Title

Antibacterial potency of Type VI amidase effector toxins is dependent on substrate topology and cellular context

Authors

Atanas Radkov¹, Anne L. Sapiro¹, Sebastian Flores², Corey Henderson³, Hayden Saunders¹, Rachel Kim⁴, Steven Massa⁵, Samuel Thompson⁶, Chase Mateusiak⁷, Jacob Biboy⁸, Ziyi Zhao¹, Lea M. Starita^{9,10}, William Hatleberg, Waldemar Vollmer⁸, Alistair B. Russell¹¹, Jean-Pierre Simorre¹², Spencer Anthony-Cahill¹³, Peter Brzovic¹⁴, Beth Hayes¹, Seemay Chou^{1*}

1 Department of Biochemistry & Biophysics, University of California – San Francisco, San Francisco, CA

2 University of Miami, Miami, FL

3 InBios International, Seattle, WA

4 Pacific Northwest University of Health Sciences, Yakima, WA

5 Department of Biology, Stanford University, Stanford, CA

6 Department of Bioengineering, Stanford University, Stanford, CA

7 Computer Science Department, University of Washington in St Louis, St. Louis, MO

8 Center for Bacterial Cell Biology, Newcastle University, Newcastle upon Tyne, United Kingdom

9 Department of Genome Sciences, University of Washington, Seattle, WA

10 Brotman Baty Institute for Precision Medicine, Seattle, WA

11 Division of Biological Sciences, University of California, San Diego, La Jolla, CA

12 Université Grenoble Alpes, CNRS, CEA, IBS, Grenoble, France

13 Chemistry Department, Western Washington University, Bellingham, WA

14 Department of Biochemistry, University of Washington, Seattle, WA

*Corresponding author: seemay.chou@ucsf.edu

Abstract

Members of the bacterial I6SS amidase effector (Tae) superfamily of toxins are delivered between competing bacteria to degrade cell wall peptidoglycan. Although Taes share a common substrate, they exhibit distinct antimicrobial potency across different competitor species. To investigate the molecular basis governing these differences, we quantitatively defined the functional determinants of Tae1 from *Pseudomonas aeruginosa* PAO1 using a combination of nuclear magnetic resonance (NMR) and a high-throughput *in vivo* genetic approach called deep mutational scanning (DMS). As expected, combined analyses confirmed the role of critical residues near the Tae1 catalytic center. Unexpectedly, DMS revealed substantial contributions to enzymatic activity from a much larger, ring-like functional hot spot extending around the entire circumference of the enzyme. Comparative DMS across distinct growth conditions highlighted how functional contribution of different surfaces is highly context-dependent, varying alongside composition of targeted cell walls. These observations suggest that Tae1 engages with the intact cell wall network through a more distributed three-dimensional interaction interface than previously appreciated, providing an explanation for observed differences in antimicrobial potency across divergent Gram-negative competitors. Further binding studies of several Tae1 variants with their cognate immunity protein demonstrate that requirements to maintain protection from Tae activity may be a significant constraint on the mutational landscape of *tae1* toxicity in the wild. In total, our work reveals that Tae diversification has likely been shaped by multiple independent pressures to maintain interactions with binding partners that vary across bacterial species and conditions.

Introduction

Bacteria live in dense polymicrobial communities where competition for nutrients and space impacts survival¹. To combat neighboring rivals, some Gram-negative bacteria use a contact-dependent toxin delivery system called the Type 6 secretion system (T6SS)^{2,3} that injects a cocktail of effectors from the donor cell directly into nearby recipient cells⁴⁻⁶. A bacterium's T6S arsenals include several types of protein effectors with distinct activities that each compromise essential cellular pathways or components, such as recipient peptidoglycan (PG) or membranes. Although many T6S bacteria deploy similar toxins, interbacterial outcomes can vary considerably depending on the bacterial species engaged in T6S-mediated competition^{7,8}. The molecular basis of this organism-level specificity of interbacterial T6SSs remains poorly understood.

A potential source of specificity could be the T6S-delivered toxins themselves⁹. T6S toxins can belong to large superfamilies comprised of diverse homologs. One such example is the T6S amidase effector (*tae*) superfamily⁷. The *taes* encode hydrolase enzymes that induce recipient cell lysis by breaking down the major structural component of bacterial cell envelopes, PG⁵. *Tae* toxins are known to be structurally and biochemically diverse¹⁰. Given that targeted PG substrates can also vary in a species-dependent manner¹¹, such diversity could potentially impact toxin antibacterial specificity. *Tae* toxins divide into four phylogenetically distinct subfamilies (*Tae*1-4)⁷ that hydrolyze different amide bonds within PG, and toxins within the same subfamily can have unique PG compositional preferences.¹⁰

These observations raise the possibility that T6SS organismal specificity may be linked, in part, to the biochemical diversity of delivered toxins. Such a model would predict that physiological features of recipient microbes may have driven divergent evolution of T6S effector diversity. To address this hypothesis, a quantitative, molecular understanding of T6S effector specificity is needed. However, comprehensive functional analyses are challenging for toxins such as the *Taes*, given that they act on cellular substrates with complex, three-dimensional structures *in vivo*. In cells, PG is a cross-linked macromolecule that forms a contiguous, mesh-like network that is structurally heterogeneous, insoluble, and dynamically regulated^{12,13}. These features are not well captured through *in vitro* functional studies that necessarily focus on soluble PG fragments.

To probe *Tae*–cell wall interactions in a physiologically relevant context, we developed a high throughput genetic approach known as deep mutational scanning (DMS)¹⁶ that surveys the entire *Tae* enzyme for functional determinants of toxicity in live *Escherichia coli* cells. We focused on *Tae*1 from the common soil bacterium *Pseudomonas aeruginosa*, which is known to interact with a wide range of microbes, including *E. coli*. We found that an extended surface across the entire *Tae*1 enzyme is important for mediating *in vivo* degradation of cell walls. Furthermore, we present evidence that *Tae*1 functional determinants can be quite distinct depending on the compositional context of PG. Finally, we discovered that the evolutionary potential for *Tae*1 potency against a broad range of cell wall types may be constrained by orthogonal functional requirements, such as *Tae*1 interactions with its cognate immunity protein during T6S-mediated competition.

RESULTS

Deep mutational scanning identifies determinants of Tae1 function *in vivo*

To understand how sequence diversity within the Tae protein superfamily impacts interbacterial toxicity, we must first probe how each amino acid residue contributes to Tae function. Biochemical assays of Tae protein binding to and degradation of peptidoglycan (PG) use isolated, purified sacculi¹⁰, which differ from the turgor-stretched PG layer in the periplasm of living cells. Thus, these assays likely identify only a subset of determinants that govern how Tae engages with PG *in vivo*. To define comprehensively and quantitatively all the amino acid residues in Tae that are important for its function *in vivo*, we developed a DMS experiment in live bacterial cells that allows us to examine the contributions of amino acid substitutions at every Tae residue in a physiologically relevant context.^{14,15}

We adapted a DMS approach^{16,17} to study the Tae1 protein from *P. aeruginosa*, which is injected into *E. coli* and leads to cell lysis (Figure 1A). We based our screen on a lysis assay that measures the toxic effects of Tae1 on *E. coli* when it is ectopically expressed in the periplasm, which contains the PG¹⁰. Periplasmic expression of wild-type Tae1 (Tae1^{WT}), but not catalytically inactive Cys30Ala Tae (Tae1^{C30A}), results in rapid cell lysis upon induction due to osmotic stress resulting cell wall degradation by the toxin. Cells expressing *tae1*^{WT} are depleted from the population faster than cells expressing inactive Tae1 (Figure 1B). Mutations that increase or decrease the rate of Tae1-induced lysis can be identified by sequencing and analyzing plasmid copy numbers in the population over time.

We generated a barcoded *tae1* plasmid library for massively parallel screening in *E. coli* (Figure 1C). Mutations were introduced by error-prone PCR, and library diversity was characterized by sequencing. We focused our analysis on copies of *tae1* that had a single mutation. In total, there were 902 different Tae1 variants, which encompassed at least one amino acid substitution at every position in the 154-residue protein. To ensure Tae1 delivery to the bacterial cell wall, we

expressed the variants as fusion proteins with N-terminal periplasm-targeting *pelB* signal sequences. We transformed *E. coli* (BW25113) grown in M9 minimal media with our *tae1* library and sequenced plasmid barcodes, both before and two hours after library induction (**Figure 1C**). *Tae1*-dependent lytic effects were quantified by using the DMS analysis pipeline Enrich2¹⁷, which calculates output-to-input ratios for individual variants, yielding fitness scores that are normalized to wild type (**Supplemental Table 1**). In our assay, the fitness score reflects a difference in the antibacterial capacity of the *Tae1* variant *in vivo* when compared to the wild type: relative enrichment of a variant indicates reduced *Tae1*-dependent cell death or loss-of-function (LOF) and yields a positive fitness score, whereas relative depletion of a variant indicates increased cell death or gain-of-function (GOF) and yields a negative fitness score.

To validate our approach, we examined the canonical catalytic dyad residues of *Tae1* to assess whether our screen accurately identified known functional determinants. Any substitutions of these catalytic residues, Cys30 and His91, should inhibit enzyme activity and lead to LOF. As anticipated, all non-synonymous amino acid changes at these sites resulted in increased fitness scores, indicating their crucial roles in *Tae1* toxicity (**Figure 1C and D**) and confirming that the DMS screen was indeed selecting *in vivo* functional determinants of *Tae* PG-degrading toxicity.

DMS identifies functional hotspots extending around the surface of *Tae1*

To obtain a global view of the amino acid residues in *Tae1* that affect fitness in our screen, we generated an averaged sequence-function map of variant fitness scores (**Figure 2A**). The majority of all unique *Tae1* protein variants (54%) were LOF variants with positive fitness scores, whereas only 6% were GOF variants. To determine the functional role of each residue of *Tae1*, we collapsed and averaged scores across all substitutions at each *Tae1* site (**Figure 2A**) and superposed these onto the three-dimensional structure of *Tae1* (**Supplemental Figure 1**). For some *Tae1* residues, position averages obscured how the degree or direction of functional change depended greatly on the specific amino acid change introduced. Therefore, we used *k*-means clustering, a statistical method for analyzing DMS data¹⁸, to classify *Tae1* residues in an unbiased way into one of four categories: residues where mutations generally decreased cell lysis (LOF_{clus}), those where mutations generally increased cell lysis (GOF_{clus}), those where mutations had variable effects on lysis (mixed), and those where mutations had little to no effect (tolerant). The vast majority of residues were categorized as

either LOF_{clus} (56%) or mixed (36%), and fewer were categorized as GOF_{clus} (5%) or tolerant (3%) (**Figure 2B**). We mapped these classifications onto the Tae1 structure and found, consistent with other mutational studies, a large majority of buried residues (78%) including the catalytic residues, were categorized as LOF_{clus} (**Figure 2C**). Residues on the enzyme surface, by contrast, were a mix of categories including most of the GOF_{clus} residues.

We anticipated that only surface residues within or near to the catalytic cleft would show loss-of-function phenotypes. Instead, we found that many surface-exposed residues more distant from the catalytic cleft were also categorized as LOF_{clus} (**Figure 2D**). Indeed, a zone of LOF_{clus} residues forms a ring around the entire enzyme, suggesting a more distributed set of surfaces may play key roles in cell lysis *in vivo* than was predicted by *in vitro* studies¹⁰. Many of the surface LOF_{clus} residues that our screen indicates are crucial for cell lysis *in vivo* are poorly conserved across the *tae1* superfamily¹⁰. We posit that while these mutations are LOF when applied to *E. coli* sacculi, divergent bacterial physiologies *in situ* may instead drive diversification of this extended surface.

Tae1 surfaces distal from catalytic core mediate binding and hydrolysis of cell wall network

Our DMS results suggested surfaces outside the Tae1 catalytic core are functionally important. We observed that several of these distal surface LOF_{clus} sites encoded serine and glycine residues, which have been previously implicated in sugar-binding^{19,20}. Thus, it is possible that the distal LOF_{clus} surfaces are important for interacting with the glycan chains of PG. Alternatively, these distal surfaces could mediate Tae1 interactions with surrounding, untargeted peptide stems within the PG network. To validate that these residues play a role in Tae1 function, we chose nine LOF_{clus} variants distributed across the Tae1 surface (**Figure 3A**) and manually assayed their individual effects through *E. coli* lysis assays (**Figure 3B and Supplemental Figure 2**). We focused mainly on serine and glycine residues and introduced the most conservative substitutions represented in the screen (Alanines or Serines, respectively).

Consistent with DMS results, all nine variants showed significantly less lytic activity than Tae1^{WT} (Figure 3B).

LOF phenotypes could stem from a variety of mechanisms, including reduced Tae1 protein expression, disruption of protein folding or stability, or direct inhibition of Tae1 PG binding or hydrolysis. To rule out expression and possibly stability-dependent effects, we performed Western blot analyses to measure protein levels of Tae1 variants expressed in *E. coli* (Figure 3B). We found that all variants were expressed; five substitutions did not exhibit major visible decreases in protein levels (G42S, S144A, S147A, G48S, G56S). Our results led us to hypothesize that these five surface substitutions may play direct roles in Tae1–PG interactions.

We hypothesized that an extended surface of Tae1 may be required for substrate binding and hydrolysis *in vivo*. Given that Tae1 likely engages with larger and even sometimes insoluble fragments of PG within the intact cell wall, we conducted a series of *in vitro* biochemical assays of LOF^{clus} variants using either intact PG sacculi or fragments that encompass more than the minimal peptide region. To address whether distal surfaces of Tae1 bind PG, we experimentally mapped Tae1 binding sites by two-dimensional (2D) nuclear magnetic resonance (NMR) using ¹⁵N-labeled, catalytically inactive Tae1^{C30A} (Tae1^{N15}) in complex with larger, soluble fragments of PG. Backbone ¹³C, ¹⁵N, and ¹H NMR resonances of Tae1 (>89% complete) were assigned using conventional heteronuclear techniques²¹. To prepare larger PG fragments for binding experiments, we digested *E. coli* cell wall sacculi using another Tae superfamily member (Tae3 from *Salmonella* Typhi)⁷, which hydrolyzes peptide crosslinks resulting in long multimeric PG fragments that include both Tae cleavage sites and intact glycan chains (Figure 3C) with an average length range of 10Å to over 300Å (1 to over 30 dimeric units)²².

To map enzyme residues that bind cell wall substrate, we compared NMR spectra of Tae1^{N15} alone or in the presence of digested multimeric PG (Supplemental Figure 3). Observed spectral differences allowed us to identify and quantitate chemical shift perturbations (CSPs) across the protein, which serve as indicators of residues involved in direct substrate interactions^{19,20}. Mapping these shifts onto the Tae1 structure, we observed CSPs for residues located within the catalytic core as well as for a variety of both buried and exposed residues outside of the core (Figure 3D). While CSPs for buried residues could be due to interaction-dependent enzyme conformational or dynamical changes²³, contiguous clusters of CSPs localized to the protein surface are suggestive of Tae1–PG binding. Such putative binding regions extended all the way to surfaces on the protein face opposite of the catalytic core. In fact, we identified shifts in six of the validated LOF^{CLUS} residues: Leu21, Gly42, Gly48, Ala52, Gly56, and Ser144 (Figure 3E), suggesting that these residues are necessary for Tae1-induced lysis because they bind PG.

To determine if these functional distal residues that bind PG are also required for hydrolysis, we focused our studies on two highly expressed variants, Tae1^{S144A} and Tae1^{G48S}, for which we expressed recombinant forms and purified. Size exclusion chromatography analysis of these variants showed similar elution volumes and apparent hydrodynamic radii compared to Tae1^{WT}. We then conducted degradative assays with PG sacculi purified from *E. coli*. Tae enzymes were incubated with PG sacculi, and the resulting fragments were quantitatively assessed using high-performance liquid chromatography (HPLC, **Figure 3F**). Tae1^{S144A} and Tae1^{G48S} cleaved 29.1 +/- 3.2% and 36.4 +/- 4.2% less PG, respectively, than observed for the wild-type enzyme. Despite being relatively far from the active site, these amino acid substitutions decreased Tae1 activity. Together, these data reveal that numerous sites distributed around the Tae1 surface physically interact with PG, and that these distal interaction sites are important for cell wall-degrading toxicity *in vivo*. Moreover, these results suggest our DMS-based approach can faithfully uncover novel and unexpected functional determinants *in vivo*, providing important clues to evolutionary pressures that may be acting on *tae1*.

The functional landscape of Tae1 is sensitive to PG differences

Our results point to a model in which Tae1 might interact with more structures of PG outside of the catalytic site, and these interactions are important for *in vivo* toxicity. This model predicts that Tae enzymes are sensitive to differences in cell wall architecture, and broad changes to PG composition

would alter Tae1–PG interactions. We hypothesized that such alterations could be captured by comparative DMS analysis across different screening conditions. To address this hypothesis, we conducted a second DMS screen in *E. coli* grown in the presence of D-methionine (D-Met), a condition known for inducing major featural changes in cell wall composition and organization (**Figure 4A**). Specifically, cell walls of *E. coli* grown in the presence of D-Met (*E. coli*^{DMet}) have generally reduced PG density and fewer modified PG crosslinks relative to cell walls of *E. coli* grown in the absence of D-Met^{24,25}. We verified these trends by HPLC analysis, and we confirmed that Tae1 maintained its ability to cleave this distinct cell wall type *in vitro* (**Supplementary Figure 4**).

We performed the same suite of DMS analyses for the second screen as for the first (**Figure 4B**) and found that the calculated fitness landscape shifted dramatically. Approximately 50% fewer residues were categorized as LOF_{clus} (38/86), while the number of mixed residues increased (94/56) as did the number of GOF_{clus} residues (18/8). The number of tolerant residues did not change (4 in both) (**Figure 4C**). Overall, 58% of residues were classified differently than in our initial screen in M9 media. We mapped the LOF_{clus} and GOF_{clus} residues onto the Tae1 structure (**Figure 4D**). We also saw that the previous ring-like band of LOF_{clus} determinants emanating from the catalytic groove was greatly diminished in *E. coli*^{DMet}. Given that our biochemical experiments suggest these sites are important for extended PG-binding, we hypothesize that such contact interfaces may be less functionally critical in cellular conditions where the cell wall matrix is more loosely packed around the enzyme. Alternatively, these changes could be indirectly due to other metabolic or physical changes to the cell under D-Met growth conditions. Overall, our comparative screening approach revealed that compositional changes to PG dramatically alter the functional landscape of Tae1, suggesting these variants are sensitive to differences in cell wall architecture.

The *E. coli*^{DMet} DMS screen also yielded a larger and distinct set of top-performing GOF_{clus} variants with amino acid substitutions that localized to both the catalytic groove and opposing faces of Tae1. When we compared fitness scores for the GOF_{clus} variants from the D-Met screen to the initial M9 screen, we found that only a small subset of substitutions showed similar effects across both conditions (**Figure 4E**). Condition-specific phenotypes suggests that Tae1 toxicity *in vivo* is highly context-dependent and optimizable for specific PG forms. To test this hypothesis, we conducted lysis assays in D-Met and M9 media for three GOF variants uniquely identified in the D-Met screen (**Figure 4F and G**). All three substitutions increased cell lysis in *E. coli* grown in D-Met but not M9 minimal media. For two of the variants, Asp134Asn and Ser147Ile, the increase in lysis may be due to higher expression levels based on Western blot analysis (**Figure 4F and G**). Together, these results suggest Tae enzymes could possibly be engineered to target specific bacteria with certain cell wall structures. Given that the structure and composition of PG varies across bacterial species, our results also potentially point to a selective pressure underlying natural sequence-based diversity in the Tae1 superfamily. Namely, cell wall features associated with particular rival species a given bacterium must antagonize may drive evolved substrate specificity differences in interbacterial competition toxins such as the Tae enzymes.

A hyperactive variant of Tae1 is naturally encoded by other *tae* homologs

A number of context-independent Tae1 variants that significantly out-performed wild-type enzyme across both DMS conditions were also identified through our comparative approach. We were

particularly struck by one variant, Tae1^{C110S}, which is modified at a partially buried residue sitting directly behind the catalytic Cys30 site (**Figure 5A**). The Cys110Ser substitution was the top-scoring GOF variant in the D-Met screen and ranked 11th amongst GOF variants in the M9 screen. Using the *E. coli* lysis assay, we first confirmed that Tae1^{C110S} indeed lysed cells at a faster rate than Tae1^{WT} in cells grown in both D-Met and M9 minimal media (**Figure 5B**). To rule out indirect effects on toxicity caused by increased protein levels, we assessed Tae1^{C110S} expression by Western blot analysis and did not observe any increase relative to Tae1^{WT} (**Figure 5B**).

The observation that Tae1 activity could generally be tuned for greater toxicity led us to ask why *tae1* from *P. aeruginosa* does not naturally encode Ser110. We considered several explanations, none of which are mutually exclusive. First, it is possible that structural rearrangements introduced by the Cys110Ser substitution may disrupt Tae1 interactions with critical molecular partners present in natural T6S-mediated interbacterial interactions but not in our DMS set-up. During T6S competition, Tae1 from *P. aeruginosa* must be delivered to neighboring cells by a complex T6S secretion apparatus that we bypassed with an engineered secretion signal in our screen. Furthermore, Tae1 injected into nearby kin cells is deactivated through binding by a cognate immunity protein, T6S amidase immunity protein 1 (Tai1), preventing lysis. Tai1 is specifically encoded by *P. aeruginosa* and not *E. coli*. Second, it is possible that we have not sampled Tae1 fitness across contexts that accurately represent bacterial rivals or competition conditions most relevant for *P. aeruginosa* in the wild.

We therefore investigated whether Ser110 naturally occurs in any Tae1 representatives outside of *P. aeruginosa* by searching for family members that encode a serine at residue positions comparable to Cys110, as predicted by structure-based alignments. Surprisingly, we found that approximately half of *tae1*-carrying bacterial species encoded serine residues at the Cys110 position, while some encoded tryptophan (**Figure 5C**). To examine how these three variant classes are taxonomically distributed, we generated a phylogenetic tree of all associated bacterial species and observed correlated divisions across taxonomic clades and the residues at the Cys110 position, with the exception of one *Pseudomonas* species (**Figure 5C**). Considering this observation, we hypothesized that *tae1* evolution may have been locally constrained due to the requirement for Tai1-dependent immunity to prevent self-antagonism.

Figure 5. A hyperactive variant of Tae1 is naturally encoded by other *tae1* homologs. (A) Ribbon structure diagram of the Tae1 catalytic site. Catalytic dyad residues Cys30 and His91 are indicated along with GOF Cys110. (B) Lysis assays with wild-type Tae1, empty vector and catalytically inactive Tae1^{C30A} as negative controls, and gain-of-function variant Tae1^{C110S} in D-Met and M9 media. Data are means $\pm$ standard deviation, $n = 3$, * $p < 0.05$. Induced with 0.0025% arabinose. Lysis experiment was repeated 3 independent times, each time with at least 3 technical replicates. Western blots below measure Tae1 protein expression with RNAP as a control. The C110S GOF substitution does not sensitize to the D-Met induced changes in cell wall architecture. (C) Tae1 family phylogenetic tree. Species with Ser at position 110 are indicated in blue, while those with Cys are in black and those with Trp are in purple. Some but not all Tae1 proteins have evolved to include a Ser at position 110.

Hyperactive variant Tae1^{C110S} evades binding and inhibition by cognate immunity

We reasoned that if Tae1 proteins are too structurally divergent, the toxin may evade binding by the cognate Tai1 partner in *P. aeruginosa* (Tai1^{PA}) and consequently antagonize kin cells or other closely related species. To assess whether the Tae1^{C110S} variant affects Tae1–Tai1 interactions, we first assessed whether Tae1^{C110S} adopted any major structural changes by solving its three-dimensional crystal structure (PDB ID: 7TVH) (Table 2). We did not observe any major structural rearrangements that would immediately implicate distinct modes of Tai1^{PA}-binding (RMSD=0.746, Figure 6A). However, we did observe some key side chain alterations near the Tae1^{C110S} active site, which is the enzyme region most intimately engaged with Tai1^{PA} upon binding (Supplemental Figure 5). Most notably, we found that His91 protruded away from the active site of Tae1^{C110S}, which would sterically clash with a loop of Tai1^{PA} that typically inserts into the catalytic cleft (Figure 6B). Our structural observations are consistent with the hypothesis that the Cys110 position is important for Tai1 recognition; therefore, serine at this position in *P. aeruginosa* may be disfavored to preserve immunity to the toxin. We generated a tree for the Tai1 protein family, similar to our tree for the Tae1 protein family. Consistent with this possibility, the tree for Tai1 has a similar branching pattern that is coincident with the residue variation at this position (Supplemental Figure 6).

We tested our prediction through a series of quantitative *in vitro* Tae1–Tai1 binding experiments. We measured Tai1-binding affinity (K_d) for recombinant Tae1^{WT} and Tae1^{C110S} proteins by isothermal titration calorimetry and found that Tae1^{WT} bound Tai1 with ~36-fold greater affinity than Tae1^{C110S} (9.5 nM K_d and 349.6 nM K_d, respectively) (Figure 6C). Toxin–immunity complexes are typically associated with binding affinities in the nM range; thus,

reduction may preclude the ability of Tai1 to effectively bind its toxin partner *in vivo*. To test whether Tae1^{C110S} evades Tai1-dependent immunity *in vivo*, we constructed a strain of *P. aeruginosa* harboring this mutation (PA^{C110S}) with a wild-type strain (PA^{WT}). We reasoned that if Tae1^{C110S} escapes Tai1-binding and therefore immune protection, we would observe a decrease in PA^{C110S} fitness compared to a wild-type strain (PA^{WT}) specifically when grown on solid media which induces T6S (Figure 6D). Indeed, for PA^{C110S} we observed substantially reduced growth compared to PA^{WT} on solid media (Figure 6E) but not in liquid media, where T6S is not induced (Supplemental Figure 7). This phenotype strongly suggests that Tae1^{C110S} lyses kin cells at a faster rate than Tae1^{WT} can bind and inhibit killing, leading to a fitness cost for this strain.

Together, our *in vitro* and *in vivo* results suggest that the fitness cost of evading Tai1-binding likely limits the proliferation of Tae1^{C110S} in the wild. This interaction presents an evolutionary bottleneck that impedes certain toxin variants, such as Tae1^{C110S}, from broadly taking hold across all species. More generally, our DMS screening approach allowed us to systematically probe function across different *in vivo* contexts, identifying two distinct but important selective pressures – substrate-specificity and toxin-immunity interactions – that may, in combination, shape diversity across the Tae1 toxin superfamily.

DISCUSSION

The Tae superfamily of interbacterial toxins is phylogenetically, structurally, and functionally diverse^{7,10}. Here, we show that sequence-based differences in *tae* homologs shape Tae proteins and their complex substrate, the PG that constitutes the bacterial cell wall. Our studies revealed that Tae1 engages PG through a circumferential interface that is more extensive than most proteases or other enzymes. This unexpectedly large interaction surface may be key for dictating species-specific toxicity of Tae proteins and consequently serve as a major driver of Tae superfamily diversification. This surface could also mediate processivity across PG. Taes have many other binding partners as well. These toxins interact with the T6SS apparatus for secretion, which likely limits effector size to a single, minimal catalytic unit^{4,26}. Therefore, unlike other larger multi-domain housekeeping PG hydrolases²⁷, Tae toxins rely on a relatively restricted set of surfaces to mediate many distinct, functional interactions. Our unbiased, high throughput DMS analyses revealed that determinants for Tae1 cell wall-binding and immunity-binding partially overlap. Our findings suggest that accommodation of multiple, distinct functions likely constrains evolution of Tae enzymes in nature despite intrinsic potential for enhanced toxicity. Extensive Tae functional surfaces may also point to the existence of additional unknown interaction partners within bacteria. This principle may prove true for other T6S

toxins and proteins; comprehensive mapping of functional determinants by DMS may power further functional discovery.

Complementing traditional biophysical approaches with DMS allowed us to probe molecular interactions in physiological states that are difficult to capture *in vitro*. Our current understanding of enzyme interactions with the bacterial cell wall draws mostly from biochemical analyses conducted in solution, posing a fundamental challenge for proteins with three-dimensional highly-crosslinked macromolecular substrates. Considering the architectural complexity of the bacterial cell wall, a large ring-like surface looping around Tae1 seems critical for toxicity *in vivo*. This observation could be due to at least two possible modes of binding that we posit may be unique to enzymes with macromolecular substrates. First, we theorize that a more extensive, non-linear substrate binding domain may enable an enzyme to specifically recognize a particular conformational state. Three-dimensional configurations of the cell wall may be an important feature for a toxin like Tae1 to detect, given that this could affect protein localization to specific cell wall features that have distinct shapes. It has been previously observed that a critical hole size in the cell wall is necessary for bacterial lysis²⁸; thus localized action of PG enzymes may provide an important advantage for Tae toxins. Another potential advantage of binding a specific substrate configuration is that this would enable a Tae enzyme to distinguish between different species or incorporation states of PG fragments. Indeed, certain endogenous enzymes, such as FtsN interact exclusively with complex PG chains of a certain length²⁹.

Another possibility is that the large functional surface comprises several different PG-binding domains that streamline substrate recognition or underlie enzyme processivity. This has been suggested for a cell wall digesting enzyme, Lysostaphin from *Staphylococcus*, which exhibits two large PG-binding surfaces on opposite sides of the enzyme²³. Multi-step binding would support the “smart autolysins” model postulated by Arthur Koch³⁰. This model hints at the ability of hydrolase enzymes to recognize “stretched” areas of PG. The resulting advantage for hydrolytic cell wall enzymes is that this type of recognition decreases substrate search dimensionality from random 3-D diffusion to facilitated diffusion by searching along the PG sacculus one plane at a time³¹. This may be particularly important for rapid, effective recognition of substrates within a large cage-like 3D target like the cell wall, which is key for potent, fast-acting toxins. This mechanism is also commonly adopted by DNA-binding proteins, which dock nonspecifically to DNA before searching for specific cut sites^{32,33}. Multi-step binding may also lead to processive activity along PG by ensuring that Tae enzymes remain tethered to the cell wall substrate after each hydrolysis event, which would allow for “stepwise chewing.” This would be yet another mechanism to increase the odds of localized, lytic activity.

In total, our combined use of biophysical and high throughput genetic approaches uncovered novel insights into Tae toxin activity inside bacterial cells, providing a window into how small, single-domain enzymes carry out several, distinct functions *in vivo*. Our leveraging of a multi-pronged strategy was critical for overcoming technical obstacles inherent to proteins that interact with complex, insoluble molecules, providing access to a class of macromolecular interactions that are not well understood. The biophysical and kinetic parameters of proteins with three-dimensional highly-crosslinked macromolecular substrates^{34,35} are dissimilar from those well-established for small, soluble substrates, and future adoption of combinatorial approaches such as ours may improve our understanding of general binding principles across these important enzymes that are ubiquitous across life. Additionally, the ability to quickly generate rich quantitative heatmaps for enzymes across different experimental conditions enables more sophisticated computational analyses. Predictive representation learning of unbiased, functional datasets across multiple biological conditions could help us anticipate *a priori* the functional consequences of sequence-based phylogenetic diversity. A deeper understanding of functional design principles would also enable more effective leveraging of evolutionary innovations to engineer new molecules tailored to physiological contexts with important, useful applications.

ACKNOWLEDGEMENTS

We would like to thank Dr Doug Fowler (University of Washington) and Dr Eric Chow (University of California – San Francisco) for their help with setting up the deep mutational scanning pipeline. We are also grateful to Dr David Mavor (University of Massachusetts Chan Medical School), Dr Robert Newberry (University of Texas – Austin), Dr Martin Kampmann (University of California – San Francisco), and Dr James Fraser (University of California – San Francisco) for their assistance with sequencing data analysis. We say thanks to Dr Kari Herrington and Dr SoYeon Kim (University of California – San Francisco) for their contribution with microscopy data collection and analysis. We also thank Dr George Meigs and Dr James Holton (Lawrence Berkeley National Lab), and Dr Michael Thompson (University of California – Merced), for their help with x-ray crystallography data collection and analysis. We are grateful to Dr Mark Kelly for his help with setting up NMR data collection (University of California – San Francisco). We also thank to Dr Parnian Lak (University of California – San Francisco) for her help with setting up the ITC experiments. Lastly, we greatly appreciate the input regarding the experiments and the manuscript from members of the Chou lab at the University of California – San Francisco (Kristine Trotta, Calla Martyn, Fauna Yarza, Elizabeth A. Martinez-Bond, Patrick Grimes, Dr Domokos Lauko). Anne Sapiro received a Life Sciences Research Foundation fellowship from the SVCF-Wave Fund. Financial support from the IR-RMN-THC FR3050 CNRS for conducting the research on the 950 MHz NMR spectrometer is gratefully acknowledged. We would also like to thank the Chan Zuckerberg Biohub as well as the Sanghvi-Agarwal Innovation Award for their financial support.

COMPETING INTERESTS

The authors declare no competing interests.

METHODS

Preparation of long chains of PG

PG isolation and HPLC analysis were conducted according to an established protocol³⁷ with the following modifications. Briefly, *E. coli* BW25113 cells were grown in 200ml LB media to mid-log phase. The cells were harvested by centrifugation at 5000 x g for 15min, after which they were suspended in 300μl of fresh LB media and added to 600μl 6% SDS solution in a 2ml Eppendorf tube. The cells were boiled for 3hrs in a heat block and then maintained in the SDS solution at 4°C until needed. The boiled sacculi were washed 5 times with MQ water to remove the SDS. Each wash consisted of 15min centrifugation at 21000 x g followed by decanting of the supernatant and resuspending in 1ml of MQ water by vortexing. After the final wash, the sacculi were suspended in Pronase E digestion buffer (10mM Tris-HCl pH 7.2 and 0.06% NaCl) and were treated for 2hrs at 60°C with activated Pronase E (100μg/ml final concentration, pre-treated for 2hrs at 60°C in the digestion buffer; supplier: VWR). The sacculi were incubated at 100°C for 10min to inactivate the Pronase E and were washed 3 times with 1ml of MQ water, with vortexing to resuspend. The sacculi were treated with purified Tae3 enzyme (10μM) overnight at 37°C in 20mM HEPES pH7.0 and 100mM NaCl buffer. The PG fragments were finally dialyzed using 500Da molecular weight cutoff tubing (Biotech CE Tubing, Spectra/Por) and dried on a Vacufuge concentrator to remove all liquid. They were stored at -80°C and were redissolved before each NMR experiment with minimal volume to yield the desired enzyme to substrate concentrations.

NMR experiments

NMR experiments for backbone assignments were performed at 25°C on a Bruker 950 MHz AVANCE (Institut de Biologie Structurale in Grenoble) or Varian INOVA spectrometers (600 and 800 MHz instruments at Pacific Northwest National Laboratories). NMR samples for determination of Tae1

resonance assignments consisted of 0.3 to 0.6 mM [^{13}C , ^{15}N]-Tae1 in citrate buffer pH 5.5 with 150 mM NaCl, 100 μM EDTA containing 5% D_2O . Assignment of backbone resonances was accomplished by analysis of standard triple-resonance experiments (^1H , ^{15}N)-HSQC, HNCA, HNCOCa, HNCACB, CBCA(CO)NH, and HNCO spectra²¹). Assignments of Tae1 resonances in 20mM HEPES or NaPi, 150 mM NaCl, 100 μM EDTA at pH 7.0 were determined by collecting a series of [^1H , ^{15}N]-HSQC spectra titrating the sample from pH 5.5 to 7.0. These assignments were confirmed by collecting an HNCO spectrum under the final conditions at pH 7.0. NMR data were processed and analyzed using NMRPipe³⁸ and NMRView^{39,40} (for the residue assignments, see **Supplemental Figure 8**). See **Supplemental Table 2** for the peak values and **Supplemental Table 3** for CSP ($\Delta\delta$) values.

All titration experiments were conducted at the Institut de Biologie Structurale in Grenoble or at the UCSF NMR facility in Genentech Hall (San Francisco). PG ligand was dialyzed against water using a float dialysis membrane device with a 100-500 Da cut-off (Spectrum Laboratories, Inc) and lyophilized before preparation of a concentrated stock solution in NMR buffer. Before aliquoting the samples or before lyophilization, the concentration of the stock solution was estimated by liquid-state NMR using the Eretic pulse sequence⁴¹ and a reference 1 mM sucrose sample in 90%:10% H_2O : D_2O . Interaction studies with ^{15}N -labelled Tae1 and PG substrates were monitored by superimposition of ^{15}N -BEST-TROSY spectra at 25°C for different substrate-to-protein ratio⁴². A 50 μM ^{15}N labeled Tae1^{C30A} was titrated with 500 and 1000 μM of purified muopeptide. CcpNmr2.4 was used to monitor protein CSP for every assigned amide resonance by superimposition of the ^{15}N -BEST-TROSY spectra. CSPs ($\Delta\delta$) were calculated on a per-residue basis for the highest substrate-to-protein ratio.

E. coli lysis assays

The pBAD vector was used to clone Tae1 variants of interest⁴³. The Tae1 wild-type was initially amplified from a previously published pET22::Tae1 construct¹⁰, together with the pelB signal sequence. All variants were constructed based on this pelB_pBAD::Tae1 template. See **Table 1** for all the primers used to clone Tae1 and the different variants. SOE-PCR was used to introduce the individual mutations, followed by restriction enzyme digestion and ligation. pBAD::Tae1 constructs were always transformed freshly into *E. coli* BW25113 chemically competent cells and a LB culture with antibiotic was immediately started after the transformation outgrowth period. The culture was diluted 1 to 100 on the following day into LB media in the presence of antibiotic selection. After reaching OD₆₀₀ of 0.2, protein expression was induced with arabinose (0.25% in Figure 3 and 0.0025% in Figure 4 and Figure 5) and cell lysis was monitored overnight on a Biotek plate reader.

DMS library construction and assay

pET22b vector was chosen for library construction. This vector contains the pelB signal sequence that would allow the expression of Tae1 variants in the cell periplasm. The preparation of variants was achieved through an error-prone PCR mutagenesis kit (GeneMorph II Random Mutagenesis Kit – Agilent Technologies), following the manufacturer's instructions in order to insure high proportion of single-mutant variants. The variants were cloned via restriction enzyme cloning. In order to track the behavior of variants, each one was barcoded with a unique DNA sequence. The barcodes were introduced immediately after the Tae1 coding sequence via Gibson assembly cloning (In-Fusion HD Cloning Kit – Takara Bio). Our Next-Generation sequencing strategy comprised two stages. First, the entire library was sequenced to determine the barcode corresponding to each variant. This was accomplished using 300bp forward and reverse reads on the MiSeq Illumina platform (MiSeq Reagent Kit v3). Using this information, a dictionary was prepared to demonstrate the relationships between variants and barcodes. The dictionary was prepared by following previously published scripts⁴⁴ with some modification (see the GitHub directory “ChouLab dictionary” for the steps and scripts). For each of the subsequent DMS experiments, 25bp reads on a NextSeq Illumina platform (conducted at Chan Zuckerberg BioHub) were used to sequence only the barcode region. For data analysis and DMS scores calculations, we used Enrich2 software¹⁷. The barcode dictionary was used to convert barcodes to variants. Please see the “Choulab DMS data analysis” GitHub directory for detailed information on data analysis in Enrich2. For the DMS experiments in this study, the DMS library was grown for 3hrs at 37°C in M9 minimal medium (or M9 supplemented with 20mM D-Methionine) before distributing across a 96-well plate. A set of 24 wells was combined for each experimental time point in order to obtain enough DNA for library amplification. Libraries were amplified for 9 cycles using Phusion high-fidelity polymerase to minimize any masking of differences in abundance. The libraries were gel purified, quantified using Qubit, and their concentrations were adjusted according to Illumina instructions before sequencing. The 0hrs time point (DMS library before any growth) was chosen as the input of the 96-well plate incubation. After growth to OD₆₀₀ 0.2, protein expression was induced with 1mM IPTG. Cells were collected 2hrs after induction that would serve as the output time point. All of the scripts used in the paper are available on GitHub at https://github.com/AtanasDRadkov/ChouLab_DMS. The data files are available on NCBI, Sequence Read Archive – PRJNA803461.

Clustering analysis

K-means clustering analysis was conducted using previously published scripts¹⁸ to group Tae1 positions into categories based on their fitness scores. We calculated spatial clusters by 1) sorting the fitness score vectors for each position 2) trimming the vectors to match vector lengths (discarding no-data values) 3) calculating a difference by subtracting the two sorted and trimmed vectors, and 4) calculating the distance – mean of the absolute value of the vector difference. The distance between a candidate position and a cluster of positions is calculated as the average of the distance between the candidate position and the three closest non-self positions in the cluster. Clustering was performed over 10 rounds following the initial seeded round, and convergence was confirmed by observing that five repetitions gave identical clusters. The results yielded five clusters, including a loss-of-function phenotype cluster which was the focus of the study. The same analysis was performed for both DMS experiments. See the GitHub directory “ChouLab clustering” for the steps and scripts. All of the scripts used in the paper are available on GitHub at https://github.com/AtanasDRadkov/ChouLab_DMS.

HPLC PG analysis

PG isolation and HPLC analysis were conducted as described³⁷ with the following modifications. Briefly, *E. coli* BW25113 cells were grown in 200ml LB (or LB plus 20mM D-Met) to mid-log phase. The cells were harvested by centrifugation at 5000 x g for 15min, after which they were suspended in 300μl fresh LB and added to 600μl of 6% SDS solution in a 2ml Eppendorf tube. The cells were boiled for 3hrs in a heat block and then maintained in the SDS solution at 4°C until needed. Before any treatments, the

boiled sacculi were washed 5 times with MQ water to remove the SDS. Each wash consisted of 15min centrifugation at 21000 x g followed by decanting of the supernatant and resuspending in 1ml MQ water by vortexing. After the final wash, the sacculi were suspended in Pronase E digestion buffer (10mM Tris-HCl pH 7.2 and 0.06% NaCl) and were treated for 2hrs at 60°C with Pronase E (100µg/ml final concentration, pre-treated for 2hrs at 60C in the digestion buffer; supplier: VWR). The sacculi were incubated at 100°C for 10min to inactivate the Pronase E. Enzymatic treatments with Tae1 variant enzymes were conducted at this point using 20mM HEPES pH7.0, 100mM NaCl buffer. Tae1 enzyme was added to 1µM final concentration and the reactions were incubated at room temperature for 10min, followed by 10min at 100°C to inactivate the enzyme. The sacculi were washed 3 times as described above. Finally the sacculi were solubilized using 25µg/ml final concentration of mutanolysin (Sigma-Aldrich) in 10mM Tris-HCl pH 7.2 and 0.06% NaCl after overnight digestion in a 37°C shaker. To prepare PG fragments for HPLC analysis, the mutanolysin treatment was incubated at 100°C for 10min, followed by centrifugation to remove any debris, alkalization with 0.5M borate buffer and reduction with sodium borohydride for 1hr at room temperature. The reduced PG fragments were acidified to pH 3-4 with 50% o-phosphoric acid and filtered (PVDF syringe filters, 0.22µm, 0.4mm - Tisch Scientific) before the HPLC injection. The HPLC instrument, method, and buffers were the same as described previously⁴⁵.

Protein purification

All Tae proteins, and Tai1 protein, used in this study were purified according to an established protocol¹⁰ using pET29b+ vector and 3hrs induction in a 37°C shaker with 1mM IPTG. Buffers for metal affinity chromatography were 20mM HEPES pH7, 100mM NaCl, 20mM imidazole and 20mM HEPES pH7, 100mM NaCl, 400mM imidazole. Buffer for size exclusion chromatography (SEC) was 20mM HEPES pH7, 100mM NaCl. The SEC was performed using an isocratic gradient on an AKTA pure instrument with a HiPrep 16/60 Sephacryl S-200 HR column. For isotopically labeled Tae1^{C30A}, the same plasmid constructs were used, but the expression was conducted in M9 minimal medium prepared with ¹⁵N-ammonium chloride (5g per liter media) (M9 minimal medium prepared according to Cold Spring Harbor Protocols). The cells were grown and induced as described above. The same purification procedure was used as described above.

Phylogenetic trees

Initial candidates of the Tae1 and Tai1 families as identified in Russell et al.⁷ were aligned and used as a seed in a PSI-BLAST search. Significant results were downloaded and aligned using MAFFT. Candidates containing the critical Cys-His dyad were kept, and other candidates discarded. Final candidates were re-aligned. Default parameters were used for all steps. Multisequence alignment (ClustalW algorithm) and phylogenetic trees (PhyML analysis) were prepared on GenomeNet (at www.genome.jp). See **Supplemental File 1** for all Tae1 proteins and **Supplemental File 2** for the Tai1 proteins.

P. aeruginosa strain construction

P. aeruginosa strains used in this study were derived from the sequenced strain PAO1⁴⁶. *E. coli* strain SM10 was used for conjugal transfer of pEXg2 plasmids for homologous recombination as described previously⁵. All deletions were in-frame and were generated by exchange with deletion alleles constructed by SOE PCR, followed by restriction enzyme cloning into pEXg2 vector. See **Table 1** for the primers used to delete genes *pppA* and *retS*, and to introduce the C110S mutation into *tae1*.

Contact-dependent fitness assays

Fitness assays were performed with overnight *P. aeruginosa* cultures. The assays were performed on solid media (LB without added NaCl containing 3% agar) using nitrocellulose membranes (Amersham Biosciences Hybond-N+, 0.45µm) to insure close contact between the cells⁵. The OD₆₀₀ was adjusted

to 1.0 for all strains. 10 μ l were spotted onto a membrane to initiate the contact-dependent fitness assays. The assays were incubated for 6 hrs at 37°C after which the spots were removed with a razor blade, placed in a tube with glass beads, shaken on a bead beater (BeadBug microtube homogenizer; Z763705, Sigma) for 20sec at 4000rpm, and the suspended cells were distributed across a 96-well plate (3 x 200 μ l wells for each competition tube) for OD₆₀₀ measurements on a Biotek plate reader.

Isothermal titration calorimetry

Isothermal titration calorimetry experiments were performed at 298°K using a Microcal ITC200 (GE Healthcare, Northampton, MA). Stock solutions of Tae1 and Tai1 were dialyzed overnight at 277°K in the same buffer (50mM NaCl, 20mM Tris 7.0). The cell was filled with 200 μ L of Tai1 protein at a concentration of 20 μ M to which 150 μ M of Tae1 protein was titrated. For 25 titrations of 1.5 μ L of Tae1 into Tai1 the heat of binding was recorded. Data were extracted and fitted using the single-site binding model following manufacturer's instructions.

Tae1^{C110S} crystal structure

Tae1^{C110S} protein crystals were generated by hanging drop vapor diffusion at 25°C from a 1:1 mixture of 10 mg/ml protein (in 20mM HEPES pH7, 100mM NaCl, 5% glycerol) with 0.2 M sodium thiocyanate, 0.1 M HEPES pH 7.5, 20% PEG3350 for 24hrs (same conditions that were used for initial structure of Tae1¹⁰). Crystals were directly used for diffraction data collection at the Lawrence Berkeley National Laboratory Advanced Light Source Beamline 8.3.1⁴⁷. Coot⁴⁸ and Maximum-Likelihood⁴⁹ refinement with PHENIX were used for iterative building and refinement. Structure was validated by MOLProbity⁵⁰ and visualized with Chimera⁵¹. Coordinates and structure factors were deposited in the Protein Data Bank (PDB ID: 7TVH).

Tables

Table1. Primers used in this study

Name	Sequence
pET22_Tae1_C110S_F1	CCGATGTGCTGGTCCGGCAGCATCGCCGGC
pET22_Tae1_C110S_F2	GGTGGTGGTGTCTGAG
pET22_Tae1_C110S_R1	GCCGGCGATGCTGCCGGACCAGCACATCGG
pET22_Tae1_C110S_R2	TGGATATCGGAATTAATTCGGATC
pET29_Tae1_C110S_F1	CCGATGTGCTGGTCCGGCAGCATCGCCGGC
pET29_Tae1_C110S_F2	GGTGGTGGTGTCTGAG
pET29_Tae1_C110S_R1	GCCGGCGATGCTGCCGGACCAGCACATCGG
pET29_Tae1_C110S_R2	GGGAATTCATATGGACAGTCTCGATCAATGC
pBAD_Tae1_F	TAGTACAGAGAATTCACCATGAAATACCTGCTGCCGACCGCTGCTGC
pBAD_Tae1_R	TCAATCAGTATCTAGATTACTCGAGACTGGCCCTGGGCAGGCTG
pEXG2_pppA_F1	TGTTAAGCTAAAGCTTGTCTGGTCTGCTATTTCCCGCTGAC
pEXG2_pppA_F2	GCATACCCGGTTAAGGACAACTGATGGTGAACCTCGAGCCGCAAGCATGCTGAA
pEXG2_pppA_R1	TCAATCAGTATCTAGAGCCGGTGAGGATCTCGTAGAGCAC
pEXG2_pppA_R2	CCGGAAGTTCTATGTCCATGTCCGTCTCAGAATTCAGCATGCTTGCGGCTCGAGTT
pEXG2_retS_F1	TCAATCAGTATCTAGACGGAAACCCAGCCGATCATGG
pEXG2_retS_F2	AACTCGAGCCGCAAGCATGCTGAAGGCGAAGTCCCTTCGAAGGTG
pEXG2_retS_R1	TGTTAAGCTAAAGCTTCCAGCATCTTCAGGTAGACGC
pEXG2_retS_R2	TTCAGCATGCTTGCGGCTCGAGTTGCCTACTGGGTCCGCGAACGC
pEXG2_Tae1C110S_F1	TCAAGTACTAGAGCTCGTGGTGCACCGCGAGGACATCTC
pEXG2_Tae1C110S_F2	CTGCCACTCCAGCACATCGGGTAC
pEXG2_Tae1C110S_R1	TATCAGAAACCTGCAGGTAGGCCAGCGACTGCATGCCGTAG
pEXG2_Tae1C110S_R2	GATGTGCTGGAGTGGCAGCATCGCC
pBAD_Tae1L21S_F	AGCTGGGACAAGAGCTACAGCGCCGGCACC
pBAD_Tae1L21S_R	CTTGTTCTGGGGTGCCGGCGCTGTAGCTCTTGTC
pBAD_Tae1G48S_F	GGCGTACCGATGCCCGCAGCAACGCCAACGCC
pBAD_Tae1G48S_R	GACCATGGCGTTGGCGTTGCTGCGGGGCATCGG
pBAD_Tae1G56S_F	GCCAACGCCATGGTCGACAGCCTGGAGCAGAGC
pBAD_Tae1G56S_R	GGTCCAGCTCTGCTCCAGGCTGTGACCATGGC
pBAD_Tae1G42S_F	TCGGTGGCCGCCGAGCTGAGCGTACCGATGCC
pBAD_Tae1G42S_R	GCCGCGGGGCATCGGTACGCTCAGCTCGGCGGC
pBAD_Tae1V43S_F	GTGGCCGCCGAGCTGGGCAGCCCGATGCCCGC
pBAD_Tae1V43S_R	GTTGCCGCGGGGCATCGGGCTGCCAGCTCGGC
pBAD_Tae1A52P_F	CCCCGCGCAACGCCAACCCGATGGTTCGACGGC
pBAD_Tae1A52P_R	CTCCAGGCCGTGACCATCGGGTTGGCGTTGCC
pBAD_Tae1S144A_F	CTCAACTACTACGTCTACGCCCTGGCCAGTTGC
pBAD_Tae1S144A_R	CAGGCTGCAACTGGCCAGGGCGTAGACGTAGTA
pBAD_Tae1S147A_F	TACGTCTACTCCCTGGCCGCTGCAGCCTGCC
pBAD_Tae1S147A_R	GGCCCTGGGCAGGCTGCAGGCGGCCAGGGAGTA
pBAD_Tae1S149A_F	TACTCCCTGGCCAGTTGCGCCCTGCCAGGGCC
pBAD_Tae1S149A_R	GAGACTGGCCCTGGGCAGGGCGCAACTGGCCAG
pBAD_Tae1A116T_F	TGCGGCAGCATCGCCGGCACTGTGCGCCAGAGC
pBAD_Tae1A116T_R	GCCCTGGCTCTGGCCGACAGTGCCGGCGATGCT
pBAD_Tae1D134N_F	CAGGTGTGGAATCGCACCAACCGCGACCGCCTC
pBAD_Tae1D134N_R	GTAGTTGAGGCGGTGCGGTTGGTGCATTCCA
pBAD_Tae1S147I_F	TACGTCTACTCCCTGGCCATTTGCAGCCTGCC
pBAD_Tae1S147I_R	GGCCCTGGGCAGGCTGCAAATGGCCAGGGAGTA
DMS_Lib_Amplify_F	AATGATACGGCGACCACCGAGATCTACACCTTTCTGGGCTTTGTTAGCAG
DMS_Lib_Amplify_R1	CAAGCAGAAGACGGCATACGAGATGGAACGGAGTGACTGGAGTTCAGACGTGTGC TCTTCCGATCTTCGATAGTAGCCTATGAAGCATGCTTACTA

DMS_Lib_Amplify_R2	CAAGCAGAAGACGGCATACGAGATCGGCACATGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTTCGATAGTAGCCTATGAAGCATGCTTACTA
DMS_Lib_Amplify_R3	CAAGCAGAAGACGGCATACGAGATTACCCTTGGTGAAGTTCAGACGTGTGCTCTTCCGATCTTCGATAGTAGCCTATGAAGCATGCTTACTA
DMS_Lib_Amplify_R4	CAAGCAGAAGACGGCATACGAGATTAGGTGTCGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTTCGATAGTAGCCTATGAAGCATGCTTACTA
DMS_Lib_Amplify_R5	CAAGCAGAAGACGGCATACGAGATGAATCACTGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTTCGATAGTAGCCTATGAAGCATGCTTACTA
DMS_Lib_Amplify_R6	CAAGCAGAAGACGGCATACGAGATCATAGAGTGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTTCGATAGTAGCCTATGAAGCATGCTTACTA
DMS_custom_seq_F	GGTGGTGTCTCGACGATAGCACGTCTCTGACACCGAATTC
DMS_custom_seq_R	GCCCAGCCGGCGATGGCCATGGATATCGGAATTAATTC

793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829

Table2. X-ray data collection and refinement statistics for Tae1^{C110S} (PDB ID: 7TVH)

	Crystal 1	832
Data collection		
Space group	P 1 21 1	
Cell dimensions		
<i>a</i> , <i>b</i> , <i>c</i> (Å)	39.121 108.512 82.402	
Resolution (Å)	65.56 - 1.71 (1.77 - 1.71)	
<i>R</i> _{sym} or <i>R</i> _{merge}	0.03896 (0.9019)	
<i>I</i> / <i>I</i>	10.48 (0.80)	
Completeness (%)	98.45 (87.02)	
Redundancy	2.00 (2.00)	
CC _{1/2}	0.999 (0.469)	
Refinement		
Resolution (Å)	65.56 - 1.71	
No. reflections	72921 (6436)	
<i>R</i> _{work} / <i>R</i> _{free}	0.2113 / 0.2430	
No. atoms	4673	
Protein	4277	
Water	396	
<i>B</i> -factors		
Protein	38.21	
Water	43.00	
R.m.s. deviations		
Bond lengths (Å)	0.006	
Bond angles	0.77	
Ramachandran favored (%)	98.09	
Ramachandran allowed (%)	1.91	
Ramachandran outliers (%)	0	
Rotamer outliers	0	

*Values in parentheses are for highest-resolution shell.

References:

1. Peterson, S. B., Bertolli, S. K. & Mougous, J. D. The central role of interbacterial antagonism in bacterial life. *Current Biology* **30**, R1203–R1214 (2020).
2. Jani, A. J. & Cotter, P. A. Type VI Secretion: Not just for pathogenesis anymore. *Cell Host & Microbe* **8**, 2–6 (2010).
3. Schwarz, S., Hood, R. D. & Mougous, J. D. What is type VI secretion doing in all those bugs? *Trends in Microbiology* **18**, 531–537 (2010).
4. Russell, A. B., Peterson, S. B. & Mougous, J. D. Type VI secretion system effectors: poisons with a purpose. *Nat Rev Microbiol* **12**, 137–148 (2014).
5. Russell, A. B. *et al.* Type VI secretion delivers bacteriolytic effectors to target cells. *Nature* **475**, 343–347 (2011).
6. Hood, R. D. *et al.* A Type VI Secretion System of *Pseudomonas aeruginosa* targets a toxin to bacteria. *Cell Host & Microbe* **7**, 25–37 (2010).
7. Russell, A. B. *et al.* A widespread bacterial Type VI Secretion effector superfamily identified using a heuristic approach. *Cell Host & Microbe* **11**, 538–549 (2012).
8. Russell, A. B. *et al.* A Type VI Secretion-related pathway in Bacteroidetes mediates interbacterial antagonism. *Cell Host & Microbe* **16**, 227–236 (2014).
9. LaCourse, K. D. Conditional toxicity and synergy drive diversity among antibacterial effectors. *Nature Microbiology* **3**, 9 (2018).
10. Chou, S. *et al.* Structure of a peptidoglycan amidase effector targeted to Gram-negative bacteria by the Type VI Secretion System. *Cell Reports* **1**, 656–664 (2012).
11. Vollmer, W., Blanot, D. & De Pedro, M. A. Peptidoglycan structure and architecture. *FEMS Microbiol Rev* **32**, 149–167 (2008).
12. Vollmer, W. & Seligman, S. J. Architecture of peptidoglycan: more data and more models. *Trends in Microbiology* **18**, 59–66 (2010).

- 878 13. Furchtgott, L., Wingreen, N. S. & Huang, K. C. Mechanisms for maintaining cell shape in rod-
879 shaped Gram-negative bacteria: Rod-shape maintenance in Gram-negative bacteria. *Molecular*
880 *Microbiology* **81**, 340–353 (2011).
- 881 14. Elazar, A. *et al.* Mutational scanning reveals the determinants of protein insertion and association
882 energetics in the plasma membrane. *eLife* **5**, e12125 (2016).
- 883 15. Barad, B. A. *et al.* Differences in the chitinolytic activity of mammalian chitinases on soluble and
884 insoluble substrates. *Protein Science* **29**, 952–963 (2020).
- 885 16. Fowler, D. M. & Fields, S. Deep mutational scanning: a new style of protein science. *Nat Methods*
886 **11**, 801–807 (2014).
- 887 17. Rubin, A. F. *et al.* A statistical framework for analyzing deep mutational scanning data. *Genome*
888 *Biol* **18**, 150 (2017).
- 889 18. Thompson, S., Zhang, Y., Ingle, C., Reynolds, K. A. & Kortemme, T. Altered expression of a quality
890 control protease in *E. coli* reshapes the in vivo mutational landscape of a model enzyme. *eLife* **9**,
891 e53476 (2020).
- 892 19. Schanda, P. *et al.* Atomic model of a cell-wall cross-linking enzyme in complex with an intact
893 bacterial peptidoglycan. *J. Am. Chem. Soc.* **136**, 17852–17860 (2014).
- 894 20. Maya-Martinez, R. *et al.* Recognition of peptidoglycan fragments by the transpeptidase Pbp4 from
895 *Staphylococcus aureus*. *Front. Microbiol.* **9**, 3223 (2019).
- 896 21. Sattler, M. Heteronuclear multidimensional NMR experiments for the structure determination of
897 proteins in solution employing pulsed field gradients. *Progress in Nuclear Magnetic Resonance*
898 *Spectroscopy* **34**, 93–158 (1999).
- 899 22. Harz, H., Burgdorf, K. & Höltje, J.-V. Isolation and separation of the glycan strands from murein of
900 *Escherichia coli* by reversed-phase high-performance liquid chromatography. *Analytical*
901 *Biochemistry* **190**, 120–128 (1990).
- 902 23. Gonzalez-Delgado, L. S. *et al.* Two-site recognition of *Staphylococcus aureus* peptidoglycan by
903 lysostaphin SH3b. *Nat Chem Biol* **16**, 24–30 (2020).

- 904 24. Caparrós, M., Pisabarro, A. G. & de Pedro, M. A. Effect of D-amino acids on structure and
905 synthesis of peptidoglycan in *Escherichia coli*. *J Bacteriol* **174**, 5549–5559 (1992).
- 906 25. Lam, H. *et al.* D-amino acids govern stationary phase cell wall remodeling in bacteria. *Science* **325**,
907 1552–1555 (2009).
- 908 26. Basler, M., Pilhofer, M., Henderson, G. P., Jensen, G. J. & Mekalanos, J. J. Type VI Secretion
909 requires a dynamic contractile phage tail-like structure. *Nature* **483**, 182–186 (2012).
- 910 27. Vollmer, W., Joris, B., Charlier, P. & Foster, S. Bacterial peptidoglycan (murein) hydrolases. *FEMS*
911 *Microbiol Rev* **32**, 259–286 (2008).
- 912 28. Daly, K. E., Huang, K. C., Wingreen, N. S. & Mukhopadhyay, R. Mechanics of membrane bulging
913 during cell-wall disruption in Gram-negative bacteria. *Phys. Rev. E* **83**, 041922 (2011).
- 914 29. Ursinus, A. *et al.* Murein (peptidoglycan) binding property of the essential cell division protein FtsN
915 from *Escherichia coli*. *J Bacteriol* **186**, 6728–6737 (2004).
- 916 30. Koch, A. L. Additional arguments for the key role of “smart” autolysins in the enlargement of the wall
917 of Gram-negative bacteria. *Research in Microbiology* **141**, 529–541 (1990).
- 918 31. Kari, J., Andersen, M., Borch, K. & Westh, P. An inverse Michaelis–Menten approach for interfacial
919 enzyme kinetics. *ACS Catal.* **7**, 4904–4914 (2017).
- 920 32. Terry, B. J., Jack, W. E. & Modrich, P. Facilitated diffusion during catalysis by EcoRI endonuclease.
921 Nonspecific interactions in EcoRI catalysis. *Journal of Biological Chemistry* **260**, 13130–13137
922 (1985).
- 923 33. Stracy, M. *et al.* Transient non-specific DNA binding dominates the target search of bacterial DNA-
924 binding proteins. *Molecular Cell* **81**, 1499-1514.e6 (2021).
- 925 34. Broendum, S. S. *et al.* High avidity drives the interaction between the streptococcal C1 phage
926 endolysin, PlyC, with the cell surface carbohydrates of Group A *Streptococcus*. *Mol Microbiol* **116**,
927 397–415 (2021).

- 928 35. McLaren, A. D. & Packer, L. Some aspects of enzyme reactions in heterogeneous systems. in
929 *Advances in enzymology and related areas of molecular biology* (ed. Nord, F. F.) vol. 33 245–308
930 (Interscience Publishers, 1970).
- 931 36. Grosdidier, A., Zoete, V. & Michielin, O. SwissDock, a protein-small molecule docking web service
932 based on EADock DSS. *Nucleic Acids Research* **39**, W270–W277 (2011).
- 933 37. Stankeviciute, G. *et al.* Differential modes of crosslinking establish spatially distinct regions of
934 peptidoglycan in *Caulobacter crescentus*. *Mol Microbiol* **111**, 995–1008 (2019).
- 935 38. Delaglio, F. *et al.* NMRPipe: A multidimensional spectral processing system based on UNIX pipes.
936 *J Biomol NMR* **6**, (1995).
- 937 39. Johnson, B. A. & Blevins, R. A. NMR View: A computer program for the visualization and analysis
938 of NMR data. *J Biomol NMR* **4**, 603–614 (1994).
- 939 40. Downing, A. K. *Protein NMR Techniques*. (Humana Press, 2004).
- 940 41. Frank, O., Kreissl, J. K., Daschner, A. & Hofmann, T. Accurate determination of reference materials
941 and natural isolates by means of quantitative ¹H NMR spectroscopy. *J. Agric. Food Chem.* **62**,
942 2506–2515 (2014).
- 943 42. Favier, A. & Brutscher, B. Recovering lost magnetization: polarization enhancement in biomolecular
944 NMR. *J Biomol NMR* **49**, 9–15 (2011).
- 945 43. Guzman, L. M., Belin, D., Carson, M. J. & Beckwith, J. Tight regulation, modulation, and high-level
946 expression by vectors containing the arabinose P_{BAD} promoter. *J Bacteriol* **177**, 4121–4130 (1995).
- 947 44. Mavor, D. *et al.* Determination of ubiquitin fitness landscapes under different chemical stresses in a
948 classroom setting. *eLife* **5**, e15802 (2016).
- 949 45. Hayes, B. M. *et al.* Ticks resist skin commensals with immune factor of bacterial origin. *Cell* **183**,
950 1562-1571.e12 (2020).
- 951 46. Stover, C. K. *et al.* Complete genome sequence of *Pseudomonas aeruginosa* PAO1, an
952 opportunistic pathogen. *Nature* **406**, 6 (2000).

953 47. MacDowell, A. A. *et al.* Suite of three protein crystallography beamlines with single superconducting
954 bend magnet as the source. *J Synchrotron Rad* **11**, 447–455 (2004).

955 48. Emsley, P. & Cowtan, K. *Coot*: model-building tools for molecular graphics. *Acta Crystallogr D Biol*
956 *Crystallogr* **60**, 2126–2132 (2004).

957 49. Adams, P. D. *et al.* PHENIX: a comprehensive Python-based system for macromolecular structure
958 solution. *Acta Crystallogr D Biol Crystallogr* **66**, 213–221 (2010).

959 50. Davis, I. W., Murray, L. W., Richardson, J. S. & Richardson, D. C. MOLPROBITY: structure
960 validation and all-atom contact analysis for nucleic acids and their complexes. *Nucleic Acids*
961 *Research* **32**, W615–W619 (2004).

962 51. Pettersen, E. F. *et al.* UCSF Chimera - A visualization system for exploratory research and
963 analysis. *J. Comput. Chem.* **25**, 1605–1612 (2004).

964