the activity of SpmX is important for its domesticated function, and, importantly,  
371 we were able to observe muramidase activity *in vitro* with purified protein. It is highly likely that  
372 the E11R mutation greatly destabilizes the structure of the muramidase domain. We found that  
373 the E11A mutant eluted in multiple fractions during purification, indicating decreased  
374 conformational stability. Moreover, the E11R mutant exactly phenocopies the  $\Delta$ mur-SpmX  
375 phenotype *in vivo* in that the protein product was no longer detectable in the cells expressing the  
376 gene [18]. In this sense, the effect of the E11R mutation is similar to using a distantly related  
377 muramidase domain (like P22Lyso) or deleting portions of the muramidase domain entirely.  
378 These data indicate that the muramidase domain plays an unanticipated role in maintaining stable  
379 levels of SpmX protein across all tested species: without the muramidase domain, the protein is  
380 misfolded or misprocessed and is therefore quickly degraded.

381 Inactivating the SpmX muramidase domain resulted in developmental defects in *C.*  
382 *crescentus* and significant loss of the bilateral stalk morphology in *A. biprosthecum*. It is curious  
383 that inactivating the enzymatic domain did not yield a null phenotype or complete delocalization.  
384 Since SpmX-E11A and N105R still bind to sacculi *in vitro*, it is possible that enough  
385 peptidoglycan-interactions are maintained in the mutants for partial SpmX function. It is also  
386 possible that SpmX recruits another protein with redundant enzymatic activity that cannot be  
387 recruited in the  $\Delta$ spmx mutant, as it is already known that SpmX interacts with targeting factors  
388 via its C-terminal domains in *Asticcacaulis* [17] and possibly via its transmembrane segments  
389 with DivJ [15]. Finally, it is hard to distinguish whether there is a direct relationship between  
390 catalytic activity and peptidoglycan binding, and therefore localization, or if cleavage of  
391 peptidoglycan by SpmX could indirectly impact SpmX localization. The multiple domains and  
392 pleiotropic effects of SpmX make it difficult to assess the effects of an individual domain on its  
393 *in vivo* function. However, our data support a model in which the muramidase domain of SpmX  
394 is still active, and this activity is used to localize SpmX directly or indirectly.

395 SpmX emerges in the genomic record at the root of Caulobacterales with the attenuating  
396 D20L mutation already in the muramidase domain (**Fig. 1B**). The D20L mutation is therefore  
397 ancestral and potentially the initial step in the co-option of the domain. The conservation of  
398 D20L throughout most of Caulobacterales suggests evolutionary constraint on this position  
399 despite no observable phenotype from the SpmX-Mur-*Cc* L20D reversion mutation *in vivo* or *in*  
400 *vitro*. There are several modifications to the active cleft, including the loss of selection on the  
401 third catalytic triad position, T26, and the invariant residue Y18. This pair is interesting in that  
402 Y18 was identified as a hotspot for spontaneous second site revertants of T26 mutants in T4L  
403 [31]. It is possible that the scatter we see at these two positions are compensatory mutations  
404 retaining attenuated activity, although a clear history of covariation is not clear. In two groups,  
405 SpmX has diverged from the ancestral D20L mutation: the marine genera *Oceanicaulis* and  
406 *Maricaulis* (D20R/G), and the freshwater/soil genus *Brevundimonas* (D20R) (**Fig. 1B, S1**).  
407 Interestingly, D20R is covariant with residue N105 (**Fig. S1**), which is a peptidoglycan-  
408 interacting residue in T4 lysozyme [33]. The covariance of peptidoglycan-interacting residues in  
409 these diverging genera further underscores the importance of this domain in peptidoglycan  
410 interactions, rather than just protein-protein interactions.

411 The SpmX gene has arisen recently enough to see the hallmarks of novel gene emergence  
412 and adaptation in a constrained bacterial clade. The gene either arose from a fusion event in the  
413 bacterial genome, or the original phage gene contained the transmembrane segments.  
414 Domestication of the muramidase appears concomitant with the occurrence of the fused SpmX

415 domains in the extant genetic record. Maintenance of the muramidase domain since the  
416 emergence of SpmX and its activity in current living members of Caulobacterales suggest that its  
417 activity was selected for in the ancestral protein and still implicated in its modern function.  
418 Given the variability of the intermediate domain throughout Caulobacterales (**Fig. 1A**), this  
419 domain is under few constraints and appears to have undergone multiple independent events of  
420 elaboration and reduction in this order. This region of charged residues and prolines drives  
421 SpmX self-oligomerization *in vitro* [18], and may also facilitate other protein interactions. For  
422 example, the intermediate domain is responsible for the targeting of SpmX to sub-polar and  
423 bilateral positions in *Asticcacaulis* [17].

424 In several reported cases bacteria have domesticated phage genes for genetic  
425 manipulation and transfer, bacterial warfare, virulence, and secretion. However, these events are  
426 distinct from that which created the novel bacterial gene SpmX. Phage genes for DNA  
427 replication and recombination have replaced bacterial functional homologues within bacterial  
428 genomes several times [38–41]. These genes retain their original function and are now used by  
429 bacteria to carry out the same tasks. Gene transfer agents (GTAs) pose an interesting case where  
430 structural proteins from domesticated cryptic prophage are used by bacteria to package random  
431 DNA from the bacterial genome to presumably share with other bacteria [42]. In  
432 Alphaproteobacteria, a specific GTA has been stably maintained across several bacterial orders  
433 but the beneficial function is not precisely known [43,44]. Regardless, this domesticated island  
434 of phage genes still shuttles DNA around, as it once did in ancestral infectious cycles. Phage tails  
435 appear to have been weaponized many times, with multiple domestication events resulting in  
436 type VI secretion systems [45,46], tailocins and phage tail-like bacteriocins [12–14], phage tail-  
437 like systems with insecticidal properties [11,47,48], and phage tail-like arrays [49]. All of these  
438 represent a “guns for hire” acquisition scheme in which phage genes are co-opted without loss of  
439 ancestral toxicity and function [2]. Many of these genes reside in genomic islands and confer  
440 environmental, niche-specific advantages that directly exploit their ancestral activity for the  
441 benefit of the host. Similarly, in two other known cases of phage lysozyme domestication in  
442 bacteria, structurally similar, active domains have been fused to colicins [50] or are predicted to  
443 be secreted with type III secretion systems [51], presumably for use in bacterial warfare or  
444 infection. In one strange case, a phage lysozyme gene has been co-opted in bivalve genomes,  
445 which apparently still use the gene for its antibacterial properties [52].

446 The domestication of the muramidase domain in SpmX is distinct from the above cases  
447 of “guns for hire” because the phage gene has been incorporated into a novel bacterial gene with  
448 new function in basic cellular processes. The SpmX muramidase domain, although active, is no  
449 longer used to lyse bacterial cells. Instead, the active domain plays a role in localizing the SpmX  
450 protein for its function in developmental regulation and morphogenesis and has become part of a  
451 core gene in Caulobacterales. The co-option of phage genes for housekeeping function is likely a  
452 common event in nature, but identifying such genes may require a careful search. We suggest  
453 that the characteristics of SpmX may recommend future strategies for their detection: searching  
454 for phage gene homologues that have long histories of vertical inheritance and signs of bacterial  
455 innovation.

#### 456 **Acknowledgements**

457 We thank members of the Brun, Vernet, and Dessen laboratories for support, advice and  
458 encouragement. Many thanks to members of the Brun Laboratory, Ernesto Vargas, Farrah  
459 Bashey-Visser, Jay Lennon, Daniel Schwartz, and Breah LaSarre for critical reading and editing  
460

461 of the manuscript. Additional thanks to critical preprint review by the journal clubs of the  
462 (Pamela) Brown Lab at the University of Missouri and the Suel Lab at UCSD. We thank David  
463 Flot (ESRF, beamline ID30a1) for support in data collection. Support for this work came from  
464 NIH Grant 2R01GM051986 and R35GM122556 (to Y.V.B.), and NIH National Research  
465 Service Award F32GM112362 (to A.M.R.). Work of Y.V.B. at the Institut de Biologie  
466 Structurale in Grenoble was supported by a Fulbright US Research Scholar Award. This work  
467 used the platforms of the Grenoble Instruct-ERIC Center (ISBG : UMS 3518 CNRS-CEA-UGA-  
468 EMBL) with support from FRISBI (ANR-10-INSB-05-02) and GRAL (ANR-10-LABX-49-01)  
469 within the Grenoble Partnership for Structural Biology (PSB).

470

## 471 **Figure Legends**

472

473 **Figure 1. The SpmX is vertically inherited in Caulobacterales.** (A) Schematic of SpmX  
474 architecture, including the conserved muramidase domain (see **Fig. S1** for alignments), the  
475 variable intermediate domain, and two C-terminal transmembrane (TM) segments. Bar indicates  
476 amino acid sequence conservation among *spmX* alleles. (B) Phylogenetic trees of representative  
477 species from Caulobacterales and other Alphaproteobacteria for concatenated housekeeping gene  
478 alignments (left) and for SpmX (right), with branch colors indicating the amino acid identity at  
479 position 20 of SpmX (D20L in yellow, D20R in red, and D20G in green). The concatenated  
480 housekeeping tree is fully supported with posterior probability of 1.0 for all clades. Asterisks  
481 indicate clades in the SpmX tree with posterior probabilities > 0.95.

482

483 **Figure 2. The SpmX muramidase domain retains the canonical GH motif but contains**  
484 **inactivating mutations in the catalytic cleft.** (A) P22 lysozyme (PDB 2ANX) as a model  
485 lysozyme colored with rainbow gradient from blue N-terminus to red C-terminus. The catalytic  
486 glutamate appears in fuchsia and the GH beta-hairpin in light blue. (B) HMM logos of GH  
487 lysozymes made using WebLogo 3 [53]. Logos were constructed from protein sequences of (i)  
488 T4 lysozyme-like genes (n = 94), (ii) representative autolysins/endolysins from the Conserved  
489 Domain Database including P22 lysozyme (n = 20) but excluding SpmX genes, (iii) closest  
490 BLAST hits from non-SpmX muramidases (n = 60), and (iv) SpmX muramidases (n = 66), and  
491 organized in a cladogram to resemble the sequence cluster tree diagram in **Figure S2**. Amino  
492 acids are color-coded according to chemical properties, with uncharged polar residues in green,  
493 neutral residues in purple, basic residues in blue, acidic residues in red, and hydrophobic residues  
494 in black. The height of each letter is proportional to the relative frequency of a given identity and  
495 the height of the stack indicates the sequence conservation at that position. T4L numbering is  
496 used for ease of comparison. Asterisks mark positions critical for enzymatic activity and open  
497 circles mark positions associated with GH motif stability [26,27].

498

499 **Figure 3. The structure of SpmX muramidase domain has a wider, more dynamic catalytic**  
500 **cleft than related phage lysins.** (A) Structural alignment of P22 lysozyme (PDB 2ANX, the  
501 model used for molecular replacement) in purple, R21 endolysin from P21 (PDB 2HDE, a  
502 distantly related GH24 T4L lysozyme) in navy blue, and SpmX-Mur-*Ae* in gold (PDB 6H9D).  
503 The catalytic glutamate is shown in red. (B) Structural alignment of the three SpmX-Mur-*Ae*  
504 molecules, chains A (green), B (light blue), and C (dark blue), from the asymmetric unit. The  
505 surface of chain B is shown in partially transparent light blue. The double-headed arrow  
506 indicates the tilt of about 16° between the GH beta-hairpins of chains B and A. (C) Overlays of

507 ribbon diagrams and surfaces of P22 lysozyme (2ANX, left) and SpmX-Mur-*Ae* (6H9D, right)  
508 illustrating the conformation of the critical residues E11 (red), D20 (dark blue), R14 (yellow),  
509 and Y18 (orange). T4L numbering is used for ease of comparison. These structures have been  
510 rotated 180° around the y axis from their representation in (A, B, D, E). **(D)** Surface  
511 representation of P22 lysozyme (2ANX) with inset showing ribbon diagram and conformation of  
512 catalytic cleft with the canonical E11/D20/T26 catalytic triad. **(E)** Surface representation of  
513 SpmX-Mur-*Ae* (6H9D) with inset showing ribbon diagram and conformation of remodeled  
514 catalytic cleft with E11/D20L/T26M.

515  
516 **Figure 4. The D20L mutation attenuates P22 hydrolytic activity.** **(A)** Remazol brilliant blue  
517 assays on *C. crescentus* sacculi using purified P22 lysozyme, P22 lysozyme D20L mutant, and  
518 *C. crescentus* SpmX muramidase. Active enzymes release peptidoglycan monomers covalently-  
519 bound to RBB into the supernatant that are detected by absorbance at 595 nm. Error bars are ±  
520 standard deviation. Lines are drawn to help guide the eye toward basic trends. Data points are  
521 from various days and sacculi preparations, but with internal normalization to Hen Egg White  
522 Lysozyme (HEWL). **(B and C)** Growth curves of Lemo21(DE3) *E. coli* expressing P22  
523 lysozyme (blue), P22 lysozyme D20L mutant (green), and *C. crescentus* SpmX muramidase  
524 (red). Proteins were expressed from pET22b with a N-terminal PelB signal sequence. In **(B)**,  
525 strains were grown in 5 mM rhamnose without IPTG for maximal repression of basal expression  
526 from the plasmids. In **(C)**, strains were grown without rhamnose and induced with 400 μM IPTG  
527 at the indicated time. **(D)** Phase/fluorescent overlays show live/dead staining of (i)  
528 Lemo21(DE3) cells expressing P22Lyso-D20L and (ii) SpmX-Mur-*Cc* after four hours of  
529 induction. Green, membrane permeable SYTO 9 stains DNA in live cells and red, membrane  
530 impermeable propidium iodide nucleic acid dyes labels released nucleoids and DNA from lysed  
531 bacteria. The rounding of the *E. coli* in (i) is characteristic of spheroplast formation and lysis by  
532 hydrolytic activity on the cell wall. Scale bars are 5 μm.

533  
534 **Figure 5. Inactivating the muramidase domain partially delocalizes SpmX in vivo.**  
535 Phase and fluorescent images of **(A)** *C. crescentus*, **(B)** *A. biprosthicum*, and **(C)** *A. excentricus*.  
536 In the top panel, phase images with derived schematics emphasizing stalks and morphologies are  
537 shown for **(i)** WT and **(ii)** Δ*spmX* cells. In **Aii**, *C. crescentus* cells exhibiting characteristic  
538 Δ*spmX* divisional defects are marked with asterisks and a cell growing stalks from both poles has  
539 its stalks marked with red arrowheads. Phase and fluorescent images of cells expressing **(iii)**  
540 SpmX-eGFP, **(iv)** SpmX-E11A-eGFP, or **(v)** SpmX-N105R-eGFP from the native chromosomal  
541 locus are shown in the lower panels. In **Aiv** and **Av**, cells with divisional defects are marked with  
542 white asterisks. In **Biii** and **Biv**, cells with one lateral or subpolar stalk are marked with white  
543 arrowheads. In **Civ**, cells with foci at the tips of stalks are marked with white arrowheads. All  
544 scale bars are 5 μm.

545  
546 **Figure 6. Replacing the muramidase domain interferes with native SpmX protein levels in vivo**  
547 **(A)** Phase and fluorescent images of strains in which the native *spmX* allele was replaced  
548 with the following gene fusions in the Δ*spmX* parent strain **(i)**: **(ii)** WT *spmX-sfGFP*, **(iii)** *spmX-*  
549 *E11A-sfGFP*, **(iv)** *spmX-L20D-sfGFP*, **(v)** *MurBs-Δmur-SpmX-sfGFP* where *MurBs* is the  
550 muramidase domain from *Brevundimonas subvibrioides* SpmX, and **(vi)** *spmX-L20R-sfGFP*. All  
551 scale bars are 5 μm. **(B)** Western blot comparing the Δ*spmX* parent strain to SpmX mutants and

552 chimeras inserted at the *spmX* locus. In all cases, the primary antibody is directed against the C-  
553 terminal GFP fusion.

554

## 555 Tables

556

### 557 Table S1. *spmX* genes

Gene ID	Genome ID	Genome Name	Aas
2579190773	2576861677	<i>Asticcacaulis benevestitus</i> DSM 16100, ATCC BAA-896	825
2515453797	2515154077	<i>Asticcacaulis benevestitus</i> DSM 16100, ATCC BAA-896	825
651318754	651285000	<i>Asticcacaulis biprotheticum</i> C19, ATCC 27554	826
649817473	649633007	<i>Asticcacaulis excentricus</i> CB 48	810
2579161412	2576861669	<i>Asticcacaulis</i> sp. AC402	861
2575314089	2574180013	<i>Asticcacaulis</i> sp. AC460	803
2573897243	2571042840	<i>Asticcacaulis</i> sp. AC466	863
2579118533	2576861656	<i>Asticcacaulis</i> sp. YBE204	851
2597242146	2596583650	<i>Asticcacaulis taihuensis</i> CGMCC 1.3431	775
2525860150	2524614798	<i>Brevundimonas aveniformis</i> DSM 17977	453
2562052386	2561511096	<i>Brevundimonas bacteroides</i> DSM 4726	507
2534536731	2531839707	<i>Brevundimonas diminuta</i> 470-4	291
2635997287	2634166514	<i>Brevundimonas diminuta</i> XGC1	291
2580800557	2579778742	<i>Brevundimonas naejangsanensis</i> B1	291
2525302517	2524614659	<i>Brevundimonas naejangsanensis</i> DSM 23858	296
2602030628	2600255394	<i>Brevundimonas nasdae</i> TPW30	393
2597236398	2596583649	<i>Brevundimonas</i> sp. 374	381
2646174451	2645727605	<i>Brevundimonas</i> sp. AAP58	493
647574062	647533115	<i>Brevundimonas</i> sp. BAL3	526
2628481828	2627853674	<i>Brevundimonas</i> sp. KM4	381
648120868	648028010	<i>Brevundimonas subvibrioides</i> ATCC 15264	514
2671799930	2671180176	<i>Brevundimonas viscosa</i> CGMCC 1.10683	420
637088217	637000061	<i>Caulobacter crescentus</i> CB15	431
643620503	643348526	<i>Caulobacter crescentus</i> NA1000	431
2522234411	2522125057	<i>Caulobacter crescentus</i> OR37	429
2585155721	2582581280	<i>Caulobacter henricii</i> CF287	419
2585146921	2582581279	<i>Caulobacter henricii</i> OK261	398
2585198327	2582581293	<i>Caulobacter henricii</i> YR570	419
646754844	646564519	<i>Caulobacter segnis</i> ATCC 21756	413
2511123337	2510917020	<i>Caulobacter</i> sp. AP07	425
2528897460	2528768059	<i>Caulobacter</i> sp. JGI 0001010-J14	422
2528898539	2528768060	<i>Caulobacter</i> sp. JGI 0001013-D04	421
641563661	641522612	<i>Caulobacter</i> sp. K31	417
2587919405	2585428106	<i>Caulobacter</i> sp. OV484	439
2686817300	2684623106	<i>Caulobacter</i> sp. UKL13	439
2593324358	2590828858	<i>Caulobacter</i> sp. UNC279MFTsu5.1	421

2565991034	2565956508	<i>Caulobacter sp. UNC358MFTsu5.1</i>	424
2522923390	2522572103	<i>Caulobacter sp. URHA0033</i>	413
2677566563	2675903340	<i>Caulobacter vibrioides T5M6</i>	426
2525553380	2524614726	<i>Hellea balneolensis DSM 19091</i>	439
2519259833	2519103046	<i>Henriciella marina DSM 19595</i>	535
644903978	644736375	<i>Hirschia baltica ATCC 49814</i>	597
2520004565	2519899518	<i>Hirschia maritima DSM 19733</i>	577
2578549711	2576861496	<i>Hyphomonas adhaerens MHS-3</i>	386
2578539236	2576861493	<i>Hyphomonas hirschiana VP5</i>	421
2578553211	2576861497	<i>Hyphomonas jannaschiana VP2</i>	385
2578549047	2576861495	<i>Hyphomonas johnsonii MHS-2</i>	409
638143841	637000135	<i>Hyphomonas neptunium ATCC 15444</i>	421
2578534396	2576861492	<i>Hyphomonas oceanitis SCH89</i>	407
2578543740	2576861494	<i>Hyphomonas polymorpha PS728</i>	418
2583291548	2582580773	<i>Hyphomonas sp. 22III-22F38</i>	390
2583281302	2582580770	<i>Hyphomonas sp. 25B14_1</i>	392
2583297317	2582580774	<i>Hyphomonas sp. BH-BN04-4</i>	407
2583284155	2582580771	<i>Hyphomonas sp. CY54-11-8</i>	386
2583256535	2582580764	<i>Hyphomonas sp. L-53-1-40</i>	398
2583287429	2582580772	<i>Hyphomonas sp. T16B2</i>	392
638141228	637000157	<i>Maricaulis maris MCS10</i>	638
2635165820	2634166316	<i>Maricaulis salignorans DSM 16077</i>	633
2554749235	2554235119	<i>Maricaulis sp. JL2009/Euryhalocaulis caribicus</i>	464
2523851441	2523533582	<i>Oceanicaulis alexandrii DSM 11625</i>	643
638892666	638341138	<i>Oceanicaulis alexandrii HTCC2633</i>	596
2609050000	2608642188	<i>Oceanicaulis sp.</i>	514
2582569245	2579779170	<i>Oceanicaulis sp. HL-87</i>	514
648155025	648028050	<i>Parvularcula bermudensis HTCC2503</i>	344
2574131646	2571042909	<i>Phenylobacterium composti DSM 19425</i>	438
642759179	642555147	<i>Phenylobacterium zucineum HLK1</i>	445
2524841340	2524614539	<i>Ponticaulis koreensis DSM 19734</i>	687
2522300904	2522125083	<i>Robiginitomaculum antarcticum DSM 21748</i>	502
2522254887	2522125070	<i>Woodsholea maritima DSM 17123</i>	853

558

559 **Table S2. Data collection and refinement statistics**

560	<u>Data collection</u>	
561	Name of dataset	3SpmX_B1_1
562	X-ray source	ID30a-1 (ESRF)
563	Scan range (°)	76
564	Oscillation (°)	0.1
565	Space group	P3 <sub>2</sub> 21
566	Unit-cell parameters	
567	a, Å	100.44
568	b, Å	100.44
569	c, Å	96.62

570	$\alpha$ , °	90
571	$\beta$ , °	90
572	$\gamma$ , °	120
573	Resolution (last shell), Å	1.90 (2.01-1.90)
574	Completeness (last shell), %	95.1 (95.0)
575	$\ I(I)\ $ (last shell)	33.99 (3.1)
576	Rsym† (last shell), %	2.0 (43.6)
577	No of unique reflections	42611
578	Wilson B factor, (Å <sup>2</sup> )	44.4
579		
580	<b>Refinement and model statistics</b>	
581	Resolution (last shell), Å	1.90 (1.95-1.90)
582	R-factor‡, R-free§	0.211, 0.256
583	Molecules/asymmetric unit	3
584	rmsd from target <sup>†</sup>	
585	Bond lengths, Å	0.008
586	Bond angle, °	1.182
587	Average B-factor, Å <sup>2</sup>	86.34
588	Mean B factor (Å <sup>2</sup> )	41.7
589	Ramachandran plot**	
590	Core, %	97.5
591	Allowed, %	2.5
592	Disallowed, %	0
593		

594 †Rsym =  $(\sum(\text{ABS}(I(h,i)-I(h)))) / (\sum(I(h,i)))$ .

595 ‡R-factor =  $\sum |F_o - F_c| / \sum |F_o|$  where  $F_o$  and  $F_c$  are the observed and calculated structure factor amplitudes, respectively.

596 §R-free is the R-factor calculated with 5% of the reflections chosen at random and omitted from refinement.

597 <sup>†</sup>rmsd of bond lengths and bond angles from ideal geometry.

600 \*\*Performed by Procheck.

601  
602

### Table S3. Plasmids

#### YB# Plasmid Name/Contents (Resistance)

#### Source or Reference

Expression Plasmids – Structural Studies

pTB146-SUMO Bernhardt Lab, [54]  
**8120** pTB146-SUMO-*spmX-Mur(1-150)-Ae* (Amp) This study

Expression Plasmids – Activity and Binding

**1170** pET28a (Kan) Novagen  
**8107** pET28a-*spmX-Mur(1-150)-Cc* (Kan) This study  
**8111** pET28a-*spmX-Mur(1-150)-Ae* (Kan) This study  
**8114** pET28a-*spmX-Mur(1-150)-Ab* (Kan) This study  
**8889** pET28a-*spmX-Mur-Cc-E19A* (Kan) “E11A” This study  
**8890** pET28a-*spmX-Mur-Cc-L28D* (Kan) “L20D” This study  
**8891** pET28a-*spmX-Mur-Cc-N91R* (Kan) “N105R” This study  
**8898** pET28a-*P22Lyso* (Kan) This study  
**410** pET28a-*P22Lyso-E11A* (Kan) This study  
**7378** pET28a-*P22Lyso-D20L* (Kan) This study

Periplasmic Expression Plasmids

**3562** pET22b (Amp) Novagen  
**7379** pET22b-*spmX-Mur(1-150)-Cc* (Amp) This study  
**7380** pET22b-*spmX-Mur-Cc-E19A* (Amp) “E11A” This study

<b>7381</b>	pET22b- <i>spmX-Mur-Cc-L28D</i> (Amp) “L20D”	This study
<b>7382</b>	pET22b- <i>P22Lyso</i> (Amp)	This study
<b>7383</b>	pET22b- <i>P22Lyso-E11A</i> (Amp)	This study
<b>7384</b>	pET22b- <i>P22Lyso-D20L</i> (Amp)	This study
<i>In vivo</i> Allelic Exchange and Insertional Plasmids		
<b>1621</b>	pNPTS138 (Kan)	M.R.K. Alley, unpublished
<b>8247</b>	pNPTS138- $\Delta$ <i>spmXAe</i>	This study
<b>8246</b>	pNPTS138- $\Delta$ <i>spmXAb</i>	This study
<b>8252</b>	pNPTS138- <i>spmXCc-E19A</i> “E11A”	This study
<b>8251</b>	pNPTS138- <i>spmXAe-E19A</i> “E11A”	This study
<b>8250</b>	pNPTS138- <i>spmXAb-E19A</i> “E11A”	This study
<b>7395</b>	pNPTS138- $\Delta$ <i>mur(1-150)-spmXCc</i>	This study
<b>7396</b>	pNPTS138- $\Delta$ <i>mur(1-150)-spmXAe</i>	This study
<b>7398</b>	pNPTS138- $\Delta$ <i>mur(1-150)-spmXAb</i>	This study
<b>5111</b>	pGFPC-1 (Spc/Str)	[55]
	pGFPC-2 (Kan)	[55]
<b>7407</b>	pGFPC-1- <i>spmXAb</i> (last 600 bp) (Spc/Str)	[17]
<b>7408</b>	pGFPC-1- <i>spmXAe</i> (last 600 bp) (Spc/Str)	[17]
<b>8271</b>	pGFPC-2- <i>spmXCc</i> (last 600 bp) (Kan)	This study
	pMCS-2 (Kan)	[55]
<b>7409</b>	pMCS-2- <i>spmXCc-sfGFP</i> (Kan)	This study
<b>7410</b>	pMCS-2- <i>spmXCc-E19A-sfGFP</i> (Kan) “E11A”	This study
<b>7417</b>	pMCS-2- <i>P22Lyso-<math>\Delta</math>mur-spmXCc-sfGFP</i> (Kan)	This study
<b>7418</b>	pMCS-2- <i>P22Lyso-E11A-<math>\Delta</math>mur-spmXCc-sfGFP</i> (Kan)	This study
<b>1041</b>	pMCS-2- <i>MBs(1-140)-<math>\Delta</math>mur-spmXCc-sfGFP</i> (Kan)	This study
<b>1042</b>	pMCS-2- <i>MPb(1-168)-<math>\Delta</math>mur-spmXCc-sfGFP</i> (Kan)	This study

603

604

**Table S4. Strains**

<b>YB#</b>	<b>Strain Genotype (Resistance)</b>	<b>Construct, Source, or Reference</b>
<i>E. coli</i> Strains		
	Alpha-Select	Bioline
	BL21(DE3) RIL	Agilent
	BL21(DE3) RIL pTB146SUMO- <i>spmXMur(1-150)-Ae</i> (Amp)	This study
<b>104</b>	BL21(DE3)	Novagen
<b>8126</b>	BL21(DE3) pET28a- <i>spmX-Mur(1-150)-Cc</i>	This study
<b>8129</b>	BL21(DE3) pET28a- <i>spmX-Mur(1-150)-Ae</i>	This study
<b>8132</b>	BL21(DE3) pET28a- <i>spmX-Mur(1-150)-Ab</i>	This study
<b>8881</b>	BL21(DE3) pET28a- <i>spmX-Mur-Cc-E19A</i>	This study
<b>8882</b>	BL21(DE3) pET28a- <i>spmX-Mur-Cc-L28D</i>	This study
<b>8899</b>	BL21(DE3) pET28a- <i>spmX-Mur-Cc-N91R</i>	This study
<b>8141</b>	BL21(DE3) pET28a- <i>P22L</i>	This study
<b>1043</b>	BL21(DE3) pET28a- <i>P22L-E11A</i>	This study
<b>1044</b>	BL21(DE3) pET28a- <i>P22L-D20L</i>	This study
<b>8880</b>	Lemo21(DE3) (Cm)	NEB



<b>1045</b>	Lemo21(DE3) pET22b- <i>spmX-Mur(1-150)-Cc</i> (Cm/Amp)	This study
<b>1046</b>	Lemo21(DE3) pET22b- <i>spmX-Mur-Cc-E19A</i> (Cm/Amp)	This study
<b>1047</b>	Lemo21(DE3) pET22b- <i>spmX-Mur-Cc-L20D</i> (Cm/Amp)	This study
	Lemo21(DE3) pET22b- <i>P22L</i> (Cm/Amp)	This study
<b>1049</b>	Lemo21(DE3) pET22b- <i>P22L-E11A</i> (Cm/Amp)	This study
<b>1050</b>	Lemo21(DE3) pET22b- <i>P22L-D20L</i> (Cm/Amp)	This study
<b><i>C. crescentus</i> Strains</b>		
<b>127</b>	<i>C. crescentus</i> CB15N (holdfast minus WT)	
<b>5873</b>	CB15N $\Delta$ <i>spmX</i>	[15]
<b>8265</b>	CB15N <i>spmX :: spmXCc-eGFP</i> (Kan)	YB127 electroporated with YB8271
<b>8262</b>	CB15N <i>spmXCc-E19A</i> “E11A”	YB127 electroporated with YB8252
<b>8266</b>	CB15N <i>spmX-E19A :: spmXCc-eGFP</i> (Kan) “E11A”	YB8262 electroporated with YB8271
<b>8275</b>	CB15N $\Delta$ <i>mur(1-150)-spmXCc</i>	YB127 electroporated with YB7395
<b>1051</b>	CB15N $\Delta$ <i>mur-spmXCc :: spmXCc-eGFP</i> (Kan)	YB8275 electroporated with YB8271
<b>1052</b>	CB15N $\Delta$ <i>spmX :: spmXCc-sfGFP</i> (Kan)	YB5873 electroporated with YB7409
<b>1053</b>	CB15N $\Delta$ <i>spmX :: spmXCc-E19A-sfGFP</i> (Kan) “E11A”	YB5873 electroporated with YB7410
<b>1054</b>	CB15N $\Delta$ <i>spmX :: P22Lyso-<math>\Delta</math>mur-spmXCc-sfGFP</i> (Kan)	YB5873 electroporated with YB7417
<b>1055</b>	CB15N $\Delta$ <i>spmX :: P22Lyso-E11A-<math>\Delta</math>mur-spmXCc-sfGFP</i> (Kan)	YB5873 electroporated with YB7418
<b>1056</b>	CB15N $\Delta$ <i>spmX :: MBs-<math>\Delta</math>mur-spmXCc-sfGFP</i> (Kan)	YB5873 electroporated with YB1041
<b>1057</b>	CB15N $\Delta$ <i>spmX :: MPb-<math>\Delta</math>mur-spmXCc-sfGFP</i> (Kan)	YB5873 electroporated with YB1042
<b><i>A. excentricus</i> Strains</b>		
<b>258</b>	<i>A. excentricus</i> AC48	[56]
<b>6662</b>	AC48 <i>spmX :: spmXAe-eGFP</i> (Spc/Str)	[17]
<b>8238</b>	AC48 $\Delta$ <i>spmX</i>	YB258 electroporated with YB8247
<b>8260</b>	AC48 <i>spmXAe-E19A</i> “E11A”	YB258 electroporated with YB8251
<b>8261</b>	AC48 <i>spmXAe-E19A :: spmXAe-eGFP</i> (Spc/Str) “E11A”	YB8260 electroporated with YB5686
<b>8276</b>	AC48 $\Delta$ <i>mur(1-150)spmXAe</i>	YB258 electroporated with YB7396
<b>1058</b>	AC48 $\Delta$ <i>mur-spmXAe :: spmXAe-eGFP</i> (Spc/Str)	YB8276 electroporated with

		YB7408
<b>A. biprosthhecum Strains</b>		
642	<i>A. biprosthhecum</i> C19	[57]
1059	C19 <i>spmX</i> :: <i>spmXAb-eGFP</i> (Spc/Str)	[17]
8237	C19 $\Delta$ <i>spmX</i>	YB642 electroporated with YB8246
8257	C19 <i>spmXE19A</i> “E11A”	YB642 electroporated with YB8250
8258	C19 <i>spmXE19A</i> :: <i>spmXAb-eGFP</i> (Spc/Str) “E11A”	YB8257 electroporated with YB5681
8277	C19 $\Delta$ <i>mur(1-150)</i> - <i>spmXAb</i>	YB642 electroporated with YB7398
1060	C19 $\Delta$ <i>mur(1-150)</i> - <i>spmXAb</i> :: <i>spmXAb-eGFP</i> (Spc/Str)	YB8277 electroporated with YB7407

605  
606

## 607 **Materials and methods**

608

609 **Bacterial strains and growth conditions.** All *C. crescentus*, *A. excentricus*, and *A.*  
610 *biprosthhecum* strains used in this study were grown in liquid PYE medium. *C. crescentus* was  
611 grown at 30°C and *Asticcacaulis* species at 26°C. Strains were maintained on PYE plates  
612 supplemented with antibiotics as necessary (kanamycin 20  $\mu\text{g mL}^{-1}$ , gentamycin 5  $\mu\text{g mL}^{-1}$ , and  
613 spectinomycin 100  $\mu\text{g mL}^{-1}$ ). For microscopy, *C. crescentus* and *A. excentricus* were inoculated  
614 from colonies, grown overnight, then diluted back 1:50 and grown for another 3-4 hours before  
615 being imaged in mid- to late-exponential phase. *A. biprosthhecum* was inoculated from colonies  
616 and grown overnight to reach mid- to late-exponential phase for imaging. A detailed list of  
617 strains is included as **Table S4**. *E. coli* strains were grown as described in the Methods sections  
618 on purification and periplasmic expression.

619

620 **Bioinformatics and gene trees.** Sequences of the *SpmX* genes in **Table S1** and members of the  
621 GH24 family were retrieved by BLAST searches on the Integrated Microbial Genomes and  
622 Microbiomes (IMG/M) database [58] and the National Center for Biotechnology Information  
623 (NCBI) “nr” database. Multiple alignments were achieved with MUSCLE [59] and manually  
624 adjusted and visualized with Jalview [60]. Sequence conservation of *SpmX* residues was  
625 determined from the multiple sequence alignment of *spmX* alleles using ConSeq [61]. In order to  
626 improve visualization of conservation patterns, the ConSeq scores were averaged across a 20-  
627 residue sliding window. For estimating bacterial species phylogeny, assembled genome data  
628 were obtained from the genome database of the National Center for Biotechnology Information  
629 [62]. Amino acid sequences of 37 conserved housekeeping genes were automatically identified,  
630 aligned, and concatenated using Phylosift [63]. All phylogenetic reconstruction was performed  
631 using MrBayes v3.2.6 [64] to estimate consensus phylogenies and clade posterior probability  
632 support values. Sequence substitution was modeled according to a WAG substitution model with  
633 gamma-distributed rate variation between sites. Trees were visualized and formatted using iTol  
634 [65]. The sequence cluster tree was built with NCBI’s Conserved Domain Database tool (CDD).  
635 This tool uses reverse position-specific BLAST, a method that compares query sequences to  
636 databases of position-specific score matrices and obtains *E*-values, such as in PSI-BLAST

637 [21,22]. WebLogo3 was used to plot the amino acid distribution at each position of the GH motif  
638 [53]. To create the alignments for logo generation, 94 T4 lysozyme-like sequences, 20  
639 endolysin/autolysins from the CDD analysis, 60 SpmX muramidase-like sequences (BLAST  
640 hits), and 66 SpmX muramidase sequences were simultaneously aligned to T4 lysozyme. Only  
641 sequences with unambiguous alignment in the GH motif were included in this analysis.  
642

643 **Recombinant DNA methods.** DNA amplification, Gibson cloning, and restriction digests were  
644 performed according to the manufacturer. Restriction enzymes and Gibson cloning mix were  
645 from New England Biolabs. Cloning steps were carried out in *E. coli* (alpha-select competent  
646 cells, Bioline) and plasmids were purified using Zyppy Plasmid Kits (Zymo Research  
647 Corporation). Sequencing was performed by the Indiana Molecular Biology Institute and  
648 Eurofins MWG Operon Technologies with double stranded plasmid or PCR templates, which  
649 were purified with a DNA Clean & Concentrator kits (Zymo Research Corporation).  
650 Chromosomal DNA was purified using the Bactozol Bacterial DNA Isolation Kit (Molecular  
651 Research Center). Plasmids were introduced into all *E. coli* strains using chemical transformation  
652 according to the manufacturer's protocols. Plasmids were introduced into *C. crescentus*, *A.*  
653 *excentricus*, and *A. biprosthicum* by electroporation based on previously published studies [66].  
654 Allelic exchange in was achieved with pNPTS138, large genetic insertions with pMCS-2 [55],  
655 and eGFP insertional fusions with pGFPC-1 and pGFPC-2 [55].  
656

#### 657 **Plasmid construction.**

658 *Expression plasmids* (pTB146SUMO, pET28a, pET22b): *spmX* gene fragments encoding amino  
659 acids 2-150 of SpmX (SpmX-Mur) were amplified from genomic DNA and inserted into  
660 linearized expression vectors using Gibson cloning (NEB) according to manufacturers protocols.  
661 P22 lysozyme (P22Lyso) was amplified from a synthetic gene strand (Eurofins) for similar  
662 construction with Gibson cloning. For pTB147SUMO, the vector was linearized with SapI and  
663 XhoI to insert SpmX-Mur-*Ae*. For pET28a, the vector was linearized with NdeI and EcoRI to  
664 insert SpmX-Mur-*Cc*, BamHI and XhoI to insert SpmX-Mur-*Ae*, SacI to insert SpmX-Mur-*Ab*,  
665 and EcoRI to insert P22Lyso. In all pET28a plasmids, the constructs were cloned in frame with  
666 the N-terminal His-tag and a stop codon to eliminate the C-terminal His-tag. For pET22b, the  
667 vector was linearized with EcoRI and the C-terminal His-tag was preserved. Point mutants in  
668 expression vectors were obtained by using standard quick-change procedures and primers with 3'  
669 single stranded overhangs for increased efficiency.

670 *Integrating plasmids for allelic exchange* (pNPTS138): For allelic exchange, the desired  
671 mutation was engineered into pNPTS138, bracketed by 1 kb up- and downstream of the  
672 corresponding genetic region. Integrants were isolated by antibiotic selection and secondary  
673 recombination events were selected by sucrose counter-selection using standard procedures. The  
674 resulting clones were confirmed by PCR and sequencing isolated genomic DNA.

675 For genomic deletions of *spmX* in *Asticcacaulis*, pNPTS138 was linearized with EcoRI  
676 and codons on either end of the gene were retained to avoid introducing frame-shifts in the  
677 surrounding area. Therefore the final gene deletion in *A. excentricus* lacks residues 5-808 and in  
678 *A. biprosthicum* lacks residues 5-815. For SpmX $\Delta$ mur truncations, residues 2-150 were removed  
679 in all three species. In all cases pNPTS138 was linearized with EcoRV, except for SpmX $\Delta$ Ab-  
680 E19A, where pNPTS138 was linearized with EcoRI. Point mutations E19A and N91R were  
681 integrated into the  $\Delta$ *spmX* background for ease of clone isolation and included full-length SpmX  
682 flanked by 1 kb genetic context. pNPTS138 containing mutated SpmX were constructed using

683 Gibson cloning with fragments on either side of the intended mutation and overlapping primers  
684 containing the mutation amplified from genomic DNA. In all cases pNPTS138 was linearized  
685 with EcoRV, except for *SpmXAb-E19A*, where pNPTS138 was linearized with EcoRI.

686 *Integrating plasmids for insertional eGFP fusions* (pGFPC): The last 600 bp of *spmX*  
687 from *C. crescentus* was amplified from genomic DNA and cloned into pGFPC-2 using Gibson  
688 cloning.

689 *Integrating plasmids for replacement at the  $\Delta$ spmX locus* (pMCS-2). These constructs  
690 were designed to allow insertion of various SpmX mutants fused to C-terminal sfGFP into the  
691  $\Delta$ spmX locus in *C. crescentus*. For SpmX-sfGFP and SpmX-E19A-sfGFP, fragments containing  
692 1 kb of genomic DNA upstream of *spmX* and *spmX* or *spmX-E19A* were amplified from existing  
693 pNPTS138 constructs and fused to a fragment containing monomeric sfGFP amplified from  
694 pSRKKm-Plac-sfgfp [67] using Gibson cloning. In the final construct, SpmX and sfGFP are  
695 connected with the linker sequence GSAGSAAGSGEF [68]. Chimeras with P22 lysozyme  
696 (P22Lyso) and its catalytic mutant were made by Gibson cloning together fragments containing  
697 1kb of upstream genomic DNA, P22Lyso (with no stop codon), SpmX $\Delta$ mur (residues 151-431)  
698 and sfGFP with the same linker. P22Lyso and P22Lyso-E11A were amplified from pET28a  
699 plasmids containing these genes. Chimeras with SpmX muramidase from *Brevundimonas*  
700 *subvibrioides* (residues 1-140) and *Parvularcula bermudensis* (residues 1-168) were similarly  
701 made with the muramidase fragments amplified from genomic DNA and synthetic gene strands  
702 (Eurofins), respectively.

703  
704 **Production of SpmX-Mur-Ae for crystallographic studies.** The muramidase domain of SpmX  
705 from *A. excentricus* (SpmX-Mur-Ae, residues 2-150) was fused to a hexahistidine tag followed  
706 by the SUMO cleavage site of the Ulp1 protease (His-SUMO tag) (**Table S3**) [69] and  
707 overexpressed in *E. coli* BL21 (DE3) RIL cells. Cells were grown at 37 °C in 2 l of Terrific  
708 Broth (BD Biosciences) supplemented with ampicillin (100  $\mu$ g/ml) until the OD<sub>600nm</sub> reached 0.8.  
709 Production of the recombinant protein was induced by the addition of isopropyl  $\beta$ -D-1-  
710 thiogalactopyranoside (IPTG) to 0.5 mM after the culture was cooled to 25°C. Cell growth was  
711 continued overnight at 25°C, and cells were harvested by centrifugation. Cell pellets were  
712 resuspended in 1/20<sup>th</sup> volume of buffer A (50 mM Tris-HCl (pH 8.0), 500 mM NaCl, 25 mM  
713 imidazole, 10% (vol/vol) glycerol) containing the Complete<sup>TM</sup> cocktail of protease inhibitors  
714 (Roche). Cells were lysed by six passages through a cell disruptor (Constant Systems Limited) at  
715 20 kPsi, and cell debris were pelleted by centrifugation at 40,000  $\times$  g for 30 min at 4 °C. The  
716 centrifugation supernatant was loaded on a Ni-NTA agarose resin (Qiagen) equilibrated with  
717 buffer A. After extensive washing with buffer A, His-SUMO-SpmX-Mur-Ae was eluted with a  
718 linear 0-100% gradient of buffer B (50 mM Tris-HCl (pH 8.0), 300 mM NaCl, 500 mM  
719 imidazole, 10% (vol/vol) glycerol) over 10 column volumes. Peak fractions were pooled, mixed  
720 with a 1:100 dilution of a His-tagged Ulp1 (SUMO) protease preparation [70] and dialyzed  
721 overnight at 4°C in buffer C (50 mM Tris-HCl (pH 8.0), 300 mM NaCl, 10% (vol/vol) glycerol).  
722 Cleavage reactions were passed through Ni-NTA resin to remove free His-SUMO tag and His-  
723 Ulp1, and untagged protein was collected in the flow through. Flow-through fractions were  
724 concentrated with Amicon Ultra Centrifugal filter units with a molecular weight cutoff of 10 kDa  
725 (Millipore) and were injected onto an ENrich<sup>TM</sup> SEC650 10x300 gel-filtration column (Biorad).  
726 SpmX-Mur-Ae was eluted with buffer D (25 mM Tris-HCl (pH 8.0), 150 mM NaCl) and again  
727 concentrated with Amicon Ultra Centrifugal filter units. Protein concentration was measured  
728 using absorbance at 280 nm.

729  
730 **Protein crystallization and structure determination.** High-throughput crystallization trials  
731 were performed with a Cartesian PixSys 4200 crystallization robot (Genomic Solutions, U.K.).  
732 Hanging drops containing 100 nl of protein (25 or 12.5 mg/ml) and 100 nl of reservoir solution  
733 were set up in 96-well Crystal Quick plates (Greiner) and incubated at 20°C. Initial crystal hits  
734 were refined manually by setting up hanging drops containing 1 µl of protein (25 or 12.5 mg/ml)  
735 and 1 µl of reservoir solution in 24-well plates (Molecular Dimensions) incubated at 20°C.  
736 Large needle-shaped crystals (dimensions of about 40 x 40 x 400 µm) were finally obtained for  
737 SpmX-Mur-*Ae* in 0.1 M Tris-HCl pH 8.5, 12% PEG 3350, 0.2 M MgCl<sub>2</sub>, at 20°C within 24–48  
738 h. SpmX-Mur-*Ae* crystals were cryoprotected by transfer into 0.1 M Tris-HCl pH 8.5, 13% PEG  
739 3350, 0.2 M MgCl<sub>2</sub>, 10% glycerol, and then flash-frozen in liquid nitrogen. X-ray diffraction  
740 data were collected at the European Synchrotron Radiation Facility (ESRF, Grenoble, France) on  
741 the ID30a1 (MASSIF-1) beamline [71,72].

742 Diffraction data were indexed and scaled using the XDS program suite [73]. SpmX-Mur-  
743 *Ae* crystals belong to the trigonal space group P3<sub>2</sub>21, with unit cell dimensions of 100.44 x  
744 100.44 x 96.62 Å and three molecules per asymmetric unit. Phase determination was carried out  
745 by the molecular replacement method with PHASER [74], using as a search model the structure  
746 of the phage P22 lysozyme (PDB entry 2ANX). The molecular replacement solution model was  
747 rebuilt de novo using PHENIX [75] to prevent bias from the model.

748 The structure of SpmX-Mur-*Ae* was completed by cycles of manual building with COOT  
749 [76] and addition of water molecules with ARP/wARP [77]. Several cycles of manual building  
750 and refinement with REFMAC [78], as implemented in the CCP4 program suite, were performed  
751 until  $R_{work}$  and  $R_{free}$  converged [79]. Stereochemical verification was performed with  
752 PROCHECK [80]. The secondary structure assignment was verified with DSSP [81], with all  
753 residues within most favorable or allowed regions of the Ramachandran plot. Figures were  
754 generated with PyMol (<http://www.pymol.org>). Coordinates of the final refined model were  
755 deposited at the Protein Data Bank (PDB, <http://www.rcsb.org>) and were assigned PDB entry  
756 code 6H9D. The data collection and refinement statistics are summarized in **Table S2**.  
757

758 **Protein production for *in vitro* assays.** Fresh BL21(DE3) competent cells (Novagen) were  
759 transformed with pET28a constructs containing various muramidase genes with N-terminal His-  
760 tags (**Table S3**) and grown overnight in LB with 1% glucose and 50 µg/mL kanamycin.  
761 Overnight cultures were diluted 100-fold in LB medium with 1% glucose and 50 µg/mL  
762 kanamycin. Typically 500 mL cultures of cells were grown for 1.5-2 hours to an OD<sub>600</sub> of 0.6-  
763 0.7 and shifted to 20°C. When the OD<sub>600</sub> reached 0.8–0.9, the cells were induced with 0.5 mM  
764 IPTG. After growing for 4 h at 20 °C, cells were harvested and resuspended in 30 mL lysis  
765 buffer (25 mM HEPES pH 7.5, 100 mM NaCl, 20 mM imidazole, 5 mM BME) with a EDTA-  
766 free Protease Inhibitor Mini Tablet (Pierce) and phenylmethanesulfonyl fluoride (PMSF, 1 mM).  
767 The 30 mL cell mixture was lysed on ice using a sonicating horn and spun down at 10,000g for  
768 20-30 min. The clarified lysate was loaded onto a 5 ml HiTrap Chelating HP cartridge (GE  
769 Healthcare) charged with Ni<sup>2+</sup> and pre-equilibrated with lysis buffer. After loading, the column  
770 was washed with lysis buffer followed by an elution via a 0-100% linear gradient of buffer B (25  
771 mM HEPES pH 7.5, 100 mM NaCl, 500 mM imidazole, 2 mM BME). Muramidase-containing  
772 fractions were pooled based on SDS-PAGE analysis and concentrated to 2.5 mL. Imidazole was  
773 removed by passing the concentrated fraction over a PD10 desalting column (GE Healthcare)  
774 equilibrated with 25 mM HEPES pH 7.5, 100 mM NaCl, 2 mM BME.

775

776 **Sacculi preparation and RBB labeling.** Sacculi were prepared from all species in the same  
777 manner. For a typical 2L prep, cells were grown to an OD of 0.5-0.7 in their respective medium  
778 (see bacterial strains and growth conditions) and harvested by centrifugation at 6,000g for 20  
779 minutes. *C. crescentus* cells usually required multiple centrifugation steps to collect all the cells.  
780 Cells were resuspended in 25 mL water (or PBS for *E. coli*) and added drop-wise into 50 mL of  
781 boiling 7.5% SDS under stirring. The mixture was boiled for 30 minutes and then allowed to  
782 cool to room temperature. Sacculi were then pelleted by ultracentrifugation at 100,000g for 30  
783 minutes at room temperature. The resulting pellets were resuspended in 100 mL pure water, and  
784 washed repeatedly until SDS was no longer detected in the supernatant. The pellet was  
785 confirmed to be clear of SDS by mixing 0.2 mL of the supernatant with 1 uL 0.5% methylene  
786 blue, 0.1 mL 0.7M NaPO<sub>4</sub> pH 7.2, and 0.6 mL chloroform and checking to make sure that, after  
787 vortexing and allowing to settle, the solution had an upper blue phase and a lower clear phase  
788 [82]. At this point, the pellets were resuspended in 10 mL PBS with 20 mM MgSO<sub>4</sub>, 250 U/uL  
789 Pierce Universal Nuclease (Thermo Fisher Scientific), and 10 mg/mL amylase (Sigma). The  
790 mixture was incubated at 37°C for 1-4 hours. Afterwards, 10 mg/mL trypsin and 10 mM CaCl<sub>2</sub>  
791 was added and the mixture incubated overnight at 37°C. 800 uL of 15% SDS were then added to  
792 the mixture and it was brought to a boil for about 10 minutes and allowed to cool to room  
793 temperature. The sample was then pelleted (100,000g, 30 min, room temperature) and  
794 resuspended in 4-5 wash steps until SDS was no longer detected. The final pellet was then  
795 resuspended in 2 mL water and added to 0.8 mL 0.2M remazol brilliant blue (Sigma), 0.4 mL  
796 5M NaOH, and additional water to 8 mL. The mixture was incubated, shaking, overnight at  
797 37°C. After neutralizing the solution with 0.4 mL 5M HCl. The mixture was then pelleted  
798 (21,000g, 20 minutes, room temperature), and resuspended in water until the supernatant became  
799 clear.

800 To calibrate the concentration of RBB-labeled sacculi for dye-release assays and  
801 peptidoglycan-binding assays, activity curves with Hen Egg White Lysozyme (Sigma) were  
802 produced using different dilutions of the RBB-labeled sacculi. The RBB-labeled sacculi were  
803 used at the dilution that resulted in an A595 of 0.5 when 5 µL of the RBB-labeled sacculi were  
804 incubated with 4 uM HEWL.

805

806 **Peptidoglycan-binding assays.** 5µL of calibrated RBB-labeled sacculi were incubated with 1 µM  
807 of purified protein (protein constructs used are shown in Fig. S4) in PBS pH 7.4 to a final  
808 volume of 50 µL for 30 minutes at 37°C and then pelleted (16,000g, 20 minutes). Fractions were  
809 separated and the pellet resuspended in 50 µL PBS. 10 µL of each fraction was loaded onto Any  
810 kD Mini-PROTEAN TGX Precast Protein Gels (BioRad) to visualize whether the protein  
811 associated with the insoluble sacculi fraction. BSA (Sigma) was used at 1 µM as a negative  
812 control.

813

814 **Remazol brilliant blue dye-release assays.** Methods were adapted from [32,83]. Assays were  
815 carried out in 25-µL reactions using 25 mM HEPES pH 7.5, 100 mM NaCl and 5 µL of  
816 calibrated RBB-labeled sacculi. Enzymes were added at various concentrations (see Figs. 5 and  
817 S5) and incubated overnight at 37°C. Reactions were then centrifuged for 20 minutes at 16,000g,  
818 and the supernatant carefully separated from the pellet. Final values in Fig. 5 and S5 are  
819 normalized against absorbances measured for reactions with HEWL that were run in tandem for  
820 every measurement to correct for differences in different sacculi preparations.

821

822 **Fluorescence microscopy and image analysis.** Fluorescence imaging was done using an  
823 inverted Nikon Ti-E microscope using a Plan Apo 60X 1.40 NA oil Ph3 DM objective with a  
824 GFP/Cy3 filter cube and an Andor DU885 EM CCD camera. Images were captured using NIS  
825 Elements (Nikon). Cells were mounted on 1% (w/v) agarose pads made with PYE (or PBS, in  
826 the case of *E. coli*) for imaging. In general, the fluorescent channel of each image was  
827 background subtracted and a Gaussian Blur filter was applied using Fiji [84]. Quantification of  
828 stalk morphotypes and stalks with multiple foci was done by hand using Fiji tools. Quantification  
829 of fluorescence data was achieved using MicrobeJ [85]. Mean stalk intensity was measured in *A.*  
830 *biprosthicum* cells by using Fiji to draw line ROIs that did not overlap with the focus at the base  
831 of the stalk, measuring mean fluorescence along the ROI. Figures and statistics were performed  
832 using GraphPad Prism version 8.00 for Mac, GraphPad Software, La Jolla California USA,  
833 [www.graphpad.com](http://www.graphpad.com).

834

835 **Western blots.** Strains were grown to saturation (overnight for *C. crescentus* and *A. excentricus*,  
836 usually 48 hours for *A. biprosthicum*). OD600 was determined and cells were collected at a  
837 normalized density of OD600 = 1/1mL. 1 mL of each normalized culture was pelleted,  
838 resuspended in 100 uL water, and prepared for analysis using standard procedures using SDS-  
839 PAGE, transfer, and western blotting. 10 µL of each sample was loaded onto Any kD Mini-  
840 PROTEAN TGX Precast Protein Gels (BioRad). The JL-8 monoclonal GFP antibody (Clontech)  
841 was used as the primary antibody and Goat Anti-mouse HRP (Pierce) was used for the secondary  
842 antibody. Transferred blots were visualized with SuperSignal West Dura Extended Duration  
843 HRP substrate (ThermoFisher Scientific) using a Bio-Rad Chemidoc.

844

845 **Periplasmic expression in *E. coli*.** Fresh Lemo(DE3) competent cells (NEB) were transformed  
846 with pET22b constructs containing various muramidase genes with N-terminal H-tags (**Table**  
847 **S3**) and plated. Lemo21(DE3) carries a rhamnose-inducible copy of LysY that inhibits T7  
848 polymerase and allows for tunable dampening of expression of toxic products. We could not  
849 transform expression strains BL21(DE3) or Tuner(DE3) with the pET22b-P22Lyso construct,  
850 but were able to isolate a few transformants carrying this construct using Lemo21(DE3) cells  
851 under high rhamnose repression (2 mM). P22Lyso-D20L, and all the SpmX-Mur-*Cc* constructs,  
852 efficiently transformed into all expression strains tested, and could be carried by Lemo21(DE3)  
853 without rhamnose.

854

855 In the case of pET22b-P22Lyso, where cells lyse from leak, cell cultures were grown  
856 directly from colonies in the presence of 5 mM rhamnose and monitored over time. Figure 4A  
857 shows the same treatment for all tested constructs. For testing induction of CCM and P22Lyso-  
858 D20L, colonies were grown overnight in LB with 100 ug/mL carbenicillin and 30 µg/mL  
859 chloramphenicol. Overnight cultures were diluted 50-fold in LB medium with 100 ug/mL  
860 carbenicillin and 30 µg/mL chloramphenicol. In experiments cases rhamnose was added at  
861 specified concentrations. Typically 4 mL cultures of cells were grown for 1-1.5 hours to an  
862 OD600 of 0.3-0.4, induced with 400 uM IPTG, and shifted to 20°C.

862

863 *Growth curves and live-dead staining:* Optical densities were measured over time and  
864 cells were routinely checked for lysing by microscopy using standard procedures. Briefly, 1 uL  
865 of a 1:1 mixture of solutions A and B from a LIVE/DEAD *BacLight* Bacterial Viability Kit  
866 (ThermoFisher Scientific) was directly added to 100 uL of cells diluted 1:10 in PBS. Cells were  
867 visualized on 1% agar pads made with PBS using the methods described in microscopy.

867 *Periplasmic expression levels:* After growing for 4 h at 20 °C, OD600 was determined and cells  
868 were collected at a normalized density of OD600 = 1/1mL. One mL of the normalized sample  
869 was pelleted at 4000g for 15 min and the pellet resuspended in 250 uL 20% sucrose, 1 mM  
870 EDTA, 30 mM TRIS pH 8 at room temperature. The sample was mixed gently by rotation at  
871 room temperature for 10 minutes before being spun down at 13,000g for 10 minutes. The  
872 supernatant was carefully removed and the pellet rapidly suspended in 250 uL ice cold pure  
873 water. The sample was mixed gently by rotation at 4°C for 10 minutes before being spun down  
874 at 13,000g at 4°C. The supernatant (periplasmic fraction) and pellet (cell fraction) were then  
875 separated and prepared for analysis using standard procedures using SDS-PAGE, transfer, and  
876 western blotting. Blots were incubated with His-Probe Antibody (H-3) sc-8136 HRP (Santa Cruz  
877 Biotechnology) and visualized with SuperSignal West Dura Extended Duration HRP substrate  
878 (ThermoFisher Scientific) using a Bio-Rad Chemidoc.

879

## 880 **Supplemental Figure Legends**

881

882 **Figure S1. Alignments of SpmX muramidases. Related to Figure 1.** Aligned protein  
883 sequences of the muramidase domains from genes listed in **Table S1**, a subset of which were  
884 used for the gene tree in **Fig. 1**. Key positions discussed in text are marked with a black dot and  
885 annotated with T4L identities.

886

887 **Figure S2. Relationships of enzyme families within the lysozyme superfamily. Related to**  
888 **Figures 1 and 2.** Sequence cluster tree diagram made using NCBI's Conserved Domain  
889 Database tool (CDD) demonstrating the inferred relationship of various lysozyme families with  
890 structural homology but no sequence similarity. This tool uses reverse position-specific BLAST,  
891 a method that compares query sequences to databases of position-specific score matrices and  
892 obtains *E*-values, such as in PSI-BLAST [21,22]. SpmX from *C. crescentus* is marked in red  
893 type; bacteriophages discussed in the text are emphasized in bold, black type. The tree was  
894 formatted using iTol [65].

895

896 **Figure S3. Binding assays of SpmX to peptidoglycan from various species. Related to**  
897 **Figure 4.** Coomassie brilliant blue stained SDS-PAGE gel of supernatant (S) and pellet (P)  
898 fractions from peptidoglycan binding assays. In brief, SpmX protein constructs from three  
899 different species and of two different lengths were incubated in PBS with or without sacculi  
900 (peptidoglycan) isolated from the three different species. Peptidoglycan is insoluble and can be  
901 pelleted out of solution. As shown here, peptidoglycan from all three species pulled down the  
902 SpmX protein constructs.

903

904 **Figure S4. *In vitro* activity and periplasmic expression/activity of SpmX and various**  
905 **mutants. Related to Figure 4.** Active enzymes release peptidoglycan monomers covalently-  
906 bound to RBB into the supernatant that are detected by absorbance at 595 nm. **(A)** RBB assays  
907 with SpmX-Mur-*Cc*-E11A, the catalytic mutant. Due to poor expression and instability, SpmX-  
908 Mur-E11A was not concentrated to higher than 12 μM and its activity is compared to WT SpmX  
909 at 12 μM. Error bars are ± standard deviation from two replicates. **(B)** RBB assays with SpmX-  
910 Mur-*Cc*-L20D, the mutant that restores the ancestral D20. **(C)** Final ODs of cultures expressing  
911 various P22Lyso and SpmX-Mur-*Cc* constructs after four hours of induction. P22Lyso-D20L  
912 showed signs of spheroplast formation and cell lysis but the other tested constructs did not.



913 Compare to growth curves and microscopy shown in Fig. 4CD. **(D)** SpmX-Mur-*Cc* is expressed  
914 at similar levels in the periplasm as P22Lyso-D20L. Attempts at expressing SpmX-Mur-*Cc* more  
915 slowly in M9G medium at lower temperatures increased periplasmic levels but did not result in  
916 lysis. **(E)** RBB assays on *E. coli* sacculi isolated from Lemo21(DE3) using purified P22  
917 lysozyme, P22 lysozyme D20L mutant, *C. crescentus* SpmX muramidase, and SpmX  
918 muramidase L20D mutant. Lines are drawn to help guide the eye toward basic trends.  
919 Normalization is to Hen Egg White Lysozyme (HEWL) assays (see methods).

920

921 **Figure S5. Quantification of fluorescence and morphology data of SpmX mutant strains.**  
922 **Related to Figure 5. (A)** Histograms of mean total cell fluorescence and integrated focal  
923 intensity of *C. crescentus* cells expressing fluorescently tagged WT (n = 1057 cells) and E11A (n  
924 = 792 cells) SpmX. Insets display the mean of these measurements with the standard deviation as  
925 error bars. Asterisks indicate  $p < 0.001$  by Mann-Whitney U test. **(B)** Histograms (as in **A**) of  
926 mean total cell fluorescence and integrated focal intensity of *A. biprosthicum* cell expressing  
927 fluorescently tagged WT (n = 494 cells) and E11A (n = 462 cells) SpmX. Far right panel:  
928 boxplot of median and quartiles with whiskers indicating minimum and maximum values of  
929 mean fluorescence intensity of stalks in *A. biprosthicum* expressing WT or E11A (n = 10 stalks  
930 each strain). **(C)** Histogram (as in **A**) of mean total cell fluorescence and integrated focal  
931 intensity of *A. excentricus* expressing fluorescently tagged WT (n = 750 cells) and E11A (n =  
932 693 cells) SpmX. Far right panel: quantification of the frequency of cells with multiple SpmX  
933 foci in the stalks of *A. excentricus* WT and E11A mutant. Chi-squared  $p < 0.001$ , n = 600 cells.  
934 **(D)** Quantification of the stalked morphotypes from *A. biprosthicum* WT and *spmX-E11A*  
935 mutant from cells without eGFP tags. Chi-squared  $p < 0.001$ , n = 500 cells. **(E)** Panels show  
936 bands from two independent blots using samples normalized by OD. Lanes have been cropped  
937 and rearranged to aid interpretation. In all cases, the primary antibody is directed against the C-  
938 terminal GFP fusion.

939

## 940 References

- 941 1. Hall, J.P.J., Brockhurst, M.A., and Harrison, E. (2017). Sampling the mobile gene pool:  
942 innovation via horizontal gene transfer in bacteria. *Phil Trans R Soc B* 372, 20160424.
- 943 2. Koonin, E.V. (2016). Viruses and mobile elements as drivers of evolutionary transitions. *Phil*  
944 *Trans R Soc B* 371, 20150442.
- 945 3. Feiner, R., Argov, T., Rabinovich, L., Sigal, N., Borovok, I., and Herskovits, A.A. (2015). A  
946 new perspective on lysogeny: prophages as active regulatory switches of bacteria. *Nat. Rev.*  
947 *Microbiol.* 13, 641–650.
- 948 4. Harrison, E., and Brockhurst, M.A. (2017). Ecological and Evolutionary Benefits of  
949 Temperate Phage: What Does or Doesn't Kill You Makes You Stronger. *BioEssays* 39,  
950 1700112.
- 951 5. Howard-Varona, C., Hargreaves, K.R., Abedon, S.T., and Sullivan, M.B. (2017). Lysogeny  
952 in nature: mechanisms, impact and ecology of temperate phages. *ISME J.* 11, 1511–1520.
- 953 6. Menouni, R., Hutinet, G., Petit, M.-A., and Ansaldi, M. (2015). Bacterial genome  
954 remodeling through bacteriophage recombination. *FEMS Microbiol. Lett.* 362, 1–10.

- 955 7. Bobay, L.-M., Touchon, M., and Rocha, E.P.C. (2014). Pervasive domestication of defective  
956 prophages by bacteria. *Proc. Natl. Acad. Sci.* *111*, 12127–12132.
- 957 8. Touchon, M., Moura de Sousa, J.A., and Rocha, E.P. (2017). Embracing the enemy: the  
958 diversification of microbial gene repertoires by phage-mediated horizontal gene transfer.  
959 *Curr. Opin. Microbiol.* *38*, 66–73.
- 960 9. Wang, X., Kim, Y., Ma, Q., Hong, S.H., Pokusaeva, K., Sturino, J.M., and Wood, T.K.  
961 (2010). Cryptic prophages help bacteria cope with adverse environments. *Nat. Commun.* *1*,  
962 147.
- 963 10. Leiman, P.G., Basler, M., Ramagopal, U.A., Bonanno, J.B., Sauder, J.M., Pukatzki, S.,  
964 Burley, S.K., Almo, S.C., and Mekalanos, J.J. (2009). Type VI secretion apparatus and  
965 phage tail-associated protein complexes share a common evolutionary origin. *Proc. Natl.*  
966 *Acad. Sci.* *106*, 4154–4159.
- 967 11. Sarris, P.F., Ladoukakis, E.D., Panopoulos, N.J., and Scoulica, E.V. (2014). A Phage Tail-  
968 Derived Element with Wide Distribution among Both Prokaryotic Domains: A Comparative  
969 Genomic and Phylogenetic Study. *Genome Biol. Evol.* *6*, 1739–1747.
- 970 12. Ghequire, M.G.K., and De Mot, R. (2015). The Tailocin Tale: Peeling off Phage Tails.  
971 *Trends Microbiol.* *23*, 587–590.
- 972 13. Hockett, K.L., Renner, T., and Baltrus, D.A. (2015). Independent Co-Option of a Tailed  
973 Bacteriophage into a Killing Complex in *Pseudomonas*. *mBio* *6*, e00452-15.
- 974 14. Scholl, D. (2017). Phage Tail–Like Bacteriocins. *Annu. Rev. Virol.* *4*, 453–467.
- 975 15. Radhakrishnan, S.K., Thanbichler, M., and Viollier, P.H. (2008). The dynamic interplay  
976 between a cell fate determinant and a lysozyme homolog drives the asymmetric division  
977 cycle of *Caulobacter crescentus*. *Genes Dev.* *22*, 212–225.
- 978 16. Lasker, K., Mann, T.H., and Shapiro, L. (2016). An intracellular compass spatially  
979 coordinates cell cycle modules in *Caulobacter crescentus*. *Curr. Opin. Microbiol.* *33*, 131–  
980 139.
- 981 17. Jiang, C., Brown, P.J.B., Ducret, A., and Brun, Y.V. (2014). Sequential evolution of bacterial  
982 morphology by co-option of a developmental regulator. *Nature* *506*, 489–493.
- 983 18. Perez, A.M., Mann, T.H., Lasker, K., Ahrens, D.G., Eckart, M.R., and Shapiro, L. (2017). A  
984 Localized Complex of Two Protein Oligomers Controls the Orientation of Cell Polarity.  
985 *mBio* *8*, e02238-16.
- 986 19. Bowman, G.R., Comolli, L.R., Gaietta, G.M., Fero, M., Hong, S.-H., Jones, Y., Lee, J.H.,  
987 Downing, K.H., Ellisman, M.H., McAdams, H.H., *et al.* (2010). *Caulobacter* PopZ forms a  
988 polar subdomain dictating sequential changes in pole composition and function. *Mol.*  
989 *Microbiol.* *76*, 173–189.

- 990 20. Holmes, J.A., Follett, S.E., Wang, H., Meadows, C.P., Varga, K., and Bowman, G.R. (2016).  
991 Caulobacter PopZ forms an intrinsically disordered hub in organizing bacterial cell poles.  
992 Proc. Natl. Acad. Sci. *113*, 12490–12495.
- 993 21. Marchler-Bauer, A., Anderson, J.B., Derbyshire, M.K., DeWeese-Scott, C., Gonzales, N.R.,  
994 Gwadz, M., Hao, L., He, S., Hurwitz, D.I., Jackson, J.D., *et al.* (2007). CDD: a conserved  
995 domain database for interactive domain family analysis. Nucleic Acids Res. *35*, D237–D240.
- 996 22. Marchler-Bauer, A., Derbyshire, M.K., Gonzales, N.R., Lu, S., Chitsaz, F., Geer, L.Y., Geer,  
997 R.C., He, J., Gwadz, M., Hurwitz, D.I., *et al.* (2015). CDD: NCBI’s conserved domain  
998 database. Nucleic Acids Res. *43*, D222–D226.
- 999 23. Blackburn, N.T., and Clarke, A.J. (2001). Identification of Four Families of Peptidoglycan  
1000 Lytic Transglycosylases. J. Mol. Evol. *52*, 78–84.
- 1001 24. Evrard, C., Fastrez, J., and Declercq, J.-P. (1998). Crystal structure of the lysozyme from  
1002 bacteriophage lambda and its relationship with V and C-type lysozymes1 Edited by K.  
1003 Nagai. J. Mol. Biol. *276*, 151–164.
- 1004 25. Vollmer, W., Joris, B., Charlier, P., and Foster, S. (2008). Bacterial peptidoglycan (murein)  
1005 hydrolases. FEMS Microbiol. Rev. *32*, 259–286.
- 1006 26. Wohlkönig, A., Huet, J., Looze, Y., and Wintjens, R. (2010). Structural Relationships in the  
1007 Lysozyme Superfamily: Significant Evidence for Glycoside Hydrolase Signature Motifs.  
1008 PLOS ONE *5*, e15388.
- 1009 27. Rennell, D., Bouvier, S.E., Hardy, L.W., and Poteete, A.R. (1991). Systematic mutation of  
1010 bacteriophage T4 lysozyme. J. Mol. Biol. *222*, 67–88.
- 1011 28. Anand, N.N., Stephen, E.R., and Narang, S.A. (1988). Mutation of active site residues in  
1012 synthetic T4-lysozyme gene and their effect on lytic activity. Biochem. Biophys. Res.  
1013 Commun. *153*, 862–868.
- 1014 29. Sun, Q., Kutty, G.F., Arockiasamy, A., Xu, M., Young, R., and Sacchettini, J.C. (2009).  
1015 Regulation of a muralytic enzyme by dynamic membrane topology. Nat. Struct. Mol. Biol.  
1016 *16*, 1192–1194.
- 1017 30. Hardy, L.W., and Poteete, A.R. (1991). Reexamination of the role of Asp20 in catalysis by  
1018 bacteriophage T4 lysozyme. Biochemistry (Mosc.) *30*, 9457–9463.
- 1019 31. Poteete, A.R., Dao-Pin, S., Nicholson, H., and Mathews, B.W. (1991). Second-site revertants  
1020 of an inactive T4 lysozyme mutant restore activity by restructuring the active site cleft.  
1021 Biochemistry (Mosc.) *30*, 1425–1432.
- 1022 32. Zhou, R., Chen, S., and Recsei, P. (1988). A dye release assay for determination of  
1023 lysostaphin activity. Anal. Biochem. *171*, 141–144.

- 1024 33. Anderson, W.F., Grütter, M.G., Remington, S.J., Weaver, L.H., and Matthews, B.W. (1981).  
1025 Crystallographic determination of the mode of binding of oligosaccharides to T4  
1026 bacteriophage lysozyme: Implications for the mechanism of catalysis. *J. Mol. Biol.* *147*,  
1027 523–543.
- 1028 34. Braga, L.P.P., Soucy, S.M., Amgarten, D.E., Silva, D., M, A., and Setubal, J.C. (2018).  
1029 Bacterial Diversification in the Light of the Interactions with Phages: The Genetic Symbionts  
1030 and Their Role in Ecological Speciation. *Front. Ecol. Evol.* *6*. Available at:  
1031 <https://www.frontiersin.org/articles/10.3389/fevo.2018.00006/full> [Accessed August 30,  
1032 2018].
- 1033 35. Canchaya, C., Fournous, G., Chibani-Chennoufi, S., Dillman, M.-L., and Brüssow, H.  
1034 (2003). Phage as agents of lateral gene transfer. *Curr. Opin. Microbiol.* *6*, 417–424.
- 1035 36. Cortez, D., Forterre, P., and Gribaldo, S. (2009). A hidden reservoir of integrative elements  
1036 is the major source of recently acquired foreign genes and ORFans in archaeal and bacterial  
1037 genomes. *Genome Biol.* *10*, R65.
- 1038 37. Daubin, V., and Ochman, H. (2004). Bacterial genomes as new gene homes: the genealogy  
1039 of ORFans in *E. coli*. *Bacterial Genomes as New Gene Homes: The Genealogy of ORFans*  
1040 *in E. coli*. *Genome Res.* *14*, 1036, 1036–1042.
- 1041 38. Brézellec, P., Vallet-Gely, I., Possoz, C., Quevillon-Cheruel, S., and Ferat, J.-L. (2016).  
1042 DciA is an ancestral replicative helicase operator essential for bacterial replication initiation.  
1043 *Nat. Commun.* *7*, 13271.
- 1044 39. Brézellec, P., Petit, M.-A., Pasek, S., Vallet-Gely, I., Possoz, C., and Ferat, J.-L. (2017).  
1045 Domestication of Lambda Phage Genes into a Putative Third Type of Replicative Helicase  
1046 Matchmaker. *Genome Biol. Evol.* *9*, 1561–1566.
- 1047 40. Forterre, P. (1999). Displacement of cellular proteins by functional analogues from plasmids  
1048 or viruses could explain puzzling phylogenies of many DNA informational proteins. *Mol.*  
1049 *Microbiol.* *33*, 457–465.
- 1050 41. Sabehi, G., Shaulov, L., Silver, D.H., Yanai, I., Harel, A., and Lindell, D. (2012). A novel  
1051 lineage of myoviruses infecting cyanobacteria is widespread in the oceans. *Proc. Natl. Acad.*  
1052 *Sci. U. S. A.* *109*, 2037–2042.
- 1053 42. Lang, A.S., Zhaxybayeva, O., and Beatty, J.T. (2012). Gene transfer agents: phage-like  
1054 elements of genetic exchange. *Nat. Rev. Microbiol.* *10*, 472–482.
- 1055 43. Lang, A.S., and Beatty, J.T. (2007). Importance of widespread gene transfer agent genes in  
1056  $\alpha$ -proteobacteria. *Trends Microbiol.* *15*, 54–62.
- 1057 44. Shakya, M., Soucy, S.M., and Zhaxybayeva, O. (2017). Insights into origin and evolution of  
1058  $\alpha$ -proteobacterial gene transfer agents. *Virus Evol.* *3*. Available at: [https://academic-oup-](https://academic-oup-com.proxyiub.uits.iu.edu/ve/article/3/2/vex036/4705971)  
1059 [com.proxyiub.uits.iu.edu/ve/article/3/2/vex036/4705971](https://academic-oup-com.proxyiub.uits.iu.edu/ve/article/3/2/vex036/4705971) [Accessed August 31, 2018].

- 1060 45. Ho, B.T., Dong, T.G., and Mekalanos, J.J. (2014). A View to a Kill: The Bacterial Type VI  
1061 Secretion System. *Cell Host Microbe* *15*, 9–21.
- 1062 46. Russell, A.B., Peterson, S.B., and Mougous, J.D. (2014). Type VI secretion system effectors:  
1063 poisons with a purpose. *Nat. Rev. Microbiol.* *12*, 137–148.
- 1064 47. Hurst, M.R.H., Glare, T.R., and Jackson, T.A. (2004). Cloning *Serratia entomophila*  
1065 Antifeeding Genes—a Putative Defective Prophage Active against the Grass Grub *Costelytra*  
1066 *zealandica*. *J. Bacteriol.* *186*, 5116–5128.
- 1067 48. Yang, G., Dowling, A.J., Gerike, U., ffrench-Constant, R.H., and Waterfield, N.R. (2006).  
1068 *Photorhabdus* Virulence Cassettes Confer Injectable Insecticidal Activity against the Wax  
1069 Moth. *J. Bacteriol.* *188*, 2254–2261.
- 1070 49. Shikuma, N.J., Pilhofer, M., Weiss, G.L., Hadfield, M.G., Jensen, G.J., and Newman, D.K.  
1071 (2014). Marine Tubeworm Metamorphosis Induced by Arrays of Bacterial Phage Tail-Like  
1072 Structures. *Science* *343*, 529–533.
- 1073 50. Patzer, S.I., Albrecht, R., Braun, V., and Zeth, K. (2012). Structural and Mechanistic Studies  
1074 of Pesticin, a Bacterial Homolog of Phage Lysozymes. *J. Biol. Chem.* *287*, 23381–23396.
- 1075 51. Michalska, K., Brown, R.N., Li, H., Jedrzejczak, R., Niemann, G.S., Heffron, F., Cort, J.R.,  
1076 Adkins, J.N., Babnigg, G., and Joachimiak, A. (2013). New sub-family of lysozyme-like  
1077 proteins shows no catalytic activity: crystallographic and biochemical study of STM3605  
1078 protein from *Salmonella* Typhimurium. *J. Struct. Funct. Genomics* *14*, 1–10.
- 1079 52. Ren, Q., Wang, C., Jin, M., Lan, J., Ye, T., Hui, K., Tan, J., Wang, Z., Wyckoff, G.J., Wang,  
1080 W., *et al.* (2017). Co-option of bacteriophage lysozyme genes by bivalve genomes. *Open*  
1081 *Biol.* *7*, 160285.
- 1082 53. Crooks, G.E., Hon, G., Chandonia, J.-M., and Brenner, S.E. (2004). WebLogo: A Sequence  
1083 Logo Generator. *Genome Res.* *14*, 1188–1190.
- 1084 54. Meisner, J., Llopis, P.M., Sham, L.-T., Garner, E., Bernhardt, T.G., and Rudner, D.Z. (2013).  
1085 FtsEX is required for CwlO peptidoglycan hydrolase activity during cell wall elongation in  
1086 *Bacillus subtilis*. *Mol. Microbiol.* *89*, 1069–1083.
- 1087 55. Thanbichler, M., Iniesta, A.A., and Shapiro, L. (2007). A comprehensive set of plasmids for  
1088 vanillate- and xylose-inducible gene expression in *Caulobacter crescentus*. *Nucleic Acids*  
1089 *Res.* *35*, e137–e137.
- 1090 56. Poindexter, J.L.S., and Cohen-Bazire, G. (1964). The Fine Structure of Stalked Bacteria  
1091 Belonging to the Family Caulobacteraceae. *J. Cell Biol.* *23*, 587–607.
- 1092 57. Pate, J.L., and Ordal, E.J. (1965). THE FINE STRUCTURE OF TWO UNUSUAL  
1093 STALKED BACTERIA. *J. Cell Biol.* *27*, 133–150.

- 1094 58. Chen, I.-M.A., Markowitz, V.M., Chu, K., Palaniappan, K., Szeto, E., Pillay, M., Ratner, A.,  
1095 Huang, J., Andersen, E., Huntemann, M., *et al.* (2017). IMG/M: integrated genome and  
1096 metagenome comparative data analysis system. *Nucleic Acids Res.* *45*, D507–D516.
- 1097 59. Edgar, R.C. (2004). MUSCLE: multiple sequence alignment with high accuracy and high  
1098 throughput. *Nucleic Acids Res.* *32*, 1792–1797.
- 1099 60. Waterhouse, A.M., Procter, J.B., Martin, D.M.A., Clamp, M., and Barton, G.J. (2009).  
1100 Jalview Version 2—a multiple sequence alignment editor and analysis workbench.  
1101 *Bioinformatics* *25*, 1189–1191.
- 1102 61. Ashkenazy, H., Erez, E., Martz, E., Pupko, T., and Ben-Tal, N. (2010). ConSurf 2010:  
1103 calculating evolutionary conservation in sequence and structure of proteins and nucleic acids.  
1104 *Nucleic Acids Res.* *38*, W529–W533.
- 1105 62. Tatusova, T., Ciufo, S., Fedorov, B., O’Neill, K., and Tolstoy, I. (2014). RefSeq microbial  
1106 genomes database: new representation and annotation strategy. *Nucleic Acids Res.* *42*,  
1107 D553–D559.
- 1108 63. Darling, A.E., Jospin, G., Lowe, E., Iv, F.A.M., Bik, H.M., and Eisen, J.A. (2014). PhyloSift:  
1109 phylogenetic analysis of genomes and metagenomes. *PeerJ* *2*, e243.
- 1110 64. Ronquist, F., and Huelsenbeck, J.P. (2003). MrBayes 3: Bayesian phylogenetic inference  
1111 under mixed models. *Bioinformatics* *19*, 1572–1574.
- 1112 65. Letunic, I., and Bork, P. (2016). Interactive tree of life (iTOL) v3: an online tool for the  
1113 display and annotation of phylogenetic and other trees. *Nucleic Acids Res.* *44*, W242–W245.
- 1114 66. Gonin, M., Quardokus, E.M., O’Donnol, D., Maddock, J., and Brun, Y.V. (2000). Regulation  
1115 of Stalk Elongation by Phosphate in *Caulobacter crescentus*. *J. Bacteriol.* *182*, 337–347.
- 1116 67. Figueroa-Cuilan, W., Daniel, J.J., Howell, M., Sulaiman, A., and Brown, P.J.B. (2016).  
1117 Mini-Tn7 Insertion in an Artificial attTn7 Site Enables Depletion of the Essential Master  
1118 Regulator CtrA in the Phytopathogen *Agrobacterium tumefaciens*. *Appl Env. Microbiol* *82*,  
1119 5015–5025.
- 1120 68. Waldo, G.S., Standish, B.M., Berendzen, J., and Terwilliger, T.C. (1999). Rapid protein-  
1121 folding assay using green fluorescent protein. *Nat. Biotechnol.* *17*, 691–695.
- 1122 69. Marblestone, J.G., Edavettal, S.C., Lim, Y., Lim, P., Zuo, X., and Butt, T.R. (2006).  
1123 Comparison of SUMO fusion technology with traditional gene fusion systems: enhanced  
1124 expression and solubility with SUMO. *Protein Sci. Publ. Protein Soc.* *15*, 182–189.
- 1125 70. Uehara, T., Parzych, K.R., Dinh, T., and Bernhardt, T.G. (2010). Daughter cell separation is  
1126 controlled by cytokinetic ring-activated cell wall hydrolysis. *EMBO J.* *29*, 1412–1422.
- 1127 71. Bowler, M.W., Nurizzo, D., Barrett, R., Beteva, A., Bodin, M., Caserotto, H., Delagenière,  
1128 S., Dobias, F., Flot, D., Giraud, T., *et al.* (2015). MASSIF-1: a beamline dedicated to the

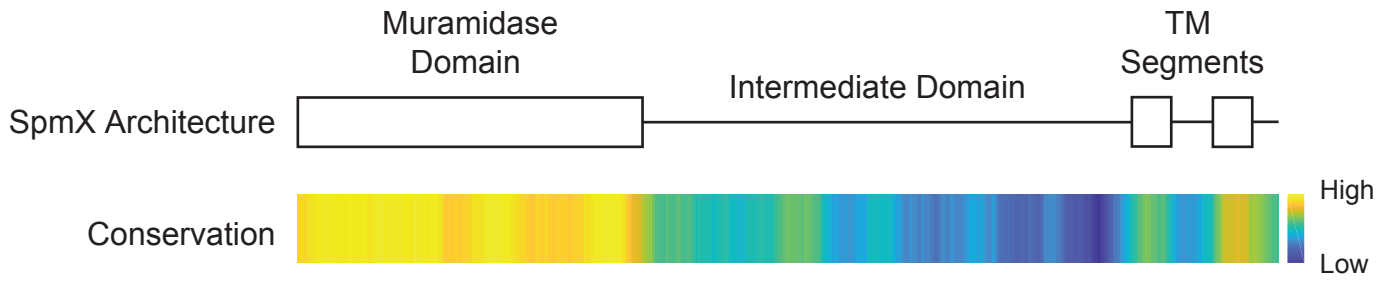
- 1129 fully automatic characterization and data collection from crystals of biological  
1130 macromolecules. *J. Synchrotron Radiat.* *22*, 1540–1547.
- 1131 72. Svensson, O., Malbet-Monaco, S., Popov, A., Nurizzo, D., and Bowler, M.W. (2015). Fully  
1132 automatic characterization and data collection from crystals of biological macromolecules.  
1133 *Acta Crystallogr. D Biol. Crystallogr.* *71*, 1757–1767.
- 1134 73. Kabsch, W. (2010). XDS. *Acta Crystallogr. D Biol. Crystallogr.* *66*, 125–132.
- 1135 74. McCoy, A.J., Grosse-Kunstleve, R.W., Adams, P.D., Winn, M.D., Storoni, L.C., and Read,  
1136 R.J. (2007). Phaser crystallographic software. *J. Appl. Crystallogr.* *40*, 658–674.
- 1137 75. Terwilliger, T.C., Grosse-Kunstleve, R.W., Afonine, P.V., Moriarty, N.W., Zwart, P.H.,  
1138 Hung, L.-W., Read, R.J., and Adams, P.D. (2008). Iterative model building, structure  
1139 refinement and density modification with the PHENIX AutoBuild wizard. *Acta Crystallogr.*  
1140 *D Biol. Crystallogr.* *64*, 61–69.
- 1141 76. Emsley, P., and Cowtan, K. (2004). Coot: model-building tools for molecular graphics. *Acta*  
1142 *Crystallogr. D Biol. Crystallogr.* *60*, 2126–2132.
- 1143 77. Langer, G.G., Cohen, S.X., Lamzin, V.S., and Perrakis, A. (2008). Automated  
1144 macromolecular model building for X-ray crystallography using ARP/wARP version 7. *Nat.*  
1145 *Protoc.* *3*, 1171–1179.
- 1146 78. Murshudov, G.N., Skubák, P., Lebedev, A.A., Pannu, N.S., Steiner, R.A., Nicholls, R.A.,  
1147 Winn, M.D., Long, F., and Vagin, A.A. (2011). REFMAC5 for the refinement of  
1148 macromolecular crystal structures. *Acta Crystallogr. D Biol. Crystallogr.* *67*, 355–367.
- 1149 79. Brünger, A.T. (1992). Free R value: a novel statistical quantity for assessing the accuracy of  
1150 crystal structures. *Nature* *355*, 472–475.
- 1151 80. Laskowski, R.A., MacArthur, M.W., Moss, D.S., and Thornton, J.M. (1993). PROCHECK: a  
1152 program to check the stereochemical quality of protein structures. *J Appl Crystallogr.* *283*–  
1153 *291*.
- 1154 81. Kabsch, W., and Sander, C. (1983). Dictionary of protein secondary structure: pattern  
1155 recognition of hydrogen-bonded and geometrical features. *Biopolymers* *22*, 2577–2637.
- 1156 82. Hayashi, K. (1975). A rapid determination of sodium dodecyl sulfate with methylene blue.  
1157 *Anal. Biochem.* *67*, 503–506.
- 1158 83. Uehara, T., Parzych, K.R., Dinh, T., and Bernhardt, T.G. (2010). Daughter cell separation is  
1159 controlled by cytokinetic ring-activated cell wall hydrolysis. *EMBO J.* *29*, 1412–1422.
- 1160 84. Schindelin, J., Arganda-Carreras, I., Frise, E., Kaynig, V., Longair, M., Pietzsch, T.,  
1161 Preibisch, S., Rueden, C., Saalfeld, S., Schmid, B., *et al.* (2012). Fiji: an open-source  
1162 platform for biological-image analysis. *Nat. Methods* *9*, 676–682.

1163 85. Ducret, A., Quardokus, E.M., and Brun, Y.V. (2016). MicrobeJ, a tool for high throughput  
1164 bacterial cell detection and quantitative analysis. *Nat. Microbiol. 1*, 16077.

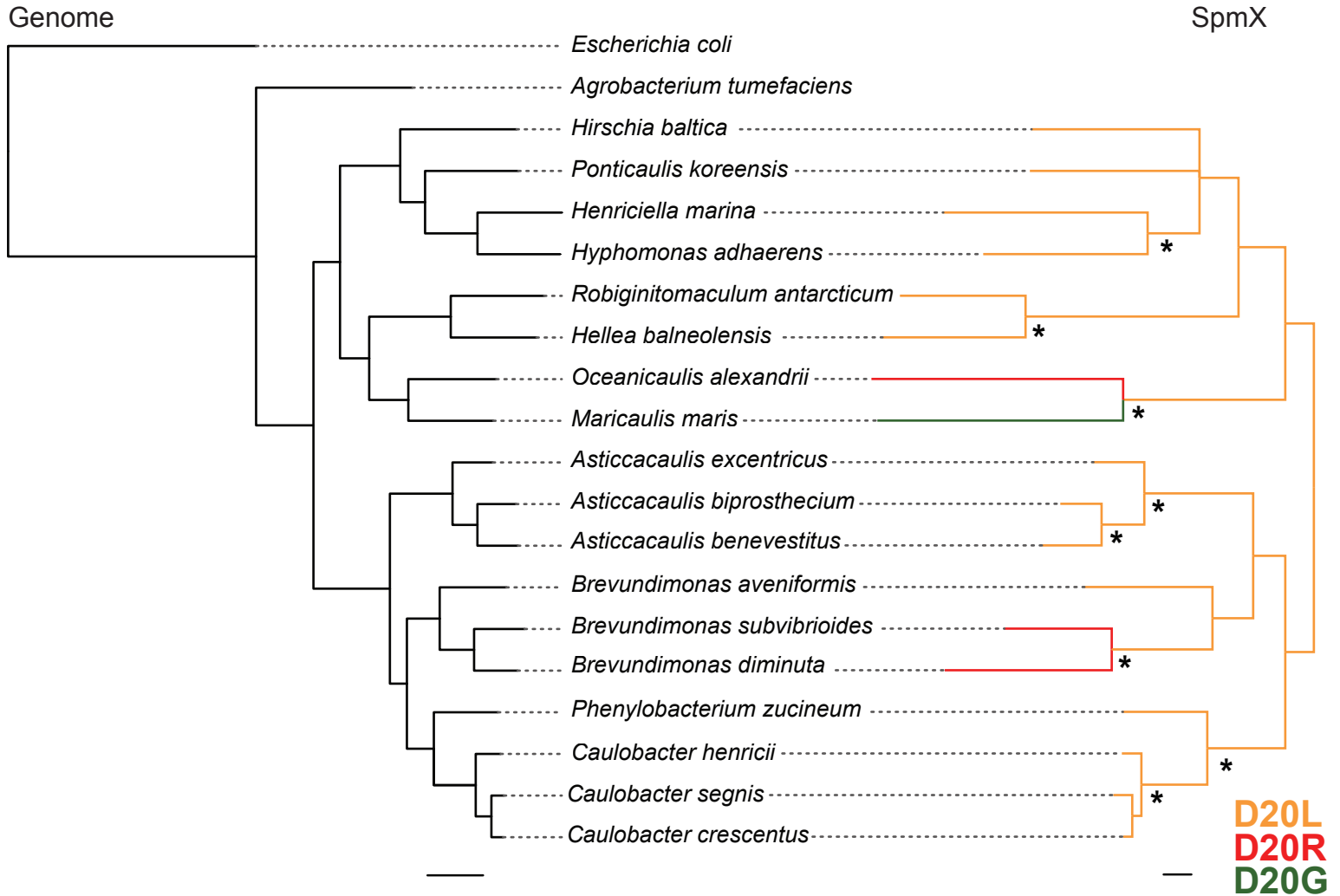
1165



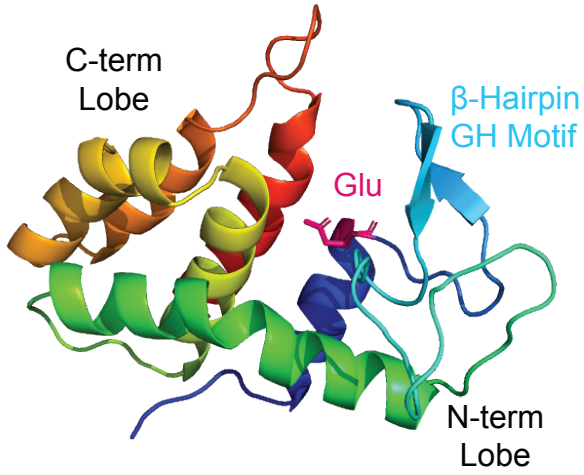
A



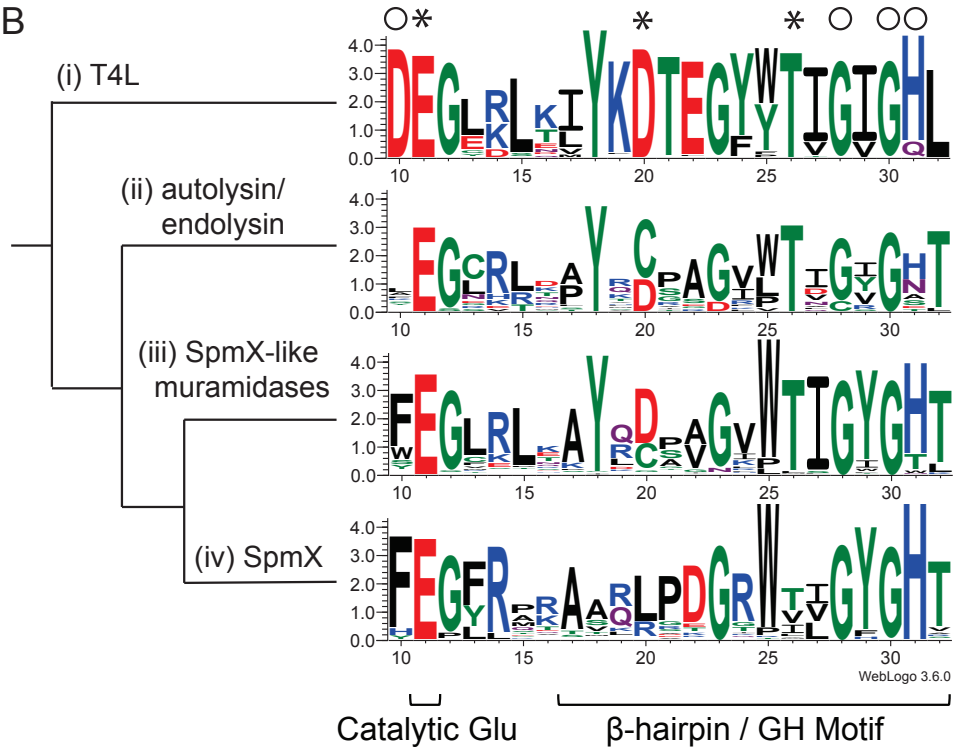
B

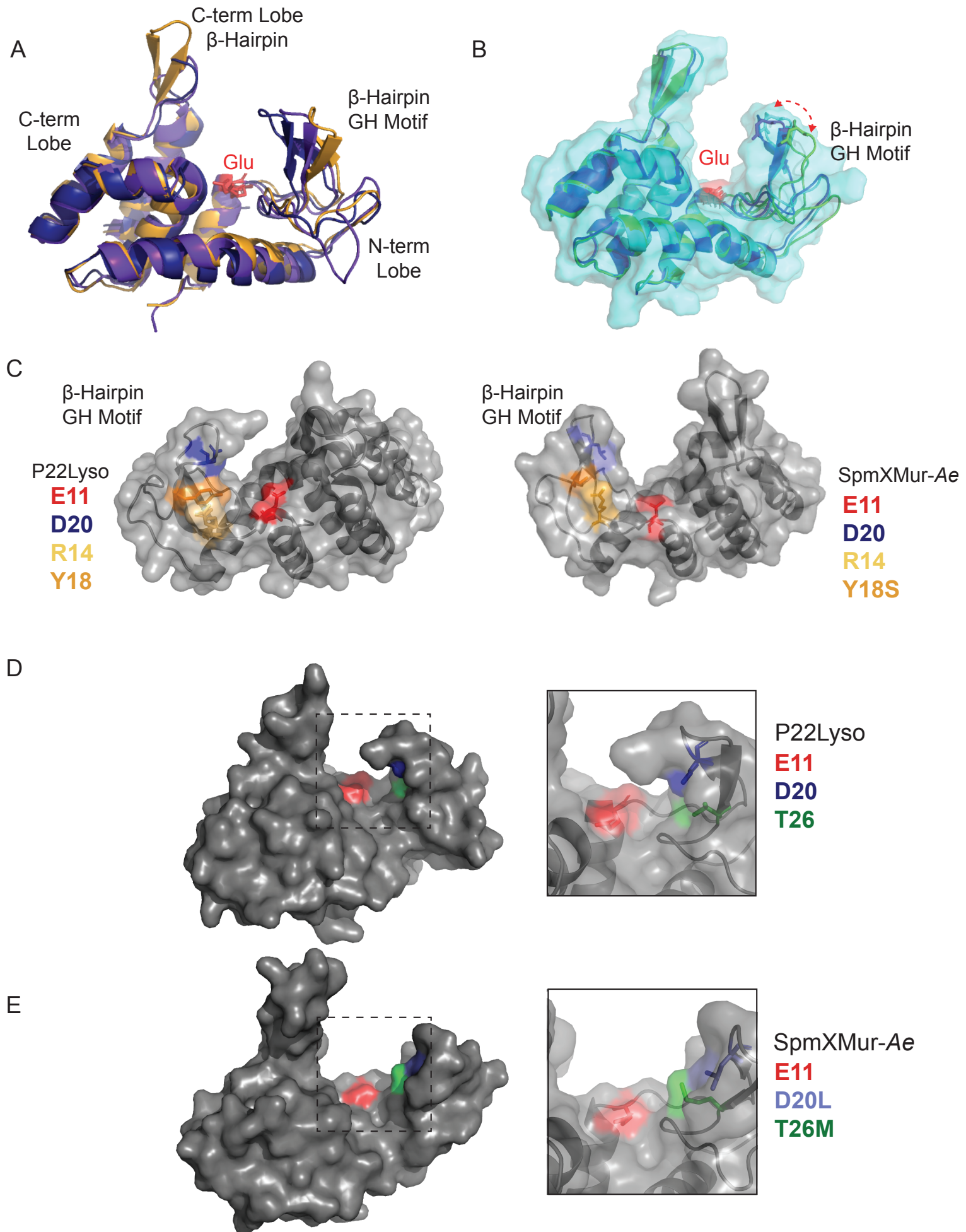


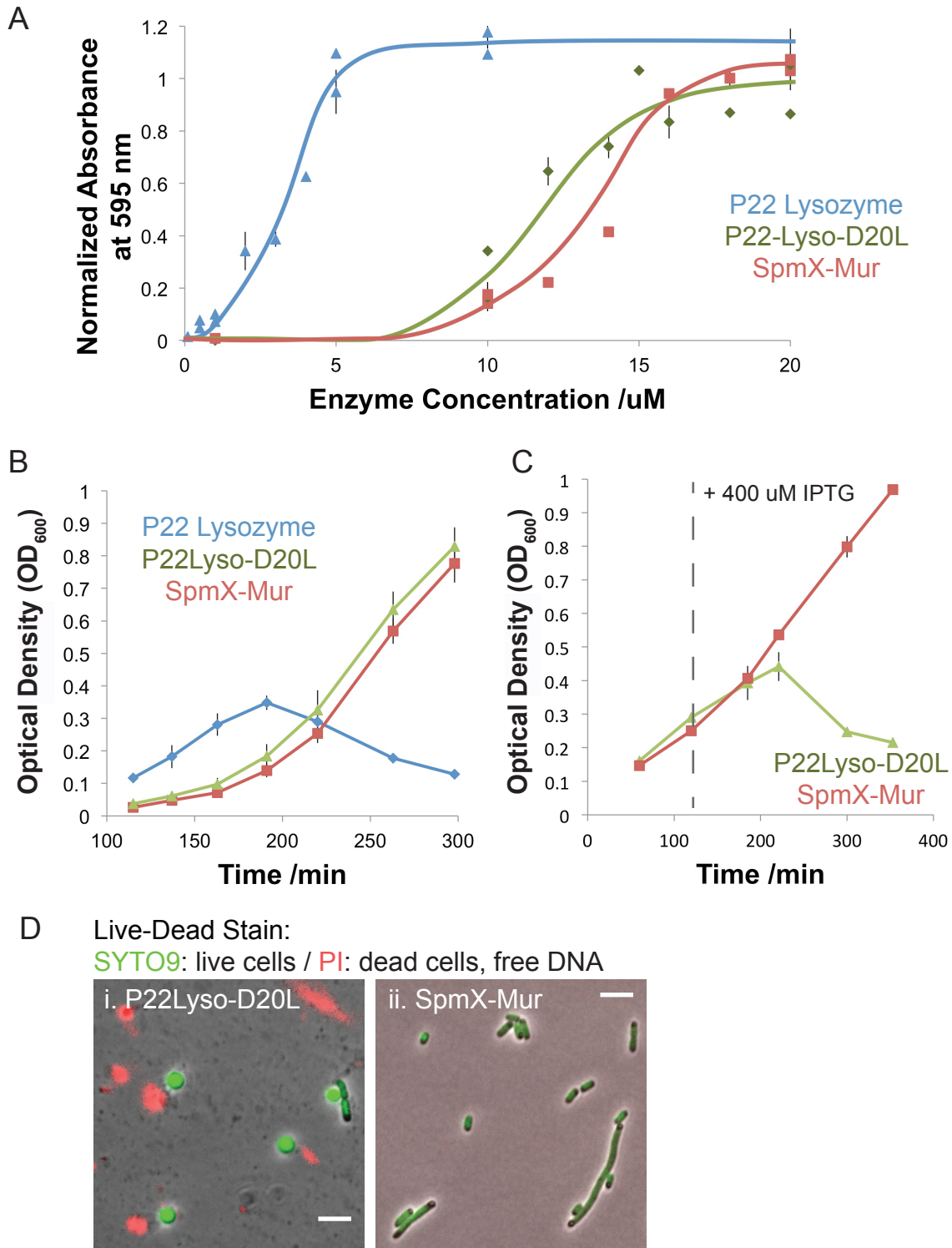
A

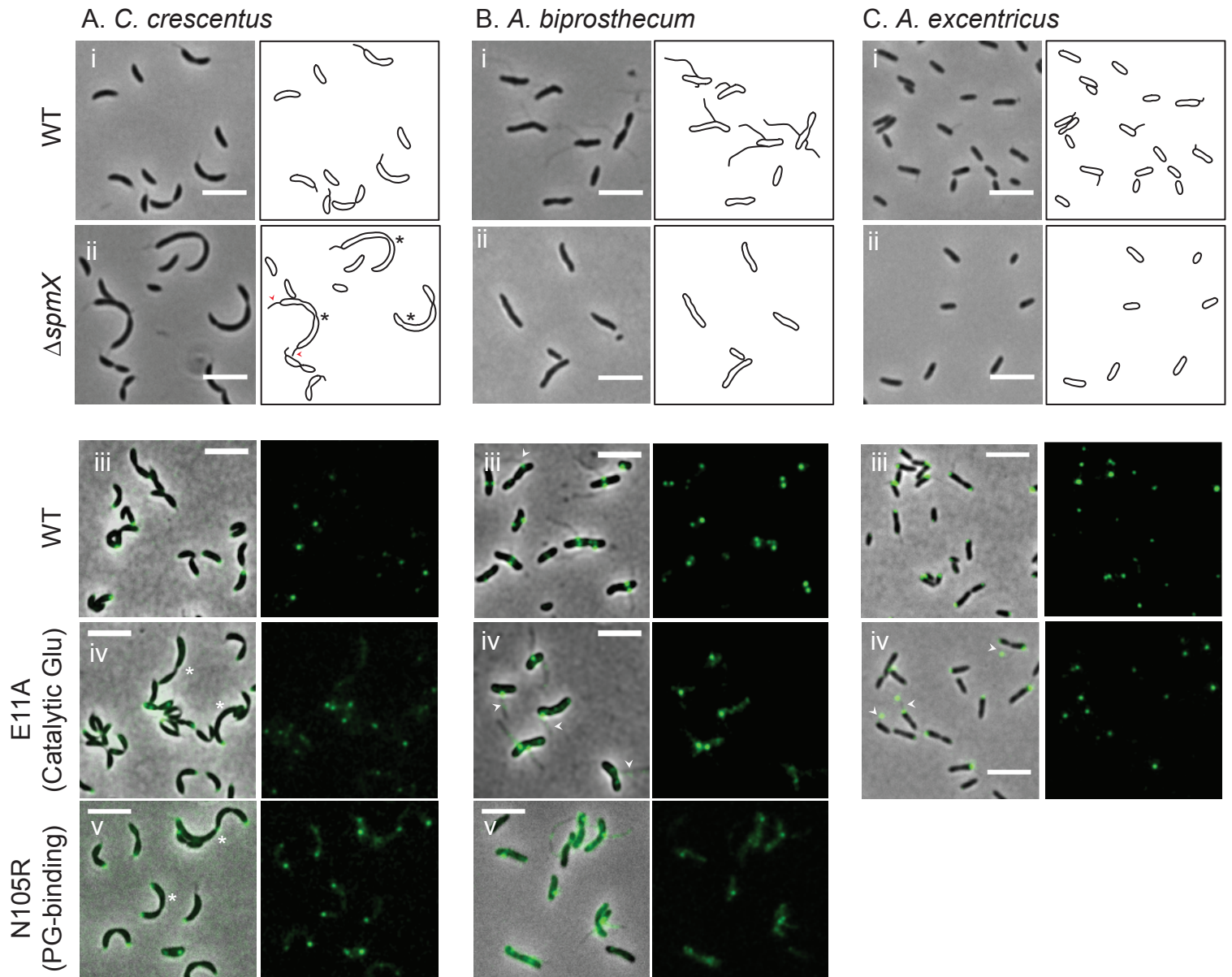


B

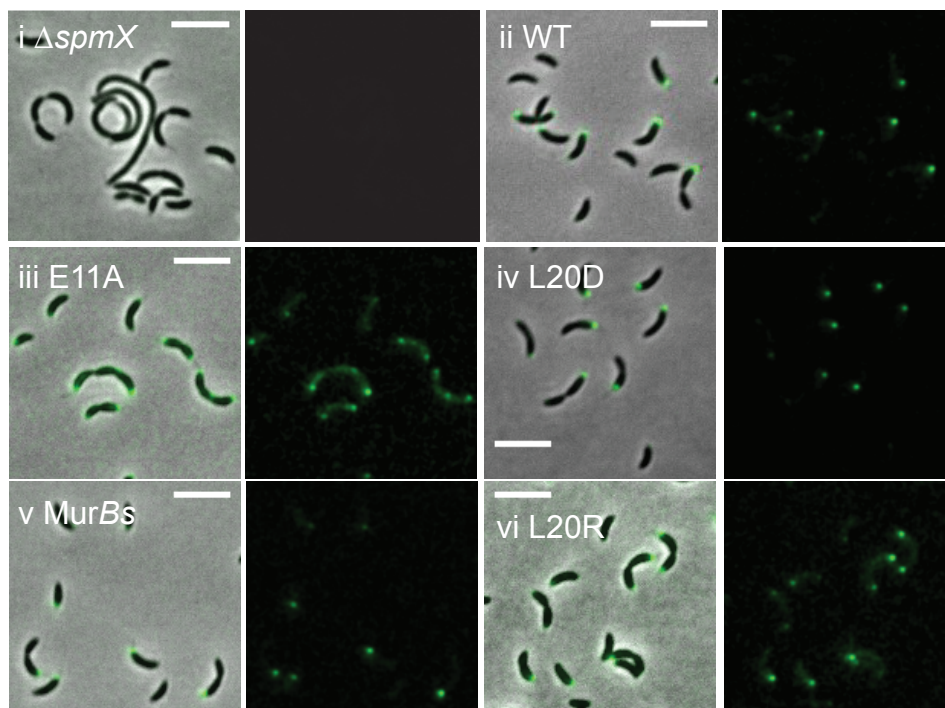








A



B

