De Novo Emergence of Peptides That Confer Antibiotic Resistance

Michael Knopp,[a] Jonina S. Gudmundsdottir,[a] Tobias Nilsson,[b] Finja König,[b] Omar Warsi,[a] Fredrika Rajer,[a] Pia Ädelroth,[b] Dan I. Andersson[a]

aDepartment of Medical Biochemistry and Microbiology, Biomedical Center, Uppsala University, Uppsala, Sweden
bDepartment of Biochemistry and Biophysics, Stockholm University, Stockholm, Sweden

ABSTRACT The origin of novel genes and beneficial functions is of fundamental interest in evolutionary biology. New genes can originate from different mechanisms, including horizontal gene transfer, duplication-divergence, and *de novo* from noncoding DNA sequences. Comparative genomics has generated strong evidence for *de novo* emergence of genes in various organisms, but experimental demonstration of this process has been limited to localized randomization in preexisting structural scaffolds. This bypasses the basic requirement of *de novo* gene emergence, i.e., lack of an ancestral gene. We constructed highly diverse plasmid libraries encoding randomly generated open reading frames and expressed them in *Escherichia coli* to identify short peptides that could confer a beneficial and selectable phenotype *in vivo* (in a living cell). Selections on antibiotic-containing agar plates resulted in the identification of three peptides that increased aminoglycoside resistance up to 48-fold. Combining genetic and functional analyses, we show that the peptides are highly hydrophobic, and by inserting into the membrane, they reduce membrane potential, decrease aminoglycoside uptake, and thereby confer high-level resistance. This study demonstrates that randomized DNA sequences can encode peptides that confer selective benefits and illustrates how expression of random sequences could spark the origination of new genes. In addition, our results also show that this question can be addressed experimentally by expression of highly diverse sequence libraries and subsequent selection for specific functions, such as resistance to toxic compounds, the ability to rescue auxotrophic/temperature-sensitive mutants, and growth on normally nonused carbon sources, allowing the exploration of many different phenotypes.

IMPORTANCE *De novo* gene origination from nonfunctional DNA sequences was long assumed to be implausible. However, recent studies have shown that large fractions of genomic noncoding DNA are transcribed and translated, potentially generating new genes. Experimental validation of this process so far has been limited to comparative genomics, *in vitro* selections, or partial randomizations. Here, we describe selection of novel peptides *in vivo* using fully random synthetic expression libraries. The peptides confer aminoglycoside resistance by inserting into the bacterial membrane and thereby partly reducing membrane potential and decreasing drug uptake. Our results show that beneficial peptides can be selected from random sequence pools *in vivo* and support the idea that expression of noncoding sequences could spark the origination of new genes.

KEYWORDS *Escherichia coli*, antibiotic resistance, aminoglycosides, *de novo*, gene evolution, membrane potential, peptides

Different mechanisms for the acquisition of new genes, such as horizontal gene transfer and duplication-divergence, have been observed in both natural systems and experimental settings (1, 2). However, these mechanisms rely on preexisting genes

as a substrate for evolution and do not explain how these preexisting genes arose *de novo* from noncoding DNA (3, 4). For a protein-coding gene to arise *de novo,* a noncoding nucleotide sequence has to be transcribed and translated and confer a selectable beneficial function. The evolution of an entirely new gene without any functional progenitor has been regarded as an exceedingly rare and improbable event (5), and considering that a random sequence generated by chance must encode a stable polypeptide with a function accessible by selection, this notion of impossibility might seem reasonable. However, comparative genomics in eukaryotes such as insects (6), yeast (7), Hydra (8), primates (9, 10), mouse (11, 12), *Plasmodium* (13), and plants (14) has provided strong evidence that this is possible and that parts of extant genomes contain genes without any recognized homologous sequences that nature had "tinkered" with.

In addition, it is becoming clear that large portions of the genomes of many organisms, including noncoding sequences, are highly transcribed (15, 16), and that these transcripts can contain randomly occurring small open reading frames (sORFs) that are translated (17–19). Historically, sORFs have been largely overlooked due to missing annotations and their incompatibility with many molecular biology detection tools. An increasing number of studies employing comparative genomics and ribosome profiling, however, show that sORFs are pervasive among all investigated organisms and can make up a large fraction of prokaryotic genomes (19–22). Modern tools for gene prediction that incorporate ribosome profiling data estimate the presence of 125 sORFs encoding proteins of ≤ 71 amino acids in *Escherichia coli* (22). Another study based on sequence conservation and ribosome binding site models predicts 217 sORFs encoding peptides of less than 50 amino acids, most of which contain predicted transmembrane domains (23). This pervasive expression of small proteins potentially can serve as a pool for the *de novo* selection of functional peptides encoded by sORFs (20, 21, 23). Furthermore, small proteins have been shown to carry out versatile functions, including attenuation of gene expression and regulation of transporters, stress response, and signal transduction (23–26). The increasing recognition of sORFs as an abundant and functionally diverse class of genes that could originate from stochastic transcription and translation of noncoding sequences make them well suited for studying their *de novo* emergence by experimental evolution. In this study, we examined if short peptides encoded by randomized DNA sequences that were expressed from a plasmid in *E. coli* cells could confer selective benefits in the form of antibiotic resistance and by which mechanisms such new phenotypes could be generated.

RESULTS AND DISCUSSION

Library construction and screening for *de novo* sORFs. We constructed a set of highly diverse plasmid libraries containing randomly generated sORFs (Fig. 1A). The plasmid used is an in-house construct with a low-copy-number origin of replication (p15A). The synthetic sORFs are under the control of an inducible P_{LlacO} promoter and encode peptides ranging from 10 to 50 amino acids. To avoid premature stop codons, we excluded adenine from the third codon position. This resulted in a 3.9-fold increase in full-length 50-mers compared to NNN (N = A, C, G, and T) repeats (see Fig. S1A in the supplemental material) while still encoding all amino acids with only mildly changed frequencies (Fig. S1B). Two libraries were designed with biased nucleotide compositions encoding either a higher fraction of primordial amino acids (27) (rnd_50b) or hydrophilic amino acids to increase intrinsic disorder and promote functional promiscuity (28) (rnd_50c) (Fig. S1C). The diversity of each library was between 0.3×10^8 and 1.7×10^8, which is the typical upper limit of transformation efficiency. However, it is noteworthy that these libraries only cover a marginal fraction of the complete possible diversity, which is approximately 9.5×10^{33} for a random 50-mer protein. All five libraries were transformed into *Escherichia coli* K-12 strain BW25113 (here designated *E. coli*) and screened for ORFs that increase antibiotic resistance. For this, cells carrying each plasmid library were spread on plates containing antibiotics above the MIC of an *E. coli* strain carrying the empty control plasmid. In total, 5.82×10^8 cells, each express-
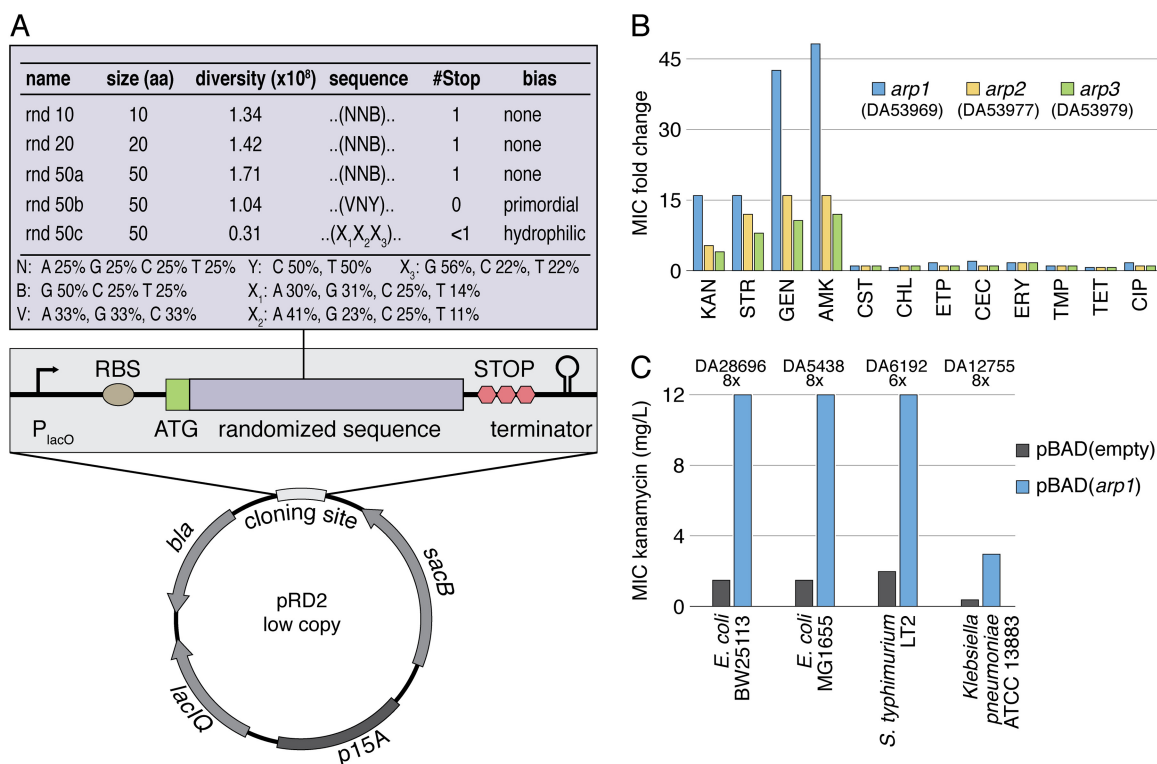
FIG 1 Selection of peptides from randomized sequence libraries that increase antibiotic resistance. (A) Expression vector and sequence libraries. Randomized sequences were cloned in the low-copy-number expression vector pRD2 with flanking transcriptional and translational start and stop sequences. The five libraries encode peptides comprised of 10 to 50 amino acids, with a total diversity of 5.82×10^8. Repeats of NNB encode random amino acid compositions, while VNY repeats are biased toward primordial amino acids (Ala, Asp, Glu, Gly, Ile, Leu, Pro, Ser, Thr, and Val) and $X_1X_2X_3$ repeats enrich for hydrophilicity to promote intrinsic disorder and functional promiscuity. (B) Cross-resistance of Arp1, Arp2, and Arp3 to different antibiotics: kanamycin (KAN), streptomycin (STR), gentamicin (GEN), amikacin (AMK), colistin (CST), chloramphenicol (CHL), ertapenem (ETP), cefaclor (CEC), erythromycin (ERY), trimethoprim (TMP), tetracycline (TET), and ciprofloxacin (CIP). A complete strain list, including custom identifiers (DA numbers), can be found in Table S1. Absolute MIC values are listed in Table S2. (C) Functionality of Arp1 in different bacterial strains and species. DA number and fold increase are indicated above the graph.

ing a unique sORF, were tested on 10 different antibiotics representing 9 antibiotic classes (kanamycin, trimethoprim, tetracycline, amikacin, fosfomycin, chloramphenicol, rifampin, cefaclor, nitrofurantoin, and ciprofloxacin) at three different concentrations, ranging from 2-fold to 16-fold of the respective MIC. Selection on plates was chosen over bulk selection in liquid culture to reduce the influence of selective sweeps and also allow the isolation of variants with low fitness.

Isolation of peptides increasing aminoglycoside resistance. We isolated three unique sORFs (designated Arp, for aminoglycoside resistance peptide) on plates containing kanamycin. Two of these sORFs were also selected on amikacin plates. To ensure that the observed phenotypes were caused by expression of the sORFs rather than by unintended resistance mutations in the chromosome, we recloned all inserts into an empty plasmid and confirmed the phenotypes in the absence and presence of the inducer isopropyl-β-D-thiogalactopyranoside (IPTG). Furthermore, we showed that the function was dependent on the encoded protein rather than the RNA transcript by demonstrating that variants without start codons or introduced frameshift mutations lost activity, and that inserts maintained activity when the identical amino acid sequence was recoded with a different nucleotide sequence (Table S3). To determine the resistance levels provided by Arp1 to Arp3 (Arp1-3) as well as their ability to confer cross-resistance to other antibiotics, we measured the MICs of 12 different antibiotics representing the major classes (Fig. 1B and Fig. S2). All peptides increased the MICs of the aminoglycosides kanamycin, amikacin, gentamicin, and streptomycin. Arp1 had the strongest effect, increasing amikacin resistance 48-fold compared to that of a strain
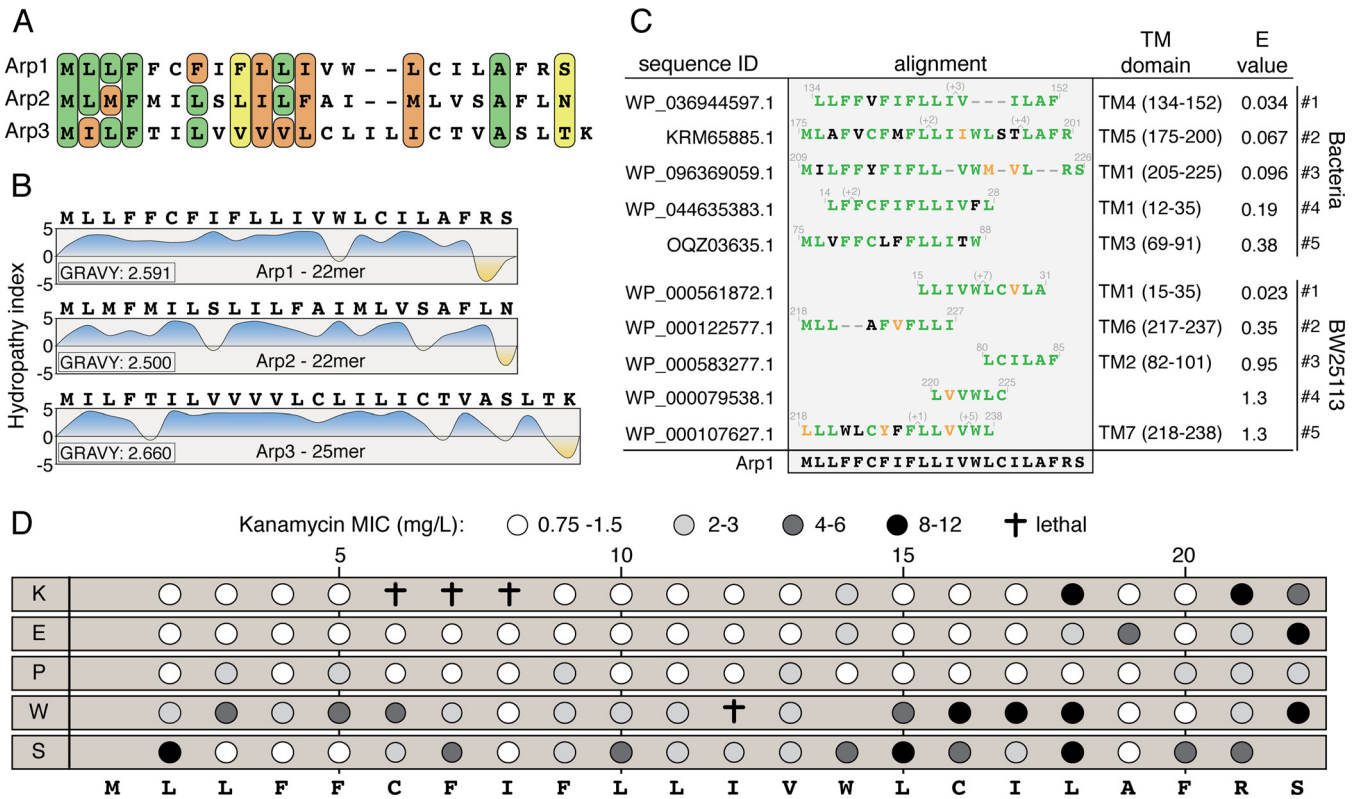
FIG 2 Sequence analysis of the resistance peptides. (A) Alignments of the selected peptides show sequence similarities for specific positions and peptide length. Green backgrounds indicate identical amino acids, while orange and yellow indicate strongly and weakly similar properties, respectively. (B) Hydropathy profiles (30). Averages of hydropathy scores (GRAVY) are indicated for each peptide. (C) BLAST-hit analysis of Arp1 against bacteria and *E. coli* K-12 BW25113. Identical amino acids are marked in green, and similar amino acids are marked in orange. (D) Site-directed mutagenesis of Arp1. The white-to-black scale indicates the level of kanamycin resistance conferred by the mutant peptide. †, the mutant peptide is toxic when expressed.

with an empty control vector. The susceptibility to other antibiotic classes was not altered by overexpression of any of the peptides, demonstrating that the effect was specific to one antibiotic class. We also tested the dependence of the phenotype on the expression vector, copy number, P_{LlacO} promoter, and host background. Overexpression of Arp1 caused a similar increase in resistance when expressed from a single chromosomal copy under the control of P_{LlacO} or when placed under an arabinose-inducible promoter on a medium-copy-number plasmid (pBAD18 backbone [29]). Increased resistance was also observed when Arp1 was overexpressed in the *E. coli* K-12 strain MG1655, *Salmonella enterica* Typhimurium LT2, and *Klebsiella pneumoniae* ATCC 13883 (Fig. 1C), showing that the mechanism of action is independent of the expression system and the specific bacterial strain/species.

Arp1-3 are small, hydrophobic membrane peptides. The sORFs originated from the rnd_50A (Arp1 and Arp3) and rnd_50B (Arp2) libraries, but sequence analysis revealed that all three sORFs encoded truncated peptides (due to inadvertent introduction of stop codons or frameshifts during synthesis/cloning), resulting in two 22mers (Arp1 and Arp2) and one 25mer (Arp3). Besides the similarities in length, all peptides had a high content (80 to 86%) of hydrophobic residues (Fig. 2A and B). The C terminus contained hydrophilic amino acids in all three peptides, suggesting a functional role in the observed phenotype. The very high average hydropathy (GRAVY) scores of all three peptides (≥2.5) indicate a probable localization in the membrane (30), which was also predicted by the transmembrane α-helix prediction tool TMHMM with high confidence. Additionally, a BLASTp search for Arp1 against all bacteria or only the *E. coli* K-12 strain BW25113 showed that 9 out of the top 10 hit sequences were located in transmembrane domains (Fig. 2C).

Since all three peptides showed strong similarities in sequence composition, namely, a highly hydrophobic sequence followed by a hydrophilic C terminus, we conducted a site-directed mutagenesis experiment based on Arp1 to identify functionally important positions (Fig. 2D). Each residue was replaced with a positively charged (lysine), negatively charged (glutamic acid), α-helix breaker (proline), bulky (tryptophan), and hydrophilic (serine) amino acid. The 103 variants generated were then analyzed for their effect on the MIC of kanamycin to test which residues in Arp1 are required for the observed resistance phenotype and which residues can be varied. Changes to charged amino acids caused a complete loss of function in most cases, except for five variants with C-terminal substitutions that maintained kanamycin MICs of over 4 mg/liter (compared to MICs of the empty vector control of 0.75 mg/liter and Arp1 of 12 mg/liter). Changes to proline resulted in 15 nonfunctional variants and 6 variants with a MIC of 2 to 3 mg/liter kanamycin. In contrast, the peptide was more tolerant to replacements with tryptophan and serine, with only 4 variants each losing complete functionality.

The strong sensitivity to charged amino acids and to the α-helix breaker proline further supports the predicted transmembrane α-helix (31) and that the membrane insertion is important for the increase in resistance. To show the membrane localization of the identified peptides, we fused a hexahistidine tag to the C terminus of Arp1 and conducted an electron microscopic analysis of immunogold-labeled samples using a His tag antibody (Fig. 3A). The micrographs revealed large inclusion bodies in the cells, a common artifact when overexpressing hydrophobic peptides (32, 33), which are strongly decorated with antibody-attached gold particles. Outside these aggregates the antibodies localized both in the cytoplasm as well as in the cell periphery, indicating that a major fraction of free Arp1 (outside the inclusion bodies) does populate the membrane. We observed strong heterogeneity of gold particle distribution among different cells, which indicates stochastic differences of expression and stability of Arp1 among different cells. Additionally, it is reasonable to assume that since the phenotype was selected in a single step, it is imperfect compared to most genes that have evolved for long times to conduct specific functions.

Resistance is mediated via reduced aminoglycoside uptake due to membrane depolarization. How can a peptide with a short transmembrane α-helix cause an increase in resistance to a whole class of antibiotics? It is known that aminoglycoside uptake relies on the membrane potential and that resistance can be conferred by reducing the membrane potential (34, 35), and conceivably these peptides could act in a similar way. To test this idea, we determined if overexpression of Arp1 causes a disruption of the electrochemical gradient by using the membrane potential-sensitive probe bis-(1,3-dibutylbarbituric acid)trimethine oxonol [DiBAC$_4$(3)] (Fig. 3B). The protonophore carbonyl cyanide m-chlorophenyl hydrazine (CCCP), which collapses the membrane potential, was used as a positive control. After CCCP treatment, a clear shift to higher fluorescence occurred due to an increased uptake of DiBAC$_4$(3). This shift in fluorescence was also observed upon induction of Arp1 expression from the high-copy-number plasmid (pCA24N), showing that Arp1 insertion in the membrane causes a reduction of membrane potential. Additionally, we observed that expression of Arp1-3 results in the formation of subpopulations of cells with different resistance levels (Fig. S2), which might be due to cellular heterogeneity in expression and/or solubility of the peptides. The resulting difference in amounts of free peptides that can insert into the membrane could cause differences in membrane depolarization across the population and thereby generate cells with different resistance levels. This heterogeneity decreases the sensitivity of the conducted assays, since only a fraction of the population shows the investigated phenotype (Fig. S2). We also determined if the uptake of streptomycin was reduced when Arp1 is expressed. As shown in Fig. 3C, levels of intracellular radiolabeled dihydrostreptomycin were significantly lower in a strain expressing the Arp1 (Fig. 3C). The observed decrease in intracellular streptomycin is relatively small, although statistically significant ($P = 0.008$), but reduced streptomycin uptake and increased resistance does not necessarily correlate in a proportional way.
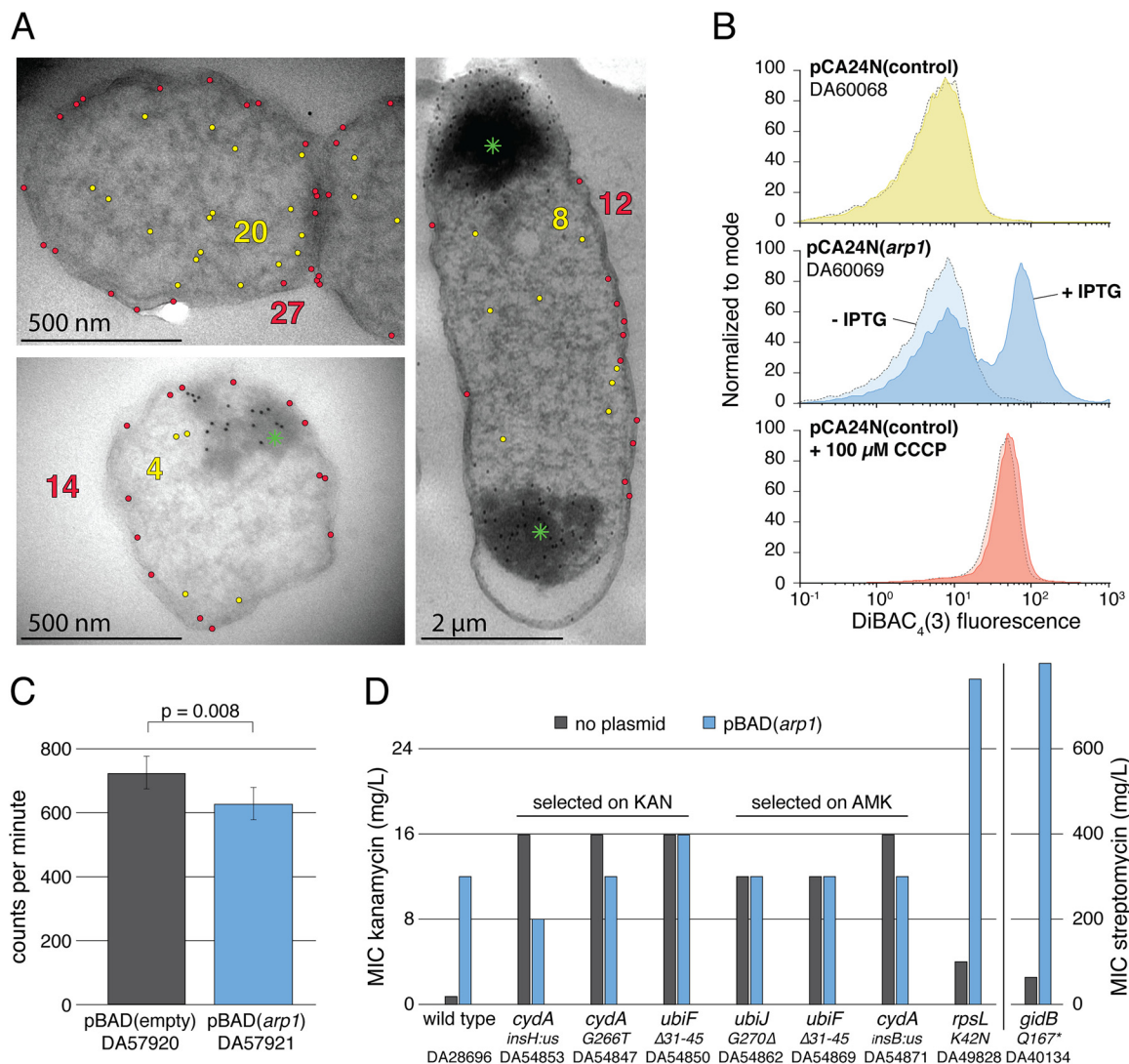
FIG 3 Localization and mechanism of action of Arp1. (A) Transmission electron microscopy of immunogold-labeled samples expressing His-tagged Arp1. The green asterisks indicate inclusion bodies at the cell poles. Gold particles outside the inclusion bodies are colored yellow or red if they are located in the cytoplasm or cell periphery, respectively. Original micrographs can be found in Fig. S5A. (B) Membrane potential determined by the slow-response-potential-sensitive probe DiBAC$_4$(3). The protonophore CCCP was used as a positive control causing a complete collapse of membrane potential. Shown is a representative data set from three independent assays (Fig. S5B). (C) Uptake of radiolabeled dihydrostreptomycin in cells carrying a control plasmid or overexpressing Arp1. Bars are the means from five biological replicates, and error bars represent standard deviations. (D) Additivity of the peptide-mediated resistance phenotype with various mutants of *E. coli* and *S.* Typhimurium LT2. The relevant genetic changes of the six chromosomal mutants are indicated below. A strain list, including custom identifiers (DA numbers) and complete genotypes of all mutants, can be found in Tables S1 and S4.

Thus, a small effect on uptake can have a large effect on resistance, as has been observed previously (36).

To determine if overexpression of Arp1 in strains with elevated aminoglycoside resistance can cause an additional increase in resistance levels, we selected chromosomal mutants with the empty vector control strain on the same concentrations of kanamycin and amikacin that were used for selection of the *de novo* resistance genes. We isolated 24 mutants (Fig. S3) and determined the resistance level to both antibiotics. Strikingly, all mutants exhibited approximately the same level of resistance as that observed for Arp1 overexpression. Furthermore, the selected mutants exhibited cross-resistance to both kanamycin and amikacin, indicating an identical mechanism of action. Whole-genome sequencing of six clones showed that all had mutations in genes associated with electron transport and generation of membrane potential (Table S4). Overexpression of Arp1 in these mutants did not cause any further increase in kana-

mycin or amikacin resistance, indicating that the chromosomal mutations and peptides act by a similar mechanism (i.e., reduction in membrane potential) and that Arp1-dependent depolarization causes the highest resistance increase achievable by this mechanism. In contrast, overexpression of Arp1 in strains with mutations in the antibiotic target (*rpsL* K42N) or a target modification enzyme (*gidB* Q167*) elevated resistance levels beyond the level that was observed for mutants with disrupted membrane potentials (Fig. 3D).

The sORFs selected from our randomized sequence libraries were shown to (i) localize in the cell membrane, (ii) disrupt membrane potential, (iii) consequently increase aminoglycoside resistance to the maximum amount that can be achieved by this resistance pathway, and (iv) show additive effects when combined with resistance mechanisms that are unrelated to the membrane potential (i.e., target modification). The results presented here are fully compatible with the depolarization model, but it should be pointed out that natural small proteins, such as MgtS (37), AcrZ (38, 39), and others (24), have been shown to regulate membrane-located transporters, including efflux pumps such as AcrAB (39). Thus, hypothetically Arp1-3 could confer increased resistance by upregulating an efflux pump. However, this explanation is unlikely, since the peptide-conferred resistance is completely specific for aminoglycosides, with no effect on more typical hydrophobic efflux pump substrates such as tetracyclines, chloramphenicol, and fluoroquinolones.

Alternatively, albeit unlikely, the increased resistance to aminoglycosides could be based on a physiological response to high expression of foreign peptides. As seen in Fig. 3A, production of Arp1 causes the formation of inclusion bodies. This will likely result in a disruption of the normal cell physiology and cause the activation of stress responses (40, 41). For example, the heat shock response induces upregulation of various chaperones and proteases that could themselves influence the formation of membrane potential. However, to our knowledge, no stress response exists in *E. coli* that results specifically in aminoglycoside resistance. Furthermore, if the effect was due to formation of inclusion bodies and generation of a stress response, one would expect that overexpression of many peptides would generate resistance (which we do not observe), since overexpression is often associated with formation of inclusion bodies. Finally, if the resistance was due to the reduced growth rate associated with peptide production, one would expect that unrelated mutations that reduce growth rate should induce aminoglycoside resistance as well. However, analysis of aminoglycoside resistance in several unrelated mutants with reduced growth rate had no effect on aminoglycoside resistance (42). In summary, it seems very unlikely that the aminoglycoside resistance is caused by an indirect physiological effect.

To further examine if the loss of membrane potential is a direct effect of peptide insertion into the membrane, we used a purified *in vitro* system to measure the effect of synthesized Arp1 peptide and a nonfunctional (i.e., confers no resistance when expressed *in vivo*) control variant (Arp1-L11K) (Fig. 2D) on the activity of the proton pump bovine cytochrome *c* oxidase (Cyt*c*O) reconstituted into proteoliposomes (Fig. 4A; also see schematic in Fig. S4). When the enzyme is active, it reduces oxygen to water, using protons from inside the vesicles; in addition, protons are pumped across the membrane. A proton electrochemical gradient thereby is rapidly built up across the liposome membrane, slowing down further activity, and the resulting slow steady-state rate is called the coupled rate. When ionophores (valinomycin, a K^+ ionophore, and/or FCCP, a protonophore similar to CCCP) are added, this gradient is partly or fully dissipated, and the so-called uncoupled rate is reached. Respiratory control ratios, dividing the uncoupled and the coupled rates (RCR; see reference 43), are a measure of the degree to which additives can uncouple the built-up potential and generally gives a value of 3 to 5 for full uncoupling by FCCP in our experimental system (see Fig. S4 for a schematic of the experiment and an example of a raw data trace). Thus, we determined the RCRs after adding the Arp-1 and Arp1-L11K peptides, and as positive controls we added valinomycin and FCCP in addition to the peptides (Fig. 4B and Fig. S4C). Increasing amounts of Arp1 had mild uncoupling effects (i.e., stimulated the
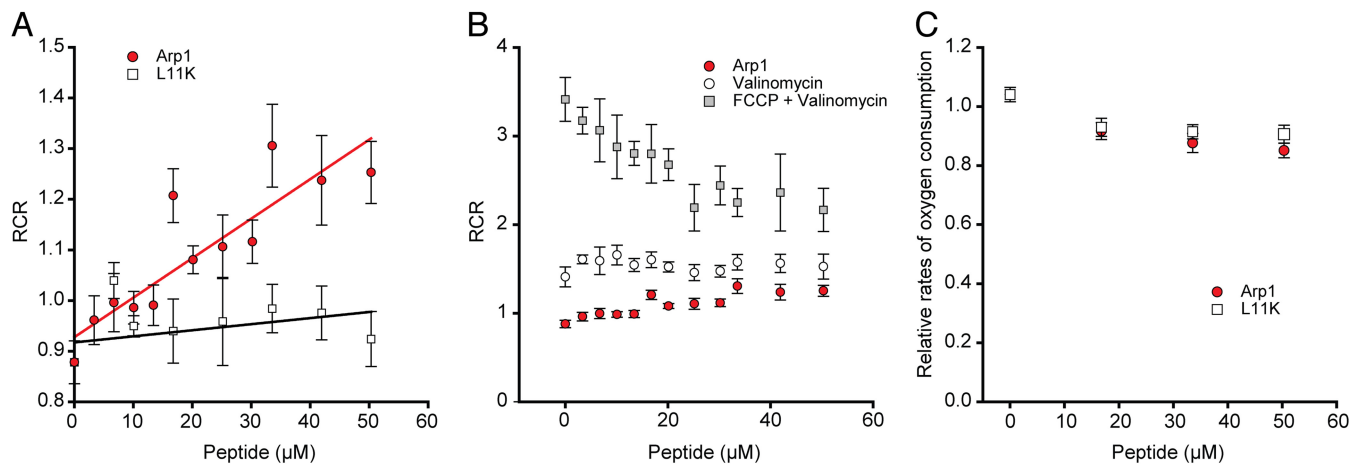
FIG 4 *In vitro* studies with synthesized Arp1 and Arp1-L11K peptides. (A) The steady-state activity of bovine Cyt*c*O in liposomes after addition of Arp1 (red circles) and L11K (white squares) divided by the coupled (original) rate gives the RCR. The data point at 0 μM peptide indicates the value obtained with only the solvent DMSO added. (B) RCRs for Arp1 with addition of valinomycin (6 μM, white circles) and FCCP (21 μM, gray squares) in addition to the peptide. (C) Relative oxygen consumption rates for Cyt*c*O in DDM detergent micelles with addition of either Arp1 (red circles) or L11K (white squares). All data points in panels A to C are averages with standard errors, $n > 3$.

activity of Cyt*c*O), with RCRs increasing up to around 1.4 (Fig. 4A and B). Valinomycin addition led to further rate increase, and FCCP fully collapsed the gradient, giving the maximum Cyt*c*O activity. The L11K variant of the peptide did not show any uncoupling effect (Fig. 4A and Fig. S4C), even at high concentrations. The maximal RCR after FCCP addition decreased at higher peptide concentrations, indicating that it affects the Cyt*c*O negatively (Fig. 4B and Fig. S4C). We further investigated if the peptides indeed affect the enzyme activity *per se* by measuring their effect on the Cyt*c*O activity in detergent. As seen in Fig. 4C, both peptides had small negative effects on the Cyt*c*O activity in detergent. Thus, the activity increase observed in the Cyt*c*O liposomes when the Arp1 peptide is added is most likely due to an uncoupling effect, as the direct effect of peptide itself on the oxidase is negative. In summary, these results show that Arp1 is capable of partly uncoupling a proton electrochemical gradient, and that the likely cause for this uncoupling is an increase in the permeability to protons. That the peptide does not completely dissipate the gradient is important, since full dissipation would stop bacterial growth.

Increased aminoglycoside resistance is specific for Arp1-3. How easily can peptide-mediated membrane depolarization be achieved? Theoretically, any peptide or domain that spans the membrane creates a lesion in the phospholipid bilayer that could cause leakage of protons or ions. However, the site-directed mutagenesis of Arp1 (Fig. 2D) showed that the function is sequence specific. Not only did most changes to charged amino acids or prolines cause a complete loss of function but also substitutions with a weak hydrophilic residue (serine) or even another hydrophobic amino acid (tryptophan) were detrimental in almost all variants. Additionally, four variants (C6K, F7K, I8K, and I12W) were toxic when expressed. While we do not know the cause of the cytotoxicity, it is possible that the substitutions cause a lethal disruption of membrane integrity. This shows that not only functionality but also viability is quickly lost upon changes of the specific sequences that were isolated.

Furthermore, the alignment of Arp1-3 suggests that, besides the requirement of a strong hydrophobicity followed by a hydrophilic C terminus, certain positions are conserved. For example, Arp1-3 all contained a phenylalanine at position 4, which indicates a functional role of this residue possibly promoting dimerization (44). To further elucidate the sequence requirements and conservation, more variants of this peptide class need to be isolated to generate a higher statistical power which allows a more detailed analysis. To further show that the phenotype is specific to Arp1-3, we overexpressed three endogenous *E. coli* peptides, TisB, IbsC, and LdrD, that are similar

to Arp1-3 in regard to size and hydrophobicity. The 29-mer toxin TisB, which is encoded by the *E. coli* K-12 genome, is part of a toxin-antitoxin system, and it inhibits growth when overexpressed. It has been shown to form channels that allow ion leakage and cause membrane depolarization (45, 46). However, TisB expression from a pBAD promoter (arabinose concentrations ranging between 0.0125 and 0.2%) conferred no increased resistance when expressed at viable levels (high expression levels of TisB were lethal). Similarly, no increased resistance could be observed when overexpressing the toxins IbsC and LdrD. We further tested if overexpression of other transmembrane domains that are not part of toxin-antitoxin systems could cause aminoglycoside resistance. For this, we chose 10 representative transmembrane domains from 5 genes with different physiological functions (Table S5) and overexpressed them from the pBAD18 vector. None of the variants increased resistance to kanamycin or amikacin. These results show that Arp1-3 are facilitating a unique function in causing depolarization to an extent that generates antibiotic resistance but still maintains cell viability.

Beneficial and deleterious effects conferred by membrane depolarization. A reduction of membrane potential is expected to be deleterious for bacteria in most environments, since the electrochemical gradient across the inner membrane promotes various physiological functions, such as ATP synthesis, nutrient flux, and pH homeostasis. However, under certain conditions, mutants with a defective electron transport chain have a strong selective advantage. For example, small-colony variants (SCVs), which often result from mutations in genes associated with electron transport, have been isolated from various infections and species (47). Additionally, SCVs can be rapidly selected *in vitro*, for example, under antibiotic pressure (48, 49), cold stress (50), hydrostatic pressure (51), or low pH (52). The reduced membrane potential and low growth rate of electron transport-defective SCVs offer various advantages to the bacterial cell. First, the uptake of toxic compounds that rely on an active membrane potential, like aminoglycoside antibiotics, is reduced and the cell will be able to persist and/or proliferate. Additionally, the reduced growth rate and, thus, reduced speed of cell division will dampen the cytostatic effect of β-lactam antibiotics (53). Besides the selective advantage in the presence of antibiotics, SCVs can have an increased survival rate *in vivo*. These mutants are often attenuated, for example, by a decreased expression of immunogens, which causes a reduction in immune response and therefore a higher survival rate inside the host (47). Other clinically important resistance mechanisms, in particular target alterations/modifications, are typically associated with a reduction in bacterial fitness in the absence of antibiotics (54). For example, mutations in *rpsL*, which confer high-level streptomycin resistance, can cause a reduction in fitness of up to 26% (55). Similarly, expression of the 16S rRNA methyltransferase ArmA or mutations in the 16S rRNA, both increasing aminoglycoside resistance, are associated with a significant reduction in fitness (56, 57).

Compared to the fitness costs discussed above, expression of Arp1 in *E. coli* causes a similar reduction in fitness, and during exponential growth in nutrient-rich growth medium fitness is reduced by 36%. Under these conditions, clones expressing this peptide would be rapidly outcompeted, and fixation of the new function would be unlikely. However, in the presence of aminoglycoside antibiotics it confers a strong fitness advantage, allowing cell proliferation at concentrations which inhibit growth of the wild type. Thus, expression of these peptides offers, similar to SCVs, a conditional benefit to the bacterial cell. Interestingly, membrane depolarization via peptides has a distinct advantage over electron transport-defective mutants, in that expression and, therefore, the phenotype could evolve to become inducible, for example, by evolution of a regulated promoter or a regulatory small RNA. It is possible that endogenous, tightly regulated toxic peptides like TisB have evolved in this manner to induce growth arrest in environments where an active metabolism and cell division are detrimental (58, 59). Furthermore, similar to what has been observed for costly antibiotic resistance mutations, compensatory evolution could reduce the deleterious fitness effects and increase the chance of fixation.

Various studies have identified potential *de novo* genes in all domains of life based on comparative genomics (13, 60, 61), and experimental studies have shown that new functions can be selected by randomization of amino acid stretches located in preexisting structural scaffolds (62, 63). Recently, even variants with enzymatic functions could be isolated from a four-helix-bundle template (64). These studies demonstrate persuasively that new functions can arise from localized randomization, but they do not sufficiently address the functionality encoded in completely random sequence pools that do not rely on preexisting scaffold. Another study using unbiased sequence libraries to select ATP binding activity *in vitro* enriched four out of 6×10^{12} proteins with the selected activity (65). However, this experimental setup does not incorporate mRNA and protein stability *in vivo* (i.e., in a living bacterial cell) or that a beneficial function was conferred to the host. Similarly, a study utilizing biased expression libraries composed of glutamine, asparagine, and arginine residues *in vivo* showed that a large fraction is stably expressed, forms α-helices, and is protected from proteolytic degradation (66), but these studies did not reveal any beneficial selectable function for the bacterial cell. Here, we present experimental selection for peptides that confer a beneficial functionality *in vivo* and that were generated from randomized nucleotide sequences, supporting the idea that expression of randomly occurring sORFs can serve as a substrate for evolution of *de novo* genes. Further work is needed to determine what fraction of sequence space can generate such beneficial functions and which factors constrain how rapidly a proto-peptide/protein can be fine-tuned in its cellular function and how the initial fitness costs can be reduced. It would be interesting to examine if a cell can evolve to reduce the deleterious side effects of Arp expression. Evolution of such variants would likely require multiple genetic alterations which could be achieved with large bacterial populations during long-term evolution experiments in laboratory settings.

Finally, an interesting question is whether Arp-like proteins could provide aminoglycoside resistance in real-life clinical isolates. To our knowledge no one has searched for such functions, but we find it unlikely that they would be significant contributors, because there are other mechanisms that are easier to acquire and less deleterious to the cell. However, it is interesting to speculate that if Arps were present on plasmids and acquired regulatory elements, this could allow horizontal transfer and a conditionally inducible temporary reduction of membrane potential and associated resistance.

MATERIALS AND METHODS

Strains and growth conditions. Most strains used in this study were derived from *Escherichia coli* K-12 BW25113 (designated *E. coli* throughout the text) and are listed in Table S1 in the supplemental material. If not indicated otherwise, Muller Hinton broth (MHB) (Becton, Dickinson) was used for growth in liquid cultures and Muller Hinton agar (MHA; MHB supplemented with 15 g/liter agar) for growth on plates. To maintain the plasmid pRD2 or pBAD18, the growth media were supplemented with 100 mg/liter ampicillin or 15 mg/liter chloramphenicol, respectively. Induction of expression was performed in the presence of 1 mM IPTG (pRD2) or 0.2% arabinose (pBAD18). For counterselection of *sacB*, no-salt lysogeny agar plates (10 g/liter tryptone, 5 g/liter yeast extract, 15 g/liter agar, and 50 g/liter sucrose) were used. Recovery after transformation was done in SOC medium (20 g/liter tryptone, 5 g/liter yeast extract, 0.5 g/liter NaCl, 10 mM MgCl$_2$, 0.25 mM KCl, and 4 g/liter glucose). For washing and dilution of cells, phosphate-buffered saline (PBS; 8 g/liter NaCl, 0.2 g/liter KCl, 1.44 g/liter Na$_2$HPO$_4$, and 0.24 g/liter KH$_2$PO$_4$) was used.

Library construction. The randomized sequences were ordered as oligonucleotide pools from MWG Eurofins, IDT, and BioSyn. The oligonucleotide design for all libraries was similar, with a random sequence (composition is depicted in Fig. 1) flanked by fixed sequences upstream (TACGCTGGATCCATG) and downstream (TAACTAAGTAACTGCAGGCTTA). The oligonucleotides were complemented in a 1-cycle PCR using Phusion DNA polymerase and the complementation primer TAAGCCTGCAGTTACTTAG. The protocol consisted of 3 min of denaturation at 98°C followed by 5 min of annealing at 55°C and a 5-min elongation at 72°C. The plasmid pRD2 was isolated using the NucleoBond Xtra Midi kit (Macherey-Nagel) according to the manufacturer's recommendations. To remove residual chromosomal DNA, the plasmid preparation was treated with Plasmid-Safe ATP-dependent DNase (Epicentre). The plasmid and double-stranded library were subsequently digested using the FastDigest restriction enzymes BamHI and PstI according to the manufacturer's recommendations (ThermoFisher) and purified using a GeneJET gel extraction micro kit (ThermoFisher). The inserts were ligated into the vector using Ready-To-Go T4 DNA ligase (GE Healthcare) at a molar ratio of 4 inserts per vector. Purified ligations were mixed with electrocompetent NEB 5-alpha *E. coli* (New England BioLabs), transferred to an electroporation cuvette

(1 mm gap), and electroporated using a Gene Pulser Xcell (Bio-Rad) at 1.8 kV, 400 Ω, and 25 μF. Cells were immediately recovered in 1 ml prewarmed SOC medium at 37°C with shaking at 200 rpm. After 1 h, a dilution series was plated on MHA containing 100 mg/liter ampicillin to determine the transformation efficiency (which equals the diversity of the cloned library). The remaining culture was split up into 300-μl aliquots, spread on MHA containing 100 mg/liter ampicillin, and incubated overnight at 37°C. The cells were scraped from the plates and plasmid libraries were extracted using a NucleoBond Xtra Midi kit (Macherey-Nagel).

To generate electrocompetent cells of *E. coli* BW25113, 1 ml culture was grown in MHB overnight and diluted 1:100 in prewarmed MHB. Cells were grown to an optical density at 600 nm (OD$_{600}$) of 0.2 and rapidly cooled down in ice water for 10 min. The cells were pelleted by centrifugation at 4,000 × *g* and 4°C for 10 min and resuspended in ice-cold 10% glycerol. After 2 additional washing steps the pellet was resuspended in 10% glycerol, and 40 μl competent cells was premixed with approximately 200 ng of plasmid library in a chilled Eppendorf tube. The cell-DNA mixture was transferred to an electroporation cuvette and electroporated. Electroporation, recovery, and plating were performed as described above. The scraped cells were not subjected to plasmid extraction but instead mixed 1:1 with 50% glycerol and cryopreserved at −80°C in 100-μl aliquots, which contained approximately 2 × 10^{10} CFU per ml.

Selection of antibiotic-resistant ORFs. Six aliquots containing the different libraries and an empty vector control strain were thawed on ice, and 50 μl was transferred to 20 ml MHB containing 100 mg/liter ampicillin and 1 mM IPTG. Cultures were incubated for 3 h, and 200 μl was spread on MHA containing 100 mg/liter ampicillin, 1 mM IPTG, and the antibiotics used for screening of functional ORFs. Plates were incubated overnight at 37°C. Since selection was typically performed slightly above the MIC of an empty vector control strain, background colonies were frequently encountered due to inoculum effects or chromosomal mutants with elevated MICs. To distinguish between these background effects and increased resistance mediated by *de novo* ORFs, we scraped all colonies of the initial selection plates and extracted the plasmids using EZNA plasmid DNA minikit I (Omega) according to the manufacturer's recommendations. The plasmids were then retransformed into *E. coli* BW25113 as described previously, recovered for 2 h in MHB containing 1 mM IPTG, and plated on MHA with 100 mg/liter ampicillin, 1 mM IPTG, and the screening antibiotic at the same concentration as that used for the initial selection. If the amount of colonies was similar to that in the initial selection, plates were discarded and this combination of antibiotic/concentration was no longer pursued. In cases were increased resistance was mediated by a plasmid-based ORF, the plasmid was enriched, resulting in a much higher abundance of resistant colonies for that particular selection. From these plates, we isolated individual clones and extracted the plasmid using EZNA plasmid DNA minikit I (Omega) according to the manufacturer's recommendations. Plasmids were sequenced (primer, CCACCTGACGTCTAAGAA) and retransformed into *E. coli* BW25113. After recovery, the cells were spread on MHA with 100 mg/liter ampicillin. An individual clone was isolated and cryopreserved in 10% dimethyl sulfoxide (DMSO). This clone was subsequently tested for growth on plates containing the antibiotic from the initial selection. Additionally, we confirmed that growth was not observed if no IPTG was present in the plate to unambiguously show that the phenotype is caused by the randomly generated ORF in the plasmid under the inducible P$_{LlacO}$ promoter.

Cloning using annealed oligonucleotides. The insert of interest was designed *in silico*, and two complementary oligonucleotides with the desired overhangs after annealing were ordered from Eurofins or Sigma-Aldrich. Oligonucleotides were diluted to 1 pmol/μl in annealing buffer (10 mM Tris, pH 7.5, 100 mM NaCl, 100 μM EDTA) and mixed 1:1 in an Eppendorf tube to a final volume of 100 μl. The mixture was placed into boiling water and allowed to cool down to room temperature. The annealed oligonucleotides were directly used for cloning into the desired vector with an insert-to-plasmid ratio of 3:1. Ligations were performed using Ready-To-Go T4 DNA ligase (GE Healthcare) according to the manufacturer's recommendations. After overnight incubation at 16°C, the reactions were cleaned and transformed as previously described.

MIC determination. MICs were determined using Etest strips (bioMérieux). Overnight cultures of the strains of interest were diluted 1:20 in PBS and spread on an MHA plate containing 100 mg/liter ampicillin and 1 mM IPTG where necessary with a cotton swab. Etest strips were applied in the middle of the plates and incubated at 37°C overnight. MIC values were scored after 20 h of incubation. Colistin MICs were determined using microdilution. For this, bacteria were grown on MHA with 100 mg/liter ampicillin, and 1 to 2 colonies were dissolved in 1 ml PBS to approximately 0.5 McFarland. The cell suspension was diluted 1:100 in Muller Hinton 2 broth (cation adjusted), and 50 μl was mixed with 50 μl Muller Hinton 2 broth containing 0.4% arabinose, 30 mg/liter chloramphenicol, and a 2-fold dilution series of colistin. The assay was performed in round-bottomed 96-well plates at 37°C overnight in a stationary position. The MIC was determined as the concentration where no growth could be observed and is based on a minimum of two biological replicates. If the determined values of the replicates were not identical, a third replicate was performed.

Growth rate determination. Exponential growth rates were determined with a Bioscreen C reader (Oy Growth Curves AB Ltd., Finland). For this, cultures were grown overnight in MH supplemented with 100 mg/liter ampicillin and diluted 1:1,000 in fresh medium. From the dilution, two 300-μl aliquots were transferred into honeycomb plates. The plates were incubated in a Bioscreen C reader for 16 h at 37°C with continuous shaking. Every 4 min the OD$_{600}$ was measured, and the interval between 0.024 and 0.09 was used to determine the slope of the growth curve. All values are based on at least 4 biological replicates with two technical replicates each and are represented as relative to the growth of an empty vector control strain.

Quantification of subpopulations with various resistance levels. Overnight cultures were diluted in PBS to 3 × 10^3 CFU/ml. One hundred μl of the cell suspension was plated on MHA plates containing

0, 0.75, 1.5, 3, and 6 mg/liter kanamycin. Additionally, all plates contained either 100 mg/liter ampicillin and 1 mM IPTG for strains carrying pRD2 variants or 15 mg/liter chloramphenicol and 0.2% arabinose for strains carrying pBAD18 variants. The plates were incubated overnight and colonies were counted. The frequencies of the subpopulations that were able to tolerate different kanamycin concentrations are relative to those of the colonies counted on plates without kanamycin.

Sequence analysis. Sequence design and analysis were performed using CLC Main Workbench (Qiagen). The peptides were aligned using Clustal Omega (67) with default parameters (Dealign input sequence, no; MBED-like clustering guide-tree, yes; MBED-like clustering iteration, yes; number of combined iterations, default [0]; maximum guide tree iterations, default; maximum hmm iterations, default; order, aligned). Transmembrane helices were predicted using the web tool TMHMM (68). BLAST searches were conducted using NCBI pBLAST with adjusted parameters for short input sequences (expect threshold, 200,000; word size, 2; matrix, PAM30; compositional adjustments, no adjustment). Grand average of hydropathy (GRAVY) values were determined with ExPASy ProtParam (69).

Site-directed mutagenesis. The site saturation mutagenesis was performed with GeneArt (Thermo-Fisher) based on the sequence of Arp1, where each position in the wild-type sequence was replaced by the amino acids lysine (K), glutamic acid (E), proline (P), tryptophan (W), and serine (S). All variants were subcloned into our modified expression vector, pBAD18-RND, using the AgeI/HindIII cloning sites, transformed into E. coli, and delivered as a total pool glycerol stock. The stock was plated out on MHA with 15 mg/liter chloramphenicol, and single colonies were restreaked on identical plates. From each clone, 2-ml overnight cultures were grown and plasmids were extracted using EZNA plasmid DNA minikit I (Omega). The plasmid variants were determined by Sanger sequencing (Eurofins, Germany). Nonredundant variants were transformed into E. coli BW25113 and subjected to MIC determination. A total of 300 clones were isolated from the pool in this manner, and variants that were not present in this set were constructed with a targeted approach using cloning of annealed oligonucleotides as described above.

Immunogold TEM. E. coli cells expressing Arp1-His tag or carrying an empty vector control were fixed for 45 min in 8% paraformaldehyde (PFA) plus 0.02% glutaraldehyde (GA) and MHB (1:1) in a 50-ml vial. The suspension was transferred to Eppendorf tubes (1.5 ml) and spun down, and the supernatant was replaced with 4% PFA plus 0.01% GA and fixed for a further 2 h. The cells were then dehydrated in alcohol and processed for embedding in K4M (Lowicryl), polymerized with UV light at $-35°C$ in 72 h. All steps were performed at room temperature, except where otherwise stated. Ultrathin sections were cut in a Leica UCT Ultramicrotome and collected on gold grids. The immunostaining was done shortly after sectioning to improve access to epitope. Sections on grids were first blocked for 30 min in 2% bovine serum albumin in Tris-buffered saline, pH 7.4 (TBS; 6.05 g/liter Tris, 8.76 g/liter NaCl, pH adjusted with 1 M HCl); this was followed by incubation overnight at 20°C with the primary antibody (THE His tag antibody, monoclonal antibody, mouse; GenScript) diluted 1:5, 1:50, and 1:500 in TBS, pH 7.4. The grids then were rinsed three times in TBS, pH 7.4, followed by once in TBS, pH 8.3. Bound antibodies were localized by incubating the sections for 2 to 3 h on anti-mouse, 12-nm Au particles diluted 1:80. Finally, grids were rinsed three times in TBS, pH 7.4, and sterile H_2O before being stained with 5% aqueous uranyl acetate and lead citrate. The grids were placed on drops of the respective reagent. Between all steps the grids were dried using filter paper. All incubations were carried out at room temperature. Grids were examined by a transmission electron microscope (TEM; FEI Tecnai G2) operated at 80 kV.

Membrane potential assay. Bacteria were grown overnight in MHB supplemented with 15 mg/liter chloramphenicol at 37°C. The following day all cultures were diluted 1:100 in MHB supplemented with 15 mg/liter chloramphenicol and grown to an OD_{600} of approximately 0.3 (time point 0). At the target OD_{600}, IPTG was added to each culture to a final concentration of 1 mM, and cells were incubated for another 60 min to allow expression of proteins (time point 1). At both time points, approximately 10^6 cells were transferred to 1 ml PBS with 10 mg/liter DiBAC$_4$(3) (D8189; Sigma-Aldrich) with or without 100 μM carbonyl cyanide 3-chlorophenylhydrazone (CCCP; Sigma-Aldrich) and incubated for 30 min at room temperature in the dark. Fluorescence was measured with a MACSQuant VYB (Miltenyi Biotec) using the GFP/FITC (green fluorescent protein/fluorescein isothiocyanate) channel and a threshold of 15,000 gated cells. The raw data were analyzed and visualized with FlowJo (BD Bioscience).

Radiolabeled dihydrostreptomycin uptake assay. Overnight cultures were diluted 1:200 in MHB containing 15 mg/liter chloramphenicol and grown to an OD_{600} of 0.6 at 37°C with continuous shaking. After reaching the target OD_{600}, a 10% arabinose solution was added to a final concentration of 0.2% and incubated for 10 min to allow expression. The cultures were then supplemented with tritiated dihydrostreptomycin (Larodan, Sweden) to a final activity of 200 nCi/ml. Control cultures were additionally supplemented with CCCP to a final concentration of 100 μM. After 20 min, 2-ml aliquots of the cultures were transferred to Eppendorf tubes and pelleted by centrifugation at 13,500 $\times$ g for 3 min, and the supernatant was decanted. The pellets were washed twice with PBS. After the last wash, the pellets were dissolved in 1 ml PBS and transferred to 2 ml Optiphase HiSafe 3 scintillation cocktail (PerkinElmer). The samples were measured with a Tri-Carb 2810 TR liquid scintillation analyzer over a duration of 1 min per sample. The counts were normalized to OD_{600} values determined after the 20-min incubation with dihydrostreptomycin to correct for differences in cell density. The counts of the CCCP control were subtracted from those of the other samples to correct for uptake in cells with a completely collapsed membrane potential and binding of dihydrostreptomycin to the cell surface.

Selection of chromosomal mutants. Four independent cultures of E. coli BW25113 carrying the empty pRD2 control vector were grown overnight in MHB with 100 mg/liter ampicillin and plated on MHA with 100 mg/liter ampicillin, 1 mM IPTG, and either 6 mg/liter kanamycin or amikacin. Plates were incubated overnight, and resistant mutants were purified on plates containing the same amount of antibiotic. Overnight cultures were grown in MHB with 100 mg/liter ampicillin and cryopreserved in 10%

DMSO. To test the additivity of the chromosomal mutants with overexpression of Arp1, the empty control plasmid was cured. For this, cultures were grown overnight in MHB without ampicillin and plated on no-salt LA plates containing 5% sucrose for counterselection of *sacB*. Colonies were restreaked on identical plates, and plasmid loss was confirmed by testing for ampicillin susceptibility. Plasmid-free strains were then subjected to transformation as described previously.

Whole-genome sequencing. Bacteria were grown overnight, and 500-μl samples were used for the preparation of genomic DNA using the MasterPure complete DNA and RNA purification kit (Epicentre) according to the manufacturer's recommendations. In brief, cells were pelleted by centrifugation for 2 min at 13,500 × *g*. The supernatant was removed and the pellet resuspended in 300 μl tissue and cell lysis solution (Epicentre) containing 50 μg proteinase K. After 15 min of incubation at 65°C, the sample was cooled to 37°C and 5 μg RNaseA was added, followed by another incubation for 30 min at 37°C. A volume of 175 μl MPC protein precipitation reagent (Epicentre) was added to the lysed sample and vortexed vigorously. Debris was pelleted by centrifugation for 20 min at 13,500 × *g* and the supernatant transferred to an Eppendorf tube. The genomic DNA was precipitated by addition of 500 μl isopropanol and washed twice with 200 μl 75% ethanol by repeated centrifugation. Finally, the DNA was resuspended in water and diluted to a final concentration of 0.2 μg/liter. Illumina's Nextera XT kit was used to generate libraries (2× 300) to be sequenced with MiSeq per the manufacturer's recommendations. All samples were dual-indexed and pooled. Average whole-genome coverage per sample was approximately 40-fold. Analysis of the fastq files obtained from MiSeq sequencing was performed using CLC Genomics Workbench. Single-nucleotide polymorphism (SNP) calling and structural variants were assessed using standard parameters. Raw sequence reads were deposited under BioProject identifier PRJNA460371.

Purification of cytochrome *c* oxidase and proteoliposome preparation. Cytochrome *c* oxidase (Cyt*c*O) was purified from bovine heart mitochondria as described before (70). Enzyme concentration was determined from a reduced-oxidized difference spectrum (400 to 700 nm) using the 603-nm peak and $\varepsilon = 24$ mM^{-1}. Proteoliposomes were prepared essentially as described in reference 71. Briefly, L-α-phosphatidylcholine (PC) from soy beans (Sigma-Aldrich) was homogenized at 20 mg/ml in 20 mM HEPES, pH 7.5, 100 mM KCl, 50 g/liter sucrose. Large liposomes were formed by rapidly freezing the lipid solution for 1 min in liquid nitrogen, followed by thawing on a 30°C heat block with 30 s of vortexing between treatments. This procedure was repeated 5 times, and the liposomes were stored at −80°C. Small unilamellar vesicles (SUVs) were formed by extruding the liposomes with a miniextruder through 100-nm Whatman extruder membranes. Proteoliposomes were formed by adding 0.6% sodium cholate and 1.3 μl Cyt*c*O (at 215 μM) and dialyzing the sample overnight against 1,000 volumes of liposome buffer to remove the detergent. This yields, on average, 2 to 3 Cyt*c*Os per liposome.

Peptide preparation. The synthetic peptides were purchased from Biosynthesis. The dry peptides were suspended at 5 mg/ml in DMSO and sonicated for 5 min in a water sonication bath to make sure they were completely dissolved.

Oxygen consumption assay. Cyt*c*O activity was measured using a Clark-type oxygen electrode (Hansatech) at 25°C. In the measuring chamber, 1 ml 20 mM HEPES, pH 7.5, 100 mM KCl, 0.1 mM EDTA was mixed with 12 mM ascorbate, 70 μM TMPD, and 60 μM cytochrome *c* under constant stirring. Oxygen consumption rates and RCRs were measured essentially as described in reference 43. Briefly, the coupled rate for Cyt*c*O activity was calculated from the slope after addition of 20 μl proteoliposomes by subtracting the baseline slope. Subsequently, 30 μl peptide solution was added and the rate was again calculated (by subtracting the baseline). As positive controls, the electric component of the proton electrochemical gradient was dissipated by addition of 6 μM valinomycin, and the gradient was subsequently fully dissipated by addition of 21 μM FCCP. To test for direct putative effects by the peptides on the Cyt*c*O activity, oxygen consumption rates were also measured for protein solubilized in 0.05% dodecyl-β-maltoside (DDM) detergent.

Data availability. All data necessary to understand and evaluate the conclusions of this study are available in the paper and supplemental materials. Raw sequence reads of the mutants that were whole-genome sequenced are available at the BioProject database under BioProject number PRJNA460371. An annotated sequence file of the expression vector (pRD2) is available at GenBank under the accession number MH298521.

SUPPLEMENTAL MATERIAL

Supplemental material for this article may be found at https://doi.org/10.1128/mBio .00837-19.

FIG S1, PDF file, 0.1 MB.
FIG S2, PDF file, 0.1 MB.
FIG S3, PDF file, 0.1 MB.
FIG S4, PDF file, 0.1 MB.
FIG S5, PDF file, 0.2 MB.
TABLE S1, PDF file, 0.1 MB.
TABLE S2, PDF file, 0.1 MB.
TABLE S3, PDF file, 0.1 MB.
TABLE S4, PDF file, 0.1 MB.
TABLE S5, PDF file, 0.1 MB.

REFERENCES

1. Boto L. 2010. Horizontal gene transfer in evolution: facts and challenges. Proc Biol Sci 277:819–827. https://doi.org/10.1098/rspb.2009.1679.
2. Nasvall J, Sun L, Roth JR, Andersson DI. 2012. Real-time evolution of new genes by innovation, amplification, and divergence. Science 338:384–387. https://doi.org/10.1126/science.1226521.
3. Wu DD, Zhang YP. 2013. Evolution and function of de novo originated genes. Mol Phylogenet Evol 67:541–545. https://doi.org/10.1016/j.ympev.2013.02.013.
4. Tautz D. 2014. The discovery of de novo gene evolution. Perspect Biol Med 57:149–161. https://doi.org/10.1353/pbm.2014.0006.
5. Jacob F. 1977. Evolution and tinkering. Science 196:1161–1166. https://doi.org/10.1126/science.860134.
6. Reinhardt JA, Wanjiru BM, Brant AT, Saelao P, Begun DJ, Jones CD. 2013. De novo ORFs in Drosophila are important to organismal fitness and evolved rapidly from previously non-coding sequences. PLoS Genet 9:e1003860. https://doi.org/10.1371/journal.pgen.1003860.
7. Li D, Dong Y, Jiang Y, Jiang H, Cai J, Wang W. 2010. A de novo originated gene depresses budding yeast mating pathway and is repressed by the protein encoded by its antisense strand. Cell Res 20:408–420. https://doi.org/10.1038/cr.2010.31.
8. Khalturin K, Anton-Erxleben F, Sassmann S, Wittlieb J, Hemmrich G, Bosch TCG. 2008. A novel gene family controls species-specific morphological traits in Hydra. PLoS Biol 6:e278. https://doi.org/10.1371/journal.pbio.0060278.
9. Wu DD, Irwin DM, Zhang YP. 2011. De novo origin of human protein-coding genes. PLoS Genet 7:e1002379. https://doi.org/10.1371/journal.pgen.1002379.
10. Xie C, Zhang YE, Chen J-Y, Liu C-J, Zhou W-Z, Li Y, Zhang M, Zhang R, Wei L, Li C-Y. 2012. Hominoid-specific de novo protein-coding genes originating from long non-coding RNAs. PLoS Genet 8:e1002942. https://doi.org/10.1371/journal.pgen.1002942.
11. Murphy DN, McLysaght A. 2012. De novo origin of protein-coding genes in murine rodents. PLoS One 7:e48650. https://doi.org/10.1371/journal.pone.0048650.
12. Neme R, Tautz D. 2013. Phylogenetic patterns of emergence of new genes support a model of frequent de novo evolution. BMC Genomics 14:117. https://doi.org/10.1186/1471-2164-14-117.
13. Yang Z, Huang J. 2011. De novo origin of new genes with introns in Plasmodium vivax. FEBS Lett 585:641–644. https://doi.org/10.1016/j.febslet.2011.01.017.
14. Donoghue MT, Keshavaiah C, Swamidatta SH, Spillane C. 2011. Evolutionary origins of Brassicaceae specific genes in Arabidopsis thaliana. BMC Evol Biol 11:47. https://doi.org/10.1186/1471-2148-11-47.
15. Dornenburg JE, Devita AM, Palumbo MJ, Wade JT. 2010. Widespread antisense transcription in Escherichia coli. mBio 1:e00024-10. https://doi.org/10.1128/mBio.00024-10.
16. Wade JT, Grainger DC. 2014. Pervasive transcription: illuminating the dark matter of bacterial transcriptomes. Nat Rev Microbiol 12:647–653. https://doi.org/10.1038/nrmicro3316.
17. Ruiz-Orera J, Verdaguer-Grau P, Villanueva-Canas JL, Messeguer X, Alba MM. 2018. Translation of neutrally evolving peptides provides a basis for de novo gene evolution. Nat Ecol Evol 2:890–896. https://doi.org/10.1038/s41559-018-0506-6.
18. Wilson BA, Masel J. 2011. Putatively noncoding transcripts show extensive association with ribosomes. Genome Biol Evol 3:1245–1252. https://doi.org/10.1093/gbe/evr099.
19. Olexiouk V, Van Criekinge W, Menschaert G. 2018. An update on sORFs.org: a repository of small ORFs identified by ribosome profiling. Nucleic Acids Res 46:D497–D502. https://doi.org/10.1093/nar/gkx1130.
20. Baek J, Lee J, Yoon K, Lee H. 2017. Identification of unannotated small genes in Salmonella. G3 (Bethesda) 7:983–989. https://doi.org/10.1534/g3.116.036939.
21. Samayoa J, Yildiz FH, Karplus K. 2011. Identification of prokaryotic small proteins using a comparative genomic approach. Bioinformatics 27:1765–1771. https://doi.org/10.1093/bioinformatics/btr275.
22. Ndah E, Jonckheere V, Giess A, Valen E, Menschaert G, Van Damme P. 2017. Reparation: ribosome profiling assisted (re-)annotation of bacterial genomes. Nucleic Acids Res 45:e168. https://doi.org/10.1093/nar/gkx758.
23. Hemm MR, Paul BJ, Schneider TD, Storz G, Rudd KE. 2008. Small membrane proteins found by comparative genomics and ribosome binding site models. Mol Microbiol 70:1487–1501. https://doi.org/10.1111/j.1365-2958.2008.06495.x.
24. Storz G, Wolf YI, Ramamurthi KS. 2014. Small proteins can no longer be ignored. Annu Rev Biochem 83:753–777. https://doi.org/10.1146/annurev-biochem-070611-102400.
25. Fozo EM, Hemm MR, Storz G. 2008. Small toxic proteins and the antisense RNAs that repress them. Microbiol Mol Biol Rev 72:579. https://doi.org/10.1128/MMBR.00025-08.
26. Hemm MR, Paul BJ, Miranda-Ríos J, Zhang A, Soltanzad N, Storz G. 2010. Small stress response proteins in Escherichia coli: proteins missed by classical proteomic studies. J Bacteriol 192:46–58. https://doi.org/10.1128/JB.00872-09.
27. Ring D, Wolman Y, Friedmann N, Miller SL. 1972. Prebiotic synthesis of hydrophobic and protein amino acids. Proc Natl Acad Sci U S A 69:765–768. https://doi.org/10.1073/pnas.69.3.765.
28. Dunker AK, Cortese MS, Romero P, Iakoucheva LM, Uversky VN. 2005. Flexible nets. The roles of intrinsic disorder in protein interaction networks. FEBS J 272:5129–5148. https://doi.org/10.1111/j.1742-4658.2005.04948.x.
29. Guzman LM, Belin D, Carson MJ, Beckwith J. 1995. Tight regulation, modulation, and high-level expression by vectors containing the arabinose PBAD promoter. J Bacteriol 177:4121–4130. https://doi.org/10.1128/jb.177.14.4121-4130.1995.
30. Kyte J, Doolittle RF. 1982. A simple method for displaying the hydropathic character of a protein. J Mol Biol 157:105–132. https://doi.org/10.1016/0022-2836(82)90515-0.
31. Ulmschneider MB, Sansom MS. 2001. Amino acid distributions in integral membrane protein structures. Biochim Biophys Acta 1512:1–14. https://doi.org/10.1016/S0005-2736(01)00299-1.
32. Lee SC, Choi YC, Yu MH. 1990. Effect of the N-terminal hydrophobic sequence of hepatitis B virus surface antigen on the folding and assembly of hybrid beta-galactosidase in Escherichia coli. Eur J Biochem 187:417–424. https://doi.org/10.1111/j.1432-1033.1990.tb15320.x.
33. Upadhyay AK, Murmu A, Singh A, Panda AK. 2012. Kinetics of inclusion body formation and its correlation with the characteristics of protein aggregates in Escherichia coli. PLoS One 7:e33951. https://doi.org/10.1371/journal.pone.0033951.

34. Damper PD, Epstein W. 1981. Role of the membrane potential in bacterial resistance to aminoglycoside antibiotics. Antimicrob Agents Chemother 20:803–808. https://doi.org/10.1128/AAC.20.6.803.

35. Taber HW, Mueller JP, Miller PF, Arrow AS. 1987. Bacterial uptake of aminoglycoside antibiotics. Microbiol Rev 51:439–457.

36. Wistrand-Yuen E, Knopp M, Hjort K, Koskiniemi S, Berg OG, Andersson DI. 2018. Evolution of high-level resistance during low-level antibiotic exposure. Nat Commun 9:1599. https://doi.org/10.1038/s41467-018-04059-1.

37. Wang H, Yin X, Wu Orr M, Dambach M, Curtis R, Storz G. 2017. Increasing intracellular magnesium levels with the 31-amino acid MgtS protein. Proc Natl Acad Sci U S A 114:5689–5694. https://doi.org/10.1073/pnas.1703415114.

38. Du D, Voss J, Wang Z, Chiu W, Luisi BF. 2015. The pseudo-atomic structure of an RND-type tripartite multidrug efflux pump. Biol Chem 396:1073–1082. https://doi.org/10.1515/hsz-2015-0118.

39. Hobbs EC, Yin X, Paul BJ, Astarita JL, Storz G. 2012. Conserved small protein associates with the multidrug efflux pump AcrB and differentially affects antibiotic resistance. Proc Natl Acad Sci U S A 109:16696–16701. https://doi.org/10.1073/pnas.1210093109.

40. Gasser B, Saloheimo M, Rinas U, Dragosits M, Rodríguez-Carmona E, Baumann K, Giuliani M, Parrilli E, Branduardi P, Lang C, Porro D, Ferrer P, Tutino ML, Mattanovich D, Villaverde A. 2008. Protein folding and conformational stress in microbial cells producing recombinant proteins: a host comparative overview. Microb Cell Fact 7:11. https://doi.org/10.1186/1475-2859-7-11.

41. Jurgen B, Lin HY, Riemschneider S, Scharf C, Neubauer P, Schmid R, Hecker M, Schweder T. 2000. Monitoring of genes that respond to overproduction of an insoluble recombinant protein in Escherichia coli glucose-limited fed-batch fermentations. Biotechnol Bioeng 70:217–224. https://doi.org/10.1002/1097-0290(20001020)70:2<217::AID-BIT11>3.0.CO;2-W.

42. Linkevicius M, Sandegren L, Andersson DI. 2013. Mechanisms and fitness costs of tigecycline resistance in Escherichia coli. J Antimicrob Chemother 68:2809–2819. https://doi.org/10.1093/jac/dkt263.

43. Miki T, Orii Y, Mukohata Y. 1987. A mechanism of respiratory control: studies with proteoliposomes containing cytochrome oxidase and bacteriorhodopsin. J Biochem 102:199–209. https://doi.org/10.1093/oxfordjournals.jbchem.a122033.

44. Unterreitmeier S, Fuchs A, Schäffler T, Heym RG, Frishman D, Langosch D. 2007. Phenylalanine promotes interaction of transmembrane domains via GxxxG motifs. J Mol Biol 374:705–718. https://doi.org/10.1016/j.jmb.2007.09.056.

45. Gurnev PA, Ortenberg R, Dorr T, Lewis K, Bezrukov SM. 2012. Persister-promoting bacterial toxin TisB produces anion-selective pores in planar lipid bilayers. FEBS Lett 586:2529–2534. https://doi.org/10.1016/j.febslet.2012.06.021.

46. Steinbrecher T, Prock S, Reichert J, Wadhwani P, Zimpfer B, Bürck J, Berditsch M, Elstner M, Ulrich AS. 2012. Peptide-lipid interactions of the stress-response peptide TisB that induces bacterial persistence. Biophys J 103:1460–1469. https://doi.org/10.1016/j.bpj.2012.07.060.

47. Proctor RA, von Eiff C, Kahl BC, Becker K, McNamara P, Herrmann M, Peters G. 2006. Small colony variants: a pathogenic form of bacteria that facilitates persistent and recurrent infections. Nat Rev Microbiol 4:295–305. https://doi.org/10.1038/nrmicro1384.

48. Musher DM, Baughn RE, Merrell GL. 1979. Selection of small-colony variants of Enterobacteriaceae by in vitro exposure to aminoglycosides: pathogenicity for experimental animals. J Infect Dis 140:209–214. https://doi.org/10.1093/infdis/140.2.209.

49. Pan X-S, Hamlyn PJ, Talens-Visconti R, Alovero FL, Manzo RH, Fisher LM. 2002. Small-colony mutants of Staphylococcus aureus allow selection of gyrase-mediated resistance to dual-target fluoroquinolones. Antimicrob Agents Chemother 46:2498–2506. https://doi.org/10.1128/AAC.46.8.2498-2506.2002.

50. Onyango LA, Dunstan RH, Gottfries J, von Eiff C, Roberts TK. 2012. Effect of low temperature on growth and ultra-structure of Staphylococcus spp. PLoS One 7:e29031. https://doi.org/10.1371/journal.pone.0029031.

51. Karatzas KA, Zervos A, Tassou CC, Mallidis CG, Humphrey TJ. 2007. Piezotolerant small-colony variants with increased thermotolerance, antibiotic susceptibility, and low invasiveness in a clonal Staphylococcus aureus population. Appl Environ Microbiol 73:1873–1881. https://doi.org/10.1128/AEM.01801-06.

52. Leimer N, Rachmühl C, Palheiros Marques M, Bahlmann AS, Furrer A,

53. von Eiff C, Friedrich AW, Becker K, Peters G. 2005. Comparative in vitro activity of ceftobiprole against staphylococci displaying normal and small-colony variant phenotypes. Antimicrob Agents Chemother 49:4372–4374. https://doi.org/10.1128/AAC.49.10.4372-4374.2005.

54. Andersson DI, Hughes D. 2010. Antibiotic resistance and its cost: is it possible to reverse resistance? Nat Rev Microbiol 8:260–271. https://doi.org/10.1038/nrmicro2319.

55. Maisnier-Patin S, Berg OG, Liljas L, Andersson DI. 2002. Compensatory adaptation to the deleterious effect of antibiotic resistance in Salmonella typhimurium. Mol Microbiol 46:355–366. https://doi.org/10.1046/j.1365-2958.2002.03173.x.

56. Lioy VS, Goussard S, Guerineau V, Yoon E-J, Courvalin P, Galimand M, Grillot-Courvalin C. 2014. Aminoglycoside resistance 16S rRNA methyltransferases block endogenous methylation, affect translation efficiency and fitness of the host. RNA 20:382–391. https://doi.org/10.1261/rna.042572.113.

57. Shcherbakov D, Akbergenov R, Matt T, Sander P, Andersson DI, Böttger EC. 2010. Directed mutagenesis of Mycobacterium smegmatis 16S rRNA to reconstruct the in vivo evolution of aminoglycoside resistance in Mycobacterium tuberculosis. Mol Microbiol 77:830–840. https://doi.org/10.1111/j.1365-2958.2010.07218.x.

58. Berghoff BA, Hoekzema M, Aulbach L, Wagner EG. 2017. Two regulatory RNA elements affect TisB-dependent depolarization and persister formation. Mol Microbiol 103:1020–1033. https://doi.org/10.1111/mmi.13607.

59. Dorr T, Vulic M, Lewis K. 2010. Ciprofloxacin causes persister formation by inducing the TisB toxin in Escherichia coli. PLoS Biol 8:e1000317. https://doi.org/10.1371/journal.pbio.1000317.

60. Levine MT, Jones CD, Kern AD, Lindfors HA, Begun DJ. 2006. Novel genes derived from noncoding DNA in Drosophila melanogaster are frequently X-linked and exhibit testis-biased expression. Proc Natl Acad Sci U S A 103:9935–9939. https://doi.org/10.1073/pnas.0509809103.

61. Cai J, Zhao R, Jiang H, Wang W. 2008. De novo origination of a new protein-coding gene in Saccharomyces cerevisiae. Genetics 179:487–496. https://doi.org/10.1534/genetics.107.084491.

62. Digianantonio KM, Hecht MH. 2016. A protein constructed de novo enables cell growth by altering gene regulation. Proc Natl Acad Sci U S A 113:2400–2405. https://doi.org/10.1073/pnas.1600566113.

63. Digianantonio KM, Korolev M, Hecht MH. 2017. A non-natural protein rescues cells deleted for a key enzyme in central metabolism. ACS Synth Biol 6:694–700. https://doi.org/10.1021/acssynbio.6b00336.

64. Donnelly AE, Murphy GS, Digianantonio KM, Hecht MH. 2018. A de novo enzyme catalyzes a life-sustaining reaction in Escherichia coli. Nat Chem Biol 14:253–255. https://doi.org/10.1038/nchembio.2550.

65. Keefe AD, Szostak JW. 2001. Functional proteins from a random-sequence library. Nature 410:715–718. https://doi.org/10.1038/35070613.

66. Davidson AR, Sauer RT. 1994. Folded proteins occur frequently in libraries of random amino acid sequences. Proc Natl Acad Sci U S A 91:2146–2150. https://doi.org/10.1073/pnas.91.6.2146.

67. Sievers F, Wilm A, Dineen D, Gibson TJ, Karplus K, Li W, Lopez R, McWilliam H, Remmert M, Söding J, Thompson JD, Higgins DG. 2011. Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. Mol Syst Biol 7:539. https://doi.org/10.1038/msb.2011.75.

68. Krogh A, Larsson B, von Heijne G, Sonnhammer EL. 2001. Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. J Mol Biol 305:567–580. https://doi.org/10.1006/jmbi.2000.4315.

69. Wilkins MR, Gasteiger E, Bairoch A, Sanchez JC, Williams KL, Appel RD, Hochstrasser DF. 1999. Protein identification and analysis tools in the ExPASy server. Methods Mol Biol 112:531–552.

70. Brandt U, Schagger H, von Jagow G. 1989. Purification of cytochrome-c oxidase retaining its pulsed form. Eur J Biochem 182:705–711. https://doi.org/10.1111/j.1432-1033.1989.tb14882.x.

71. Nilsson T, Lundin CR, Nordlund G, Ädelroth P, von Ballmoos C, Brzezinski P. 2016. Lipid-mediated protein-protein interactions modulate respiration-driven ATP synthesis. Sci Rep 6:24113. https://doi.org/10.1038/srep24113.

Eichenseher F, Seidl K, Matt U, Loessner MJ, Schuepbach RA, Zinkernagel AS. 2016. Nonstable Staphylococcus aureus small-colony variants are induced by low pH and sensitized to antimicrobial therapy by phagolysosomal alkalinization. J Infect Dis 213:305–313. https://doi.org/10.1093/infdis/jiv388.