JACS
JOURNAL OF THE AMERICAN CHEMICAL SOCIETY

ARTICLE

pubs.acs.org/JACS

Site-Specific Incorporation of Photo-Cross-Linker and Bioorthogonal Amino Acids into Enteric Bacterial Pathogens

Shixian Lin,[†] Zhenrun Zhang,[†] Hao Xu,[§] Lin Li,[§] She Chen,[§] Jie Li,[†] Ziyang Hao,[†] and Peng R. Chen*[,†,‡]

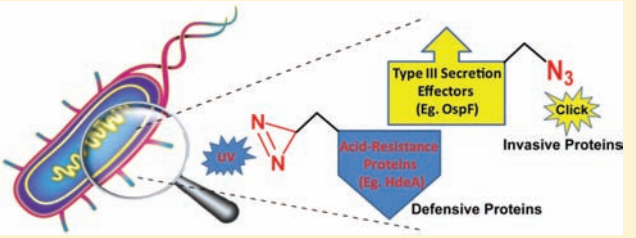
[†]Beijing National Laboratory for Molecular Sciences, College of Chemistry and Molecular Engineering, Peking University, Beijing 100871, China

[‡]Center for Life Sciences, Beijing 100871, China

[§]National Institute of Biological Sciences (NIBS), Beijing 102206, China

Supporting Information

ABSTRACT: Enteric bacterial pathogens are known to effectively pass through the extremely acidic mammalian stomachs and cause infections in the small and/or large intestine of human hosts. However, their acid-survival strategy and pathogenesis mechanisms remain elusive, largely due to the lack of tools to directly monitor and manipulate essential components (e.g., defense proteins or invasive toxins) participating in these processes. Herein, we have extended the pyrrolysine-based genetic code expansion strategy for encoding unnatural amino acids in enteric bacterial species, including enteropathogenic *Escherichia coli*, *Shigella*, and *Salmonella*. Using this system, a photo-cross-linking amino acid was incorporated into a *Shigella* acid chaperone HdeA (shHdeA), which allowed the identification of a comprehensive list of in vivo client proteins that are protected by shHdeA upon acid stress. To further demonstrate the application of our strategy, an azide-bearing amino acid was introduced into a *Shigella* type 3 secretion effector, OspF, without interruption of its secretion efficiency. This site-specifically installed azide handle allowed the facile detection of OspF's secretion in bacterial extracellular space. Taken together, these bioorthogonal functionalities we incorporated into enteric pathogens were shown to facilitate the investigation of unique and important proteins involved in the pathogenesis and stress-defense mechanisms of pathogenic bacteria that remain exceedingly difficult to study using conventional methodologies.



INTRODUCTION

Proteins are the most abundant biomolecules within living organisms and participate in essentially all cellular processes. Beyond their indispensible physiological functions, a repertoire of proteins have also been evolved by bacterial pathogens to fulfill diverse roles during pathogenesis.[1,2] For example, a variety of bacterial toxic effectors are employed to invade host cells and cause diseases, while numerous defensive proteins are recruited to assist these invaders to survive and replicate in the host environment.[3,4] Enteric bacterial pathogens, including virulent *Escherichia coli* species (e.g., enteropathogenic *E. coli* (EPEC), enterohemorrhagic *E. coli* (EHEC), etc.), *Shigella* species (abbreviated here to *Shigella*), and *Salmonella* species (abbreviated here to *Salmonella*), continuously possess a formidable threat to human health, as exemplified by the most recent deadly outbreak of virulent *E. coli* strain O104:H4 in Europe.[5,6] These microorganisms have developed highly efficient acid-protection mechanisms for passage through the extremely acidic mammalian stomachs (pH 1−3).[7−9] Such acid-resistance systems are essential for these pathogens to reach their primary infection sites in the small intestine, where diverse bacterial toxins are injected into host cells to cause infections.[10,11] Labeling and manipulation of these unique invasive or defensive proteins under living conditions will substantially enhance our ability to dissect their important biological processes.

Noninvasive photochemical reactions have been extensively utilized for the study of protein function and dynamics. For example, a panel of protein photo-cross-linking probes have been created for covalently trapping transient protein−protein interactions that remain challenging for other methods to capture in living systems.[12−19] In addition, bioorthogonal ligation reactions have emerged as a powerful strategy for tagging proteins with diverse labeling probes in vivo. Such methods rely on the widely used bioorthogonal reactions, such as the copper-assisted azide−alkyne cycloaddition (CuAAC), that have found broad biological applications in recent years.[20−26] A key requirement in applying these strategies for protein labeling is to selectively incorporate the chemical functional groups into target proteins in living cells. Both enzymatic modification and genetic-incorporation approaches have been developed in recent years toward installation of bioorthogonal handles into proteins of interest. In particular, a genetic code expansion strategy has allowed the site-specific introduction of bioorthogonal functionalities in the form

Received: September 29, 2011
Published: November 15, 2011

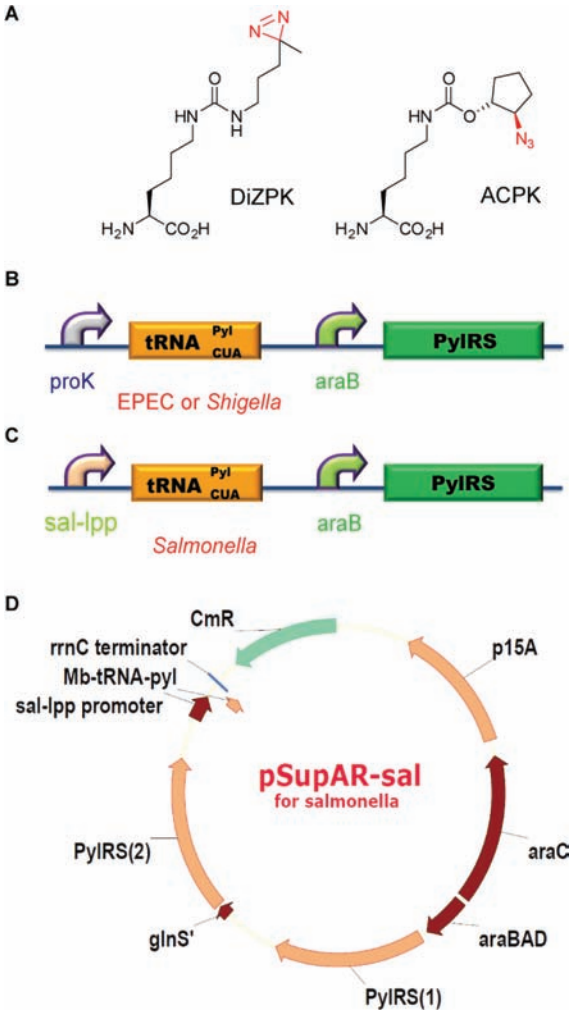


Figure 1. Expanding the genetic code of enteric bacterial pathogens. (A) Structures of unnatural amino acids used in this study. The photo-cross-linking amino acid DiZPK (((3-(3-methyl-3*H*-diazirin-3-yl)propamino)carbonyl)-N^{ϵ}-L-lysine) and the azide-bearing amino acid ACPK (N^{ϵ}-((((1*R*,2*R*)-2-azidocyclopentyl)oxy)carbonyl)-L-lysine) are shown with the bioorthogonal functional groups colored in red. (B, C) Promoters used for expressing $tRNA^{Pyl}_{CUA}$ in EPEC (proK promoter), *Shigella* (proK promoter), or *Salmonella* (sal-lpp promoter) as well as the promoter used for expressing PylRS. (D) Plasmid map of pSupAR used in *Salmonella* (pSupAR-sal).

of unnatural amino acids (UAAs) into proteins in both prokaryotic and eukaryotic cells.[27−30] However, although UAAs carrying bioorthogonal groups have been successfully incorporated into proteins in the laboratory stains of *E. coli* cells, this strategy has never been used in pathogenic strains of *E. coli* as well as other Gram-negative bacterial species.

While being closely related, enteric bacterial species possess distinct physiological and pathological features. For example, most *E. coli* are motile and indole-positive and ferment lactose. By contrast, the human-adapted pathogens *Shigella* are nonmotile and indole- and lactose-negative.[31] *Shigella* species are the causative agent of bacillary dysentery with an estimated 165 million cases and about 1.1 million deaths worldwide annually.[32] Most outbreaks of food-borne infections in developed countries are also caused by *Shigella* as well as *E. coli* O157:H7, a notoriously known virulent strain of *E. coli* that is highly similar to *Shigella*.[31]

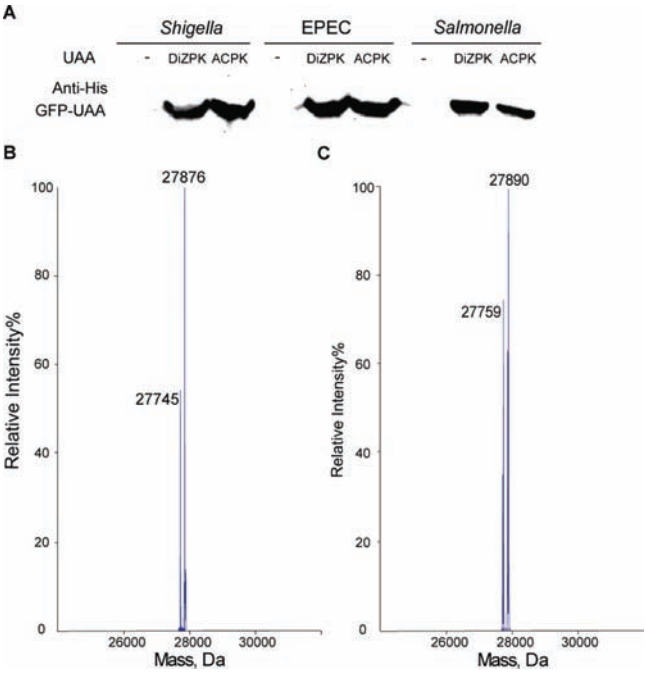


Figure 2. (A) Incorporation of DiZPK and ACPK into the model protein GFP-149TAG in EPEC, *Shigella*, and *Salmonella*. Mass spectrometry of GFP with DiZPK (B) and ACPK (C) incorporated at residue 149 in *Salmonella* cells. MW of GFP149-DiZPK: calculated, 27 875; found, 27 876. MW of GFP149-ACPK: calculated, 27 890; found, 27 890. See Figures S4 and S5 (Supporting Information) for MS/MS results of GFP149-DiZPK and GFP149-ACPK proteins purified from *Salmonella*.

To be highly virulent, *Shigella* has evolved a series of toxins to infect intestinal epithelial cells of the human host, and many of them are not found in *E. coli*. In particular, several recently identified type 3 secretion (T3S) effectors, such as OspF (outer *Shigella* protein F), were shown to directly trigger the epigenetic modifications of host cells, which will ultimately lead to the suppression of the host innate immune response.[33,34] This novel strategy employed by *Shigella* to subvert host-cell signaling pathways is not found in *E. coli* cells. Another variation exhibited by these bacterial species is that whereas *E. coli* and *Shigella* are both highly acid-resistant, *Salmonella* is only moderately acid-tolerant, due to the lack of extreme acid-survival mechanisms and the existence of an alternative, moderately effective acid-tolerant system.[9] In addition, virulent *E. coli* strains such as O157:H7 have also been shown to be more acid-resistant than laboratory stains.[9] Finally, the increase in antibiotic resistance observed in certain virulent strains of enteric pathogens further underscores the importance of studying the pathogenesis of distinct enteric bacterial species.[35,36]

Herein, we report the extension of the genetic code expansion strategy into Gram-negative enteric bacterial pathogens, including *Shigella*, *Salmonella*, and virulent *E. coli* strains, using the pyrrolysine (Pyl) based system. For proof-of-concept, two recently developed UAAs from our laboratory carrying a photoaffinity group (DiZPK) or an azide handle (ACPK) were genetically encoded in EPEC, *Shigella flexneri*, and *Salmonella typhimurium* (Figure 1) by employing the pyrrolysyl-tRNA synthetase (PylRS)—$tRNA^{Pyl}_{CUA}$ pair. These two UAAs were previously used in laboratory strains of *E. coli* cells for capturing the acid-mediated chaperone—substrate interactions as well as in situ protein labeling.[19,37] In the

current study, we site-specifically incorporated DiZPK and ACPK into the *Shigella* acid chaperone HdeA (shHdeA) and the T3S effector OspF, respectively. This enabled the investigation of the acid-resistance mechanism as well as the toxin-secretion system in *Shigella*, a pathogenic microorganism known to be more acid-resistant and virulent than the laboratory strains of *E. coli*.[9,31] Such UAAs can be generally applicable for the investigation of invasion and defense strategies in distinct species of enteric pathogens that remain difficult to study using conventional methodologies.

RESULTS AND DISCUSSION

We started by expanding the genetic code of enteric bacteria to introduce unnatural functionalities into proteins in these microorganisms. The PylRS−$tRNA^{Pyl}_{CUA}$ pair has recently emerged as a powerful tool for expanding the genetic code of both prokaryotic and eukaryotic cells.[27,28,38−44] The wild-type PylRS and its mutant forms were shown to be directly used by the translation machinery of *E. coli*, yeast, or mammalian cells to encode Pyl and diverse Pyl analogues at an in-frame amber mutation in response to $tRNA^{Pyl}_{CUA}$. There are considerable genetic differences among these enteric bacterial species that might affect the expression and utilization of this PylRS−$tRNA^{Pyl}_{CUA}$ pair. Therefore, we first tested the compatibility of the *E. coli* vector pSupAR for expressing the PylRS−$tRNA^{Pyl}_{CUA}$ pair in EPEC, *Shigella*, and *Salmonella*. The plasmid pSupAR carries two copies of PylRS and one copy of $tRNA^{Pyl}_{CUA}$ (Figure 1B−D; Figure S1, Supporting Information).[28] Cotransformation of this plasmid with a plasmid carrying GFP-149TAG in EPEC or *Shigella* cells generated the full-length green fluorescent protein (GFP) with ACPK or DiZPK site-specifically incorporated at residue 149 (Figure 2A). However, we could not detect the full-length GFP production in *Salmonella* harboring these two plasmids (data not shown). Since one of the PylRS genes was under the control of the arabinose promoter, which had been routinely used to express proteins in all these bacterial species,[38,45] we suspected that the problem was likely due to the incompatibility of the tRNA promoter on pSupAR with *Salmonella* cells.

Indeed, our sequence alignment showed that the proK promoter and terminator used on pSupAR for expressing $tRNA^{Pyl}_{CUA}$ in *E. coli* are conserved in both EPEC and *Shigella*, but not in *Salmonella*. Searching the genomic sequence of *Salmonella* revealed that an lpp promoter is used for tRNA expression and its sequence is highly similar to that of the lpp promoter in *E. coli*. (Figure 1C; Figures S2 and S3, Supporting Information). Therefore, we replaced the proK promoter with the lpp promoter from *Salmonella* (sal-lpp) to generate a new plasmid named pSupAR-sal (Figure 1D). Both DiZPK and ACPK were found to be incorporated into GFP at residue 149 in *Salmonella* by using pSupAR-sal (Figure 2A). These results can be further confirmed by mass spectrometry, which showed that these two UAAs were incorporated into proteins in EPEC, *Shigella*, and *Salmonella* without being modified inside cells (Figure 2; Figures S4−S7, Supporting Information).

Next, we used the photo-cross-linking amino acid DiZPK to study the in vivo protein−protein interactions in enteric bacteria under acid-stress conditions. Surviving through the extremely acidic mammalian stomach is a prerequisite for these bacteria to infect the host; therefore, various acid-resistance strategies have been developed to protect their internal proteins from acid-induced aggregation.[9,11] Among these defense mechanisms, acid chaperones such as HdeA in *E. coli* play essential roles in preserving the integrity of proteins upon acid attack.[46,47] Studying the interactions between these acid-protection chaperones and their acid-vulnerable client proteins is highly desired to understand the mechanism underlying such bacterial acid-defense systems. We reason that protein photo-cross-linking is particularly well-suited for studying the transient protein interactions in the bacterial periplasm under acid stress (pH < 3), since this method is not affected by the highly acidic pH whereas most conventional methods (e.g., co-immunoprecipitation) are incompatible with such a harsh condition. Furthermore, the covalently captured protein−substrate complexes are able to sustain the subsequent sample manipulation procedures, such as extensive washing with detergents. Our previous study demonstrated that the diazirine-bearing amino acid DiZPK has higher cross-linking efficiency and fidelity than the currently used probes (e.g., *p*-benzoylphenylalanine, pBpa) for identifying the in vivo client proteins of HdeA.[8,19] Employing this strategy, we aim to uncover the in vivo substrates of shHdeA, the HdeA archetype protein originally found in *Shigella* that is essential in supporting its acid resistance.[46]

Existing as a dimer at neutral pH, shHdeA can be converted to highly flexible monomers when the pH drops to below 3, enabling its two hydrophobic regions to interact with the client proteins.[46] The incorporation of DiZPK into shHdeA at different sites did not interrupt its dimer formation, as confirmed by native polyacrylamide gel electrophoresis (PAGE) analysis (Figure S8A, Supporting Information). We then focused on the shHdeA variant carrying DiZPK at residue V58 in one of its hydrophobic regions for in vivo photo-cross-linking experiments. *Shigella* cells expressing this variant (shHdeA58-DiZPK) were treated at pH 2.3 for 30 min before being subjected to 365 nm light irradiation for 5 min. Immunoblotting analysis of lysates from the acid-treated cells detected multiple shHdeA-containing protein bands with higher molecular weights than those of the shHdeA monomer and dimer, representing the cross-linked complexes of shHdeA (Figure 3B). By contrast, only the monomeric and dimeric forms of shHdeA were detected at pH 7 upon UV irradiation at residue 58 as well as residue 38 located at the dimer interface (Figure S8B).

To characterize the nature of the proteins cross-linked to shHdeA, we performed gel-based proteomic analysis on the cross-linked bands with molecular weights above that of the shHdeA dimer in two independent experimental runs. The mass spectrometry identified a total of 69 in vivo client proteins of shHdeA, and a large portion of these proteins were not found in our previous study on *E. coli* HdeA (Figure 3C). Among these identified proteins, SurA and DegP were two common substrates shared by shHdeA and *E. coli* HdeA. The interactions between shHdeA and SurA were subsequently verified by immunoblotting analysis using an anti-SurA antibody (Figure 3D). Furthermore, many of these identified client proteins have functions similar to those of *E. coli* HdeA's substrates. For example, AraF (a component of the L-arabinose transport system) was a client protein of *E. coli* HdeA, and similarly, amino acid transporters, including GlnH, ArgT, HisJ, and ProX, were among shHdeA's substrates. Notably, our mass specrometry data showed that alkaline phosphatase (AP) and MDH (methanol dehydrogenase) were among the in vivo client proteins of shHdeA. The interaction between shHdeA and AP was further verified by immunoblotting assay using an anti-AP antibody (Figure 3E). Both AP and MDH are routinely used in vitro model proteins to evaluate the refolding efficiency of HdeA and its homologue HdeB. However,

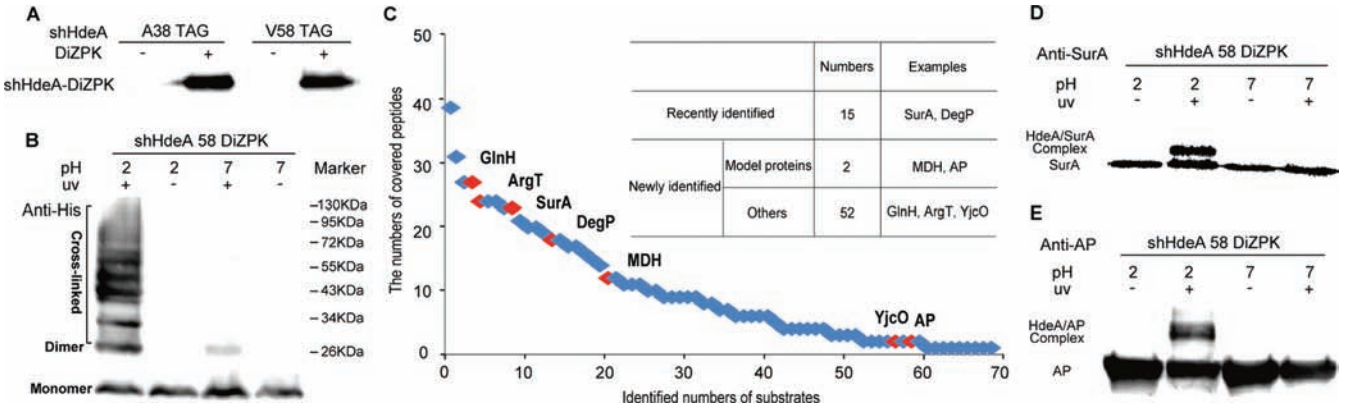


Figure 3. Identification of shHdeA's client proteins under acid stress. (A) Incorporation of DiZPK into A38 and V58 positions on shHdeA in *Shigella*. (B) Photo-cross-linking of *Shigella* cells expressing shHdeA58-DiZPK at pH 7 and 2 with or without UV irradiation for 5 min. (C) Summary of shHdeA client proteins in *Shigella*. For each client protein, the numbers of covered peptides are plotted with blue squares, while the red squares represent the client proteins highlighted in the paper. See Table S1 (Supporting Information) for detailed information on each shHdeA client protein. (D, E) Verification of the cross-linked complexes of shHdeA with SurA or AP by immunoblotting analysis.

whether these two proteins interact with such acid chaperones in vivo remains under debate. Our results for the first time revealed the interaction between shHdeA and AP or MDH upon acid stress in vivo, which warranted future usage of AP and MDH as the physiologically relevant client proteins to assess the working mechanisms of this family of acid chaperones. Finally, our mass spectrometry data also uncovered dozens of client proteins that have never been identified before. Some of them are metal- and radical-related proteins such as SodB, AhpC, NikR, and FUR, while the rest of the proteins remain uncharacterized. Taken together, the identified native substrates of shHdeA in this study will be valuable for expanding our knowledge on the in vivo functions of acid-stress chaperones and thus their roles in acid defense of *Shigella*.

One of the major advantages for bioorthogonal protein labeling is the small-sized labeling fluorophores that can be incorporated at virtually any desired position on the protein.[43] This feature is extremely valuable when the bulky imaging or affinity tags (e.g., GFP and glutathione S-transferase (GST)) possess significant perturbations to the target protein's structure and/or function.[48] Many bacterial toxins are such examples that their function or infection capability will be significantly affected upon fusion with these large protein tags.[49,50] To demonstrate the utilization of bioorthogonal chemistry for "precise" labeling of infectious proteins, we chose to introduce bioorthogonal handles into OspF in *Shigella*.

The T3S system is widely used by Gram-negative bacteria to infect mammalian hosts.[1,2,51] The needle-like T3S complex mediates the injection of toxins from the bacterial cytoplasm directly into host cells. These virulent proteins are believed to travel through this pathway in a largely unfolded manner, which will render the bulky and unfolding-resistant proteins incompatible for labeling such toxins. OspF is a recently identified T3S effector from *Shigella*, the model organism for studying the mechanism and functions of the T3S apparatus. Upon being secreted into host cells, OspF works as an epigenetic modulator by irreversibly dephosphorylating MAPKs (mitogen-activated protein kinases) such as Erk that will alter transcription genes involved in host immune responses.[52−55] Homologue proteins of OspF were only found in *Salmonella*, and not in *E. coli*, indicating that this is a unique invasive strategy in *Shigella* and *Salmonella* and not in *E. coli*. We therefore chose OspF as the

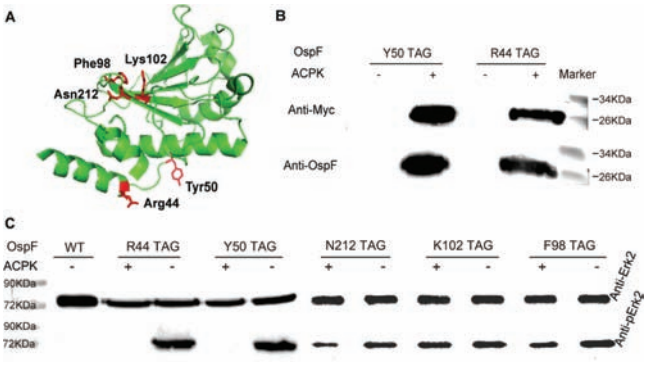


Figure 4. Expression of OspF carrying an azide-bearing unnatural amino acid (ACPK) in *Shigella*. (A) Crystal structure of OspF with mutation sites we used in this study labeled. (B) Incorporation of ACPK into the full-length OspF at residue R44 or Y50 in the *ospF* knockout *Shigella* strain. The cells were cotransformed with plasmids pBAD24-OspF-TAG-myc and pSupar, and the expression of OspF was verified by Western blotting analysis using both anti-OspF and anti-Myc antibodies. (C) In vitro MAPK phosphate removal assay. The same amounts of WT-OspF and OspF variants with or without ACPK were incubated with phosphor-GST-Erk2 for 45 min, and the results were analyzed by immunoblotting analysis using both anti-Erk2 and anti-pErk2 antibodies (Figure S9, Supporting Information).

model protein, and our azide-bearing UAA-ACPK was site-specifically incorporated into OspF at different positions in *S. flexneri* cells (Figure 4A,B; Figure S9, Supporting Information). Interestingly, we found that the activity of OspF was highly dependent on the ACPK-incorporation site (Figure 4C). When ACPK was incorporated at positions that are located away from the catalytic center, including residues Arg 44 and Tyr 50, the resulting OspF variants (OspF-R44-ACPK and OspF-Y50-ACPK) exhibited the dephophorylation activity on Erk at a level essentially comparable to that of wild-type OspF (WT-OspF). By contrast, for incorporation sites surrounding the catalytic center of OspF, including residues Phe 98, Lys 102, and Asn 212, their dephophorylation activities on Erk were significantly disrupted (Figure 4C).

We then performed the toxin-secretion assay to compare the secretion efficiency of a panel of OspF variants, including

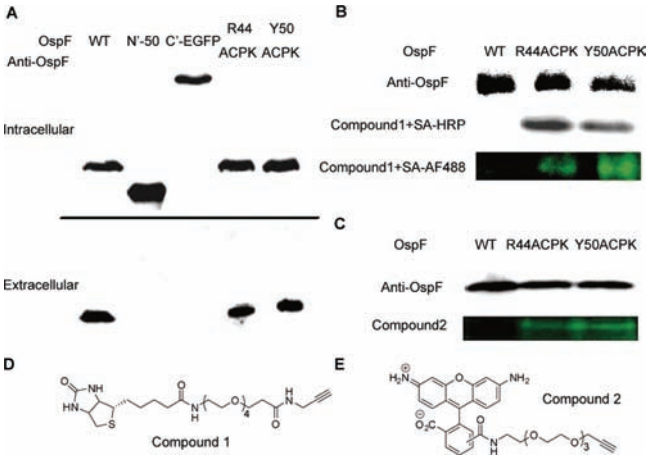


Figure 5. Verification of the secreted ACPK-bearing OspF from *Shigella* cells to extracellular space. (A) Congo red secretion assay of *Shigella* OspF knockout strains complemented with WT-OspF, OspF-GFP, N′50-OspF, or ACPK-bearing OspF variants. The loading of OspF variants was 4 times more than that of WT-OspF. (B) The secreted OspF-ACPK variants were "click" labeled with compound **1** by CuAAC. The samples were separated by SDS−PAGE and stained with streptavidin−HRP and streptavidin−AF488, respectively. (C) The secreted OspF-ACPK variants were reacted with compound **2** via CuAAC, and the reaction products were verified by in-gel detection. The immunoblotting analysis using anti-OspF antibody was performed as the loading control. (D, E) Structures of compound **1** and compound **2** used in this study.

OspF-WT, OspF-EGFP (OspF with a C-terminal-fused enhanced green fluorescent protein (EGFP)), N′50-OspF (OspF with the N-terminal 50 amino acids truncated), and the two ACPK-incorporated OspF variants (OspF-R44-ACPK and OspF-Y50-ACPK). These variants were successfully expressed inside *S. flexneri* bacteria which were then subjected to the Congo red induced toxin secretion. The extracellularly secreted OspF was detected by immunoblotting analysis using the anti-OspF antibody (Figure 5A). Our results showed that the two ACPK-incorporated OspF variants can be secreted into the extracellular space with an efficiency similar to that of WT-OspF. In contrast, the secretion of OspF-EGFP as well as N′50-OspF was completely abandoned. We reasoned that this could be due to the bulky and unfolding-resistant GFP tag that markedly affected OspF's passage through the narrow T3S needle compartment. Next, by using CuAAC in conjunction with the biotin-$(PEG)_4$-alkyne, these secreted OspF-ACPK variants were effectively enriched and detected by immunoblotting analysis with streptavidin−horseradish peroxidase (HRP) or streptavidin−AF488 conjugates (Figure 5B). Furthermore, these secreted variants were shown to be readily captured and detected by AF488-alkyne using in-gel fluorescent detection (Figure 5C,D). Taken together, this bioorthogonal tagging strategy might permit visualization of the secretion process of OspF in living *Shigella* cells, which is under way in our laboratory.

CONCLUSIONS

In summary, we have extended the pyrrolysine-based genetic code expansion strategy into enteric bacterial pathogens. This for the first time, to our knowledge, allowed the utilization of UAAs in investigating essential proteins involved in the pathogenesis and stress-defense processes of virulent Gram-negative bacteria distinct from laboratory strains of *E. coli*. By introducing the photo-cross-linking amino acid DiZPK into shHdeA, we obtained a comprehensive list of native client proteins of this major acid-defense protein in *Shigella*. A large portion of these identified substrates were not found in our previous study on *E. coli* HdeA. Among them, we confirmed shHdeA's interactions with AP and MDH, two routinely used model proteins to evaluate the refolding efficiency of acid chaperones in vitro.[56,57] Such information laid the ground for further investigation of the acid-resistance mechanism of *Shigella* upon passing through the highly acidic mammalian stomachs. Since the levels of acid tolerance significantly varied among different enteric bacterial species, this highly efficient photoaffinity probe in conjunction with the cross-linking strategy we applied here can be easily adapted to probe the acid-survival mechanisms of diverse intestinal microorganisms.

The utilization of the azide-bearing UAA-ACPK enabled the precise tagging of a bacterial T3S effector, OspF, with a bioorthogoanl handle, which facilitated the detection and labeling of a *Shigella* toxin without disruption of its enzymatic activity. Such a method may help us to further understand how unique T3S effectors such as OspF are utilized by *Shigella* to manipulate host transcriptional responses. Furthermore, adopting this click labeling and detection strategy in various secretion toxins might give deep insight into how virulence factors are delivered into intestinal epithelia cells to subvert the host immune response and cause infections.

Finally, our strategy could serve as a template for those who wish to expand the genetic code of various bacteria species. This work also further demonstrated the "global" orthogonality of the PylRS−$tRNA^{Pyl}_{CUA}$ pair in diverse living organisms. By using the proper promoters for $tRNA^{Pyl}_{CUA}$ and PylRS, this attractive system can be applied for incorporation of UAAs into various prokaryotic and eukaryotic species. Expanding the genetic code of opportunistic Gram-negative or Gram-positive pathogens such as *Pseudomonas aeruginosa* and *Staphylococcus aureus* is currently under way in our laboratory.

EXPERIMENTAL SECTION

Expression of UAA-Bearing Proteins in EPEC, *Shigella*, or *Salmonella*. The plasmid pSupAR-Mb was cotransformed with a plasmid carrying the target genes with an amber mutation into EPEC, *Shigella*, or *Salmonella* cells by electrotransformation. Overnight bacterial cultures were grown at 37 °C in Luria−Bertani (LB) medium for about 2.5 h until OD_{600} reached 0.5, at which point 1 mM (final concentration) DiZPK or ACPK was added to the culture. The bacteria were continuously grown at 37 °C for another 30 min before being transferred to 30 °C for induction in the presence of 0.02% arabinose for 8 h.

In Vitro and in Vivo Photo-Cross-Linking Assay. For in vitro photo-cross-linking assay, the shHdeA58-DiZPK protein (10 nM final concentration) and the substrate protein (10 nM final concentration) were mixed together, and the solution was titrated to pH 2.3 with 1 M HCl carefully, followed by incubation at 4 °C for 1 h. The sample was UV irradiated (365 nm) for 5 min with CL-1000 ultraviolet cross-linker (UVP) on ice. Then the pH was recovered to 7 by adding 1 M NaOH very carefully. The samples were separated by sodium dodecyl sulfate (SDS)−PAGE and analyzed by immunoblotting with an anti-His antibody. For in vivo photo-cross-linking assay, the bacterial (*Shigella*) cells harboring DiZPK-incorporated shHdeA were centrifuged at 4000 rpm and 4 °C for 12 min. The supernatant was then discarded, and the bacterial pellet was resuspended by 10 mM citrate acid containing 100 mM NaCl (pH 2.3), incubated at 4 °C for 30 min, and then

irradiated with UV 365 nm with CL-1000 ultraviolet cross-linker on ice. After being centrifuged at 4000 rpm and 4 °C again, the bacterial cells were washed by phosphate-buffered saline (PBS) before being subjected to SDS−PAGE and immunoblotting.

Mass Spectrometry Measurement of DiZPK- or ACPK-Incorporated GFP. The purified GFP carrying DiZPK or ACPK (about 1 mg/mL) was loaded into an analytical capillary column (75 μm, 5 cm) packed with Poros 20 R1 packing material (Applied Biosystems, Foster City, CA). An Agilent 1100 binary pump was used to generate an HPLC gradient as follows: 0−100% solvent B in 60 min (solvent A is 0.1 M acetic acid in water; solvent B is 0.1 M acetic acid/70% acetonitrile). The eluted proteins were sprayed into a QSTAR XL mass spectrometer (AB Sciex, Foster City, CA). The spray voltage was set at 2100 V, and the data were acquired in MS mode. The protein charge envelope from the raw spectrum was deconvoluted into noncharged form by the BioAnalyst software provided by the manufacturer.

Gel-Based Proteomic Analysis by Mass Spectrometry. *Shigella* cells expressing shHdeA58-DiZPK were grown in 1 L of LB broth to the stationary phase followed by in vivo photo-cross-linking assay as described above. The cells were then lysed by sonication in buffer A (20 mM sodium phosphate, 0.5 M NaCl, 20 mM imidazole, and 8 M urea, pH 7.4), and the cross-linked products were purified with a Ni−nitrilotriacetic acid (NTA) agarose column (GE Healthcare). Collected fractions were concentrated with Amicon Ultra-15 centrifugal filter units (Millipore) with a molecular weight cutoff of 3000 to a final volume of 500−1000 μL. Samples were then subjected to tricine−glycine SDS−PAGE, and the gel was stained with Coomassie blue. After gel separation, the corresponding protein bands were sliced and then digested in-gel with sequencing grade trypsin (10 ng/μL trypsin, 50 mM ammonium bicarbonate, pH 8.0) overnight at 37 °C. The resulting peptides were extracted with 5% formic acid/50% acetonitrile and 0.1% formic acid/75% acetonitrile sequentially and then concentrated to ∼20 μL. The extracted peptides were separated on an analytical capillary column (50 μm × 10 cm) packed with 5 μm spherical C18 reversed-phase material (YMC, Kyoyo, Japan). An Agilent 1100 binary pump was used to generate an HPLC gradient as follows: 0−5% solvent B in 5 min, 5−40% sovent B in 25 min, 40−100% solvent B in 15 min. For GFP-derived peptides, the eluted peptides were sprayed into a QSTAR XL mass spectrometer equipped with a nano electrospray ionization (nano-ESI) ion source. The mass spectrometer was operated in information-dependent mode with one MS scan followed by three MS/MS scans for each cycle. For shHdeA-interacting proteins, the resulting peptides were analyzed on an LTQ mass spectrometer (Thermo Fisher Scientific, San Jose, CA) in data-dependent mode (the MS scan mass range was from 350 to 2000 Da, and the top five most abundant precursor ions from each MS scan were selected for MS/MS scans). Database searches were performed on an in-house Mascot server (Matrix Science Ltd., London, U.K.). The methionine oxidation was included as a variable modification. For DiZPK or ACPK substitution site mapping on GFP, DiZPK or ACPK substitution for asparagine was also included as a variable modification. The Mascot run was performed on Sept 1, 2011, against the database SwissProt (SwissProt 57.15; 515 203 sequences, 181 334 896 residues).

In Vitro MAPK Phosphate Removal Assay. The OspF-WT and OspF-ACPK proteins expressed in the *Shigella ospF* deletion strain were quantified by anti-myc Western blotting and adjusted to the same level (Figure S9, Supporting Information) with lysis buffer (150 mM NaCl and 20 mM Tris−HCl, pH 7.4). Then 4 μL (1 μg/mL) of phosphor-GST-Erk2 was added to 200 μL of the lysate solutions, and the solutions were incubated at 30 °C for 45 min. Then the reactions were quenched by addition of 4× SDS loading buffer, determined by phosphor-Erk1/2 and Erk1/2 antibodies.[53]

Congo Red Assay[58,59]. The overnight culture of *Shigella* cells harboring different OspF variants (20 mL) was collected by centrifugation at 3000 rpm and 4 °C. The supernatant was discarded, and the bacteria were resuspended in 4 mL of precooled PBS buffer. Then the cells were collected by centrifugation at 4 °C and resuspended in 4 mL of precooled PBS buffer again. After the addition of 10 μL of Congo red (1% stock),[58,59] the solution was cultured at 37 °C for 30 min, followed by centrifugation at 12 000 rpm. The supernatant was filtered with a 0.22 μm filter, and proteins were precipitated by adding trichloroacetic acid (TCA; 8% final concentration) for overnight incubation at 4 °C. Then the precipitates were collected by centrifugation at 13 000 rpm for 30 min and washed with acetone carefully three times. Samples were dissolved in 1× sample buffer and analyzed by Western blot.

In Vitro Labeling and In-Gel Labeling Assay. The secreted WT-OspF (from a 50 mL bacterial culture) or ACPK-incorporated OspF (from a 200 mL bacterial culture) proteins induced by Congo red were dissolved in 300 μL of PBS buffer containing 2% SDS. The biotin-$(PEG)_4$-alkyne (50 μM final concentration) and $Cu(TBTA)_2^{2+}$ (TBTA = tris-(benzyltriazolylmethyl)amine; 50 μM final concentration) were added to the 50 μL PBS solution. The CuAAC reaction was performed on the shaker at 37 °C for 1 h. Then the labeled proteins were separated by SDS−PAGE, transferred to a poly(vinylidene fluoride) (PVDF) film, and stained with streptavidin−HRP or streptavidin−Alexa Fluor 488, respectively. For in-gel labeling assay, Alexa Fluor 488 alkyne (50 μM final concentration) and $Cu(TBTA)_2^{2+}$ (50 μM final concentration) were added to the 50 μL PBS solution. CuAAC was performed under the same conditions. The labeled proteins were separated by SDS−PAGE and imaged with Chem-Doc (BioRad) directly.

■ ASSOCIATED CONTENT

Supporting Information. Experimental details, supplemental data, and complete ref 35. This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author
pengchen@pku.edu.cn

■ ACKNOWLEDGMENT

This work was supported by the National Key Basic Research Foundation of China (Grant 2010CB912300) and National Natural Science Foundation of China (Grant 20932006). We thank Prof. Feng Shao for the OspF antibody and plasmid and Dr. Meng Zhang for the alkaline phosphatase antibody. We acknowledge support from the Peking University Principal Foundation.

■ REFERENCES

(1) Troisfontaines, P.; Cornelis, G. R. *Physiology* **2005**, *20*, 326–339.
(2) Bhavsar, A. P.; Guttman, J. A.; Finlay, B. B. *Nature* **2007**, *449*, 827–834.
(3) Cossart, P.; Sansonetti, P. J. *Science* **2004**, *304*, 242–248.
(4) Nomura, K.; DebRoy, S.; Lee, Y. H.; Pumplin, N.; Jones, J.; He, S. Y. *Science* **2006**, *313*, 220–223.
(5) Bielaszewska, M.; Mellmann, A.; Zhang, W.; Koeck, R.; Fruth, A.; Bauwens, A.; Peters, G.; Karch, H. *Lancet Infect. Dis.* **2011**, *11*, 671–676.
(6) Scheutz, F.; Nielsen, E. M.; Frimodt-Moller, J.; Boisen, N.; Morabito, S.; Tozzoli, R.; Nataro, J. P.; Caprioli, A. *Euro Surveill.* [Online] **2011**, *16*, Article 19889. http://www.eurosurveillance.org/ViewArticle.aspx?ArticleId=19889.
(7) Lin, J. S.; Lee, I. S.; Frey, J.; Slonczewski, J. L.; Foster, J. W. *J. Bacteriol.* **1995**, *177*, 4097–4104.
(8) Waterman, S. R.; Small, P. L. C. *Mol. Microbiol.* **1996**, *21*, 925–940.

(9) Foster, J. W. *Nat. Rev. Microbiol.* **2004**, *2*, 898–907.
(10) Bearson, S.; Bearson, B.; Foster, J. W. *FEMS Microbiol. Lett.* **1997**, *147*, 173–180.
(11) Zhao, B. Y.; Houry, W. A. *Biochem. Cell. Biol.* **2010**, *88*, 301–314.
(12) Chin, J. W.; Martin, A. B.; King, D. S.; Wang, L.; Schultz, P. G. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 11020–11024.
(13) Suchanek, M.; Radzikowska, A.; Thiele, C. *Nat. Methods* **2005**, *2*, 261–267.
(14) Hino, N.; Okazaki, Y.; Kobayashi, T.; Hayashi, A.; Sakamoto, K.; Yokoyama, S. *Nat. Methods* **2005**, *2*, 201–206.
(15) Tippmann, E. M.; Liu, W.; Surnmerer, D.; Mack, A. V.; Schultz, P. G. *ChemBioChem* **2007**, *8*, 2210–2214.
(16) Tanaka, Y.; Bond, M. R.; Kohler, J. J. *Mol. Biosyst.* **2008**, *4*, 473–480.
(17) Majmudar, C. Y.; Lee, L. W.; Lancia, J. K.; Nwokoye, A.; Wang, Q.; Wands, A. M.; Wang, L.; Mapp, A. K. *J. Am. Chem. Soc.* **2009**, *131*, 14240–14242.
(18) Ai, H.-w.; Shen, W.; Sagi, A.; Chen, P. R.; Schultz, P. G. *ChemBioChem* **2011**, *12*, 1854–1857.
(19) Zhang, M.; Lin, S.; Song, X.; Liu, J.; Fu, Y.; Ge, X.; Fu, X.; Chang, Z.; Chen, P. R. *Nat. Chem. Biol.* **2011**, *7*, 671–677.
(20) Rostovtsev, V. V.; Green, L. G.; Fokin, V. V.; Sharpless, K. B. *Angew. Chem., Int. Ed.* **2002**, *41*, 2596–2599.
(21) Tornoe, C. W.; Christensen, C.; Meldal, M. *J. Org. Chem.* **2002**, *67*, 3057–3064.
(22) Prescher, J. A.; Bertozzi, C. R. *Nat. Chem. Biol.* **2005**, *1*, 13–21.
(23) Ning, X.; Guo, J.; Wolfert, M. A.; Boons, G. J. *Angew. Chem., Int. Ed.* **2008**, *47*, 2253–2255.
(24) Hong, V.; Steinmetz, N. F.; Manchester, M.; Finn, M. G. *Bioconjugate Chem.* **2010**, *21*, 1912–1916.
(25) Jewett, J. C.; Bertozzi, C. R. *Chem. Soc. Rev.* **2010**, *39*, 1272–1279.
(26) Kennedy, D. C.; McKay, C. S.; Legault, M. C. B.; Danielson, D. C.; Blake, J. A.; Pegoraro, A. F.; Stolow, A.; Mester, Z.; Pezacki, J. P. *J. Am. Chem. Soc.* **2011**, *133*, 17993–18001.
(27) Wang, Q.; Parrish, A. R.; Wang, L. *Chem. Biol.* **2009**, *16*, 323–336.
(28) Chen, P. R.; Groff, D.; Guo, J.; Ou, W.; Cellitti, S.; Geierstanger, B. H.; Schultz, P. G. *Angew. Chem., Int. Ed.* **2009**, *48*, 4052–4055.
(29) Liu, C. C.; Schultz, P. G. *Annu. Rev. Biochem.* **2010**, *79*, 413–444.
(30) Chin, J. W. *EMBO J.* **2011**, *30*, 2312–2324.
(31) Johnson, J. R. *J. Med. Microbiol.* **2000**, *49*, 583–585.
(32) Zurawski, D. V.; Mumy, K. L.; Faherty, C. S.; McCormick, B. A.; Maurelli, A. T. *Mol. Microbiol.* **2009**, *71*, 350–368.
(33) Ogawa, M.; Honda, Y.; Ashida, H.; Suzuki, M.; Sasakawa, C. *Nat. Rev. Microbiol.* **2008**, *6*, 11–16.
(34) Brennan, D. F.; Barford, D. *Trends Biochem. Sci.* **2009**, *34*, 108–114.
(35) Holden, M. T. G.; et al. *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101*, 9786–9791.
(36) Cameron, D. R.; Howden, B. P.; Peleg, A. Y. *Clin. Infect. Dis.* **2011**, *53*, 576–582.
(37) Hao, Z.; Song, Y.; Lin, S.; Yang, M.; Liang, Y.; Wang, J.; Chen, P. R. *Chem. Commun.* **2011**, *47*, 4502–4504.
(38) Newman, J. R.; Fuqua, C. *Gene* **1999**, *227*, 197–203.
(39) Mukai, T.; Kobayashi, T.; Hino, N.; Yanagisawa, T.; Sakamoto, K.; Yokoyama, S. *Biochem. Biophys. Res. Commun.* **2008**, *371*, 818–822.
(40) Fekner, T.; Li, X.; Chan, M. K. *Eur. J. Org. Chem.* **2010**, *2010*, 4171–4179.
(41) Hancock, S. M.; Uprety, R.; Deiters, A.; Chin, J. W. *J. Am. Chem. Soc.* **2010**, *132*, 14819–14824.
(42) Wan, W.; Huang, Y.; Wang, Z.; Russell, W. K.; Pai, P. J.; Russell, D. H.; Liu, W. R. *Angew. Chem., Int. Ed.* **2010**, *49*, 3211–3214.
(43) Hao, Z.; Hong, S.; Chen, X.; Chen, P. R. *Acc. Chem. Res.* **2011**, *44*, 742–751.
(44) Greiss, S.; Chin, J. W. *J. Am. Chem. Soc.* **2011**, *133*, 14196–14199.
(45) Ahmer, B. M. M.; van Reeuwijk, J.; Timmers, C. D.; Valentine, P. J.; Heffron, F. *J. Bacteriol.* **1998**, *180*, 1185–1193.
(46) Gajiwala, K. S.; Burley, S. K. *J. Mol. Biol.* **2000**, *295*, 605–612.
(47) Hong, W. Z.; Jiao, W. W.; Hu, J. C.; Zhang, J. R.; Liu, C.; Fu, X. M.; Shen, D.; Xia, B.; Chang, Z. Y. *J. Biol. Chem.* **2005**, *280*, 27029–27034.
(48) Meyer, E.; Fromherz, P. *Eur. J. Neurosci.* **1999**, *11*, 1105–1108.
(49) Enninga, J.; Mounier, J.; Sansonetti, P.; Van Nhieu, G. T. *Nat. Methods* **2005**, *2*, 959–965.
(50) Akeda, Y.; Galan, J. E. *Nature* **2005**, *437*, 911–915.
(51) Parsot, C. *Curr. Opin. Microbiol.* **2009**, *12*, 110–116.
(52) Arbibe, L.; Kim, D. W.; Batsche, E.; Pedron, T.; Mateescu, B.; Muchardt, C.; Parsot, C.; Sansonetti, P. J. *Nat. Immunol.* **2007**, *8*, 47–56.
(53) Li, H.; Xu, H.; Zhou, Y.; Zhang, J.; Long, C.; Li, S.; Chen, S.; Zhou, J.-M.; Shao, F. *Science* **2007**, *315*, 1000–1003.
(54) Zhu, Y.; Li, H.; Long, C.; Hu, L.; Xu, H.; Liu, L.; Chen, S.; Wang, D.-C.; Shao, F. *Mol. Cell* **2007**, *28*, 899–913.
(55) Reiterer, V.; Grossniklaus, L.; Tschon, T.; Kasper, C. A.; Sorg, I.; Arrieumerlou, C. *Cell. Signalling* **2011**, *23*, 1188–1196.
(56) Tapley, T. L.; Koerner, J. L.; Barge, M. T.; Hupfeld, J.; Schauerte, J. A.; Gafni, A.; Jakob, U.; Bardwell, J. C. A. *Proc. Natl. Acad. Sci. U.S.A.* **2009**, *106*, 5557–5562.
(57) Tapley, T. L.; Franzmann, T. M.; Chakraborty, S.; Jakob, U.; Bardwell, J. C. A. *Proc. Natl. Acad. Sci. U.S.A.* **2010**, *107*, 1071–1076.
(58) Qadri, F.; Hossain, S. A.; Ciznár, I.; Haider, K.; Ljungh, A.; Wadstrom, T.; Sack, D. A. *J. Clin. Microbiol.* **1988**, *26*, 1343–1348.
(59) Demers, B.; Sansonetti, P. J.; Parsot, C. *EMBO J.* **1998**, *17*, 2894–2903.