The production of antibiotics by microbes in the environment and their use in medicine and agriculture select for existing and emerging resistance. To address this inevitability, prudent development of antibiotic drugs requires careful consideration of resistance evolution. Here, we identify the molecular basis for expanded substrate specificity in MphI, a macrolide kinase (Mph) that does not confer resistance to erythromycin, in contrast to other known Mphs. Using a combination of phylogenetics, drug-resistance phenotypes, and in vitro enzyme assays, we find that MphI and MphK phosphorylate erythromycin poorly resulting in an antibiotic-sensitive phenotype. Using likelihood reconstruction of ancestral sequences and site-saturation combinatorial mutagenesis, supported by Mph crystal structures, we determine that two non-obvious mutations in combination expand the substrate range. This approach should be applicable for studying the functional evolution of any antibiotic resistance enzyme and for evaluating the evolvability of resistance enzymes to new generations of antibiotic scaffolds.
Antibiotic resistance is difficult to overcome because its evolution is persistent and multifactorial. In addition to genetic mutations that confer drug insensitivity, mobile genetic elements disperse resistance genes in pathogens, adding to the complexity and unpredictability of the emergence of resistance in bacterial populations. Moreover, the prediction and detection of emerging resistance phenotypes resulting from mutations in resistance genes themselves remains a barrier to clinical surveillance and strategic antibiotic use, along with antibiotic drug development. A focus of the antibiotic pharmaceutical industry is on identifying new antibiotic scaffolds with decreased susceptibility to current and widespread resistance mechanisms. Another fruitful avenue in drug discovery is the development of successive generations of existing antibiotic scaffolds that bypass antibiotic-inactivating enzymes and other resistance mechanisms. For these strategies to be productive in the face of increasing rates of resistance, the functional and evolutionary landscape of resistance enzymes must be interrogated to inform new drug discovery efforts.

Macrolides such as azithromycin are among the most successful and highly prescribed antibiotics in the world. They are first-line treatments for community-acquired respiratory tract infections and gonorrhea, and increasingly used to treat infections caused by drug-resistant Enterobacteriaceae. Actinobacteria produce numerous variants of this antibiotic class but clinical implementation is mostly limited to erythromycin and its semi-synthetic derivatives, azithromycin and clarithromycin. The defining features of macrolide antibiotics include a 12-membered macro lactone ring and a dimethylamino sugar linked to the C5 position, which is essential for interacting with its target, the large subunit of the bacterial ribosome (Fig. 1). Some macrolides are decorated with additional sugars at C3 (azithromycin, erythromycin), 4'-OH of the dimethylamino sugar (carbomycin, josamycin), or at multiple positions on the macrolactone (megalomicin, spiramycin, tylosin). Telithromycin is a semi-synthetic derivative of erythromycin and a first-generation ketolide with a C11-C12 carbamate and an N-linked alkyl-aryl substituent, and a ketone (ketolide) at C3 in place of a cladinose. The inspiration for removing the cladinose in telithromycin development was natural ketolides that bypass the induction of macrolide resistance. Nevertheless, bacteria have evolved to overcome this modification. Macrolide resistance is abundant in pathogenic bacteria, and is most often the result of GTP-dependent macrolide kinases (Mph), ribosomal methyltransferases, or efflux pumps. Mph enzymes inactivate macrolides by phosphorylating the 2'-OH of the essential dimethylamino sugar, preventing it from binding the ribosome, and providing the chemical rationale for the resistance phenotype.

The Eukaryotic-like kinase superfamily is a structural family related to Eukaryotic protein kinases, and includes antibiotic kinases that modify macrolides (Mph), aminoglycosides (APH), and tuberactinomycins (APH). Mphs have undergone substantial functional radiation, leading to homologs with different substrate-specificities and regio-specificities. Four Mph homologs (MphA, MphB, MphC, MphE) are mobilized in human pathogens and confer resistance to a wide range of macrolide substrates. In contrast, Mphs from non-pathogenic bacteria are genetically and functionally diverse. We recently described an Mph (MphI) from the environmental bacterium Paenibacillus sp. LC231 that does not confer resistance to macrolides with a C3 cladinose. Paenibacillus sp. LC231 was isolated from Lichuguilla Cave where it was spatiotemporally separated from surface bacteria for over four million years. MphI shares high sequence identity (94%) to homologs found in related surface Paenibacillus sp., indicating the functional divergence of MphI is not recent. The Bacillus cereus group have two genetically and functionally distinct Mph enzymes; one that modifies a broad range of macrolides and another that cannot modify macrolides with 16-membered rings. The closest experimentally validated homolog to MphI is MphJ (51% identity), which confers resistance to erythromycin and tylosin, but not to spiramycin or josamycin. Mph enzymes are therefore adaptable resistance enzymes with complex evolutionary paths.

In this study, we show that narrow range Mphs are common in Bacillales, and their phosphorylation rates of cladinose containing macrolides is insufficient to confer resistance. We reconstruct the evolutionary path to Mph functional divergence using ancestral sequence reconstruction and identify four candidate residues responsible for substrate specificity. Applying phylogenetics, structural biology, protein engineering, and in vitro enzyme assays to the Mph family, we evolve increased substrate range in MphJ, which results from multiple non-obvious mutations that, in tandem, increase the catalytic rate towards cladinose containing macrolides. This work demonstrates that while the antibiotic kinases have evolved several distinct functions, there exists significant barriers to acquiring a very related, but non-intrinsic function.

Results

Multiple independent macrolide kinase mobilization events. Of the 13 Mph homologs that have been experimentally validated for macro lide kinase activity, 7 are captured by mobile genetic elements. This implies that Mphs are at a high risk for mobilization (Supplementary Fig. 1, Supplementary Table 1). These Mphs share low sequence identity and it is therefore likely that each homolog was mobilized independently; if they originated from a single mobilized Mph, then they should cluster together in phylogenetic analysis. However, only MphE, MphF, and MphG are phylogenetically clustered, while MphA, MphB, MphC, and MphN are isolated from other known Mphs (Fig. 2; Supplementary Fig. 2). These results show that Mph genes may

Fig. 1 Structure of macrolide antibiotics. The blue arrow indicates the site of phosphorylation by Mph enzymes. The C3 position is highlighted with a blue line.
have been captured by mobile genetic elements at least five independent times.

The taxonomic distribution of Mphs remains unclear despite several studies detailing their phenotypic, biochemical, and structural characterizations. We expanded the known genetic and taxonomic diversity of Mph enzymes by surveying bacterial strains with genomes deposited in the RefSeq database. Mphs were detected in 4732 genomes across 71 genera, and were predominantly found in Staphylococcus, Acinetobacter, Escherichia, Bacillus, and Klebsiella (Supplementary Fig. 3). Gram-negative pathogens, such as A. baumannii (38% of all A. baumannii strains), K. pneumoniae (38%), and E. coli (13%) are significantly enriched for Mphs (Fig. 3a). Ancestral A. baumannii, K. pneumoniae, and E. coli strains did not harbor an Mph homolog, and therefore it is very likely that these were acquired on mobile elements. Analysis of the taxonomic distribution of each Mph homolog reveals that MphA, MphB, and MphE are widespread in Gram-negative bacteria, and that the Mph homologs are limited to Bacillus and Paenibacillus and have not been mobilized (Fig. 3b; Supplementary Fig. 4). Our analysis demonstrates that Mphs are widely distributed across bacterial taxa, and the mobilized homologs are accumulating in Gram-negative pathogens.

**Substrate specificity correlates with genetic diversity.** MphI is the only known Mph that does not confer C3 cladinose macrolide resistance36. If MphI lost the ability to phosphorylate macrolides with a C3 cladinose, then we would expect it to share a common ancestor with Mphs that have wider substrate specificity, and therefore not positioned near the root of a phylogenetic reconstruction. The closest known homolog to MphI, MphJ (51% sequence identity), can phosphorylate C3 cladinose macrolides38 and is an outgroup to the MphI clade (Fig. 2; Supplementary Fig. 2). The most parsimonious explanation is that MphI lost the ability to phosphorylate C3 cladinose macrolides. Adjacent to MphI is another clade with uncharacterized enzymes from Bacillus that we reasoned may show MphI-like specificity if phylogeny predicts substrate specificity. We heterologously expressed one of the members of this clade in E. coli TOP10 (mphK/ycbJ, 54% identical to mphI) from Bacillus subtilis 168 and verified that it confers resistance to macrolide antibiotics (Supplementary Table 2). Moreover, like mphI, mphK does not confer resistance to C3 cladinose macrolides, indicating that Mph substrate specificity correlates with phylogeny.

To further investigate Mph substrate specificity, we performed in vitro enzyme assays with purified MphB, MphI, and MphK. Azithromycin and telithromycin are semisynthetic derivatives of erythromycin, with telithromycin lacking a C3 cladinose (Fig. 1). MphI and MphK both phosphorylate telithromycin but also phosphorylate azithromycin (Supplementary Fig. 5, Supplementary Table 2), suggesting that resistance phenotype may not correlate with biochemical analysis of drug modification. To confirm this result, we used tandem mass spectrometry to identify the site of phosphorylation on erythromycin, a macrolide
antibiotic with a similar structure to azithromycin and with well-established mass fragmentation patterns. Both MphI and MphK phosphorylate the 2′-OH of erythromycin (Supplementary Fig. 6). Using steady-state kinetics, we demonstrated that MphI and MphK can indeed use C3 cladinose macrolides as substrates, but not efficiently ($k_{cat}/K_m \approx 10^1$–$10^2$) (Supplementary Tables 4 and 5). The difference in catalytic constant ($k_{cat}/K_m$) values between C3 cladinose macrolides and descladinose macrolides is about $10^3$ and $10^1$–$10^3$ respectively for MphI and MphK. This poor biochemical specificity translates to the lack of resistance phenotype to these antibiotics when the genes are expressed in E. coli. To confirm whether this low activity against C3 cladinose macrolides impacts resistance phenotypes, we expressed mphI and mphK in an antibiotic hyper-susceptible strain of E. coli (E. coli BW25113 ΔbamBΔtolC). Both mph genes confer only a modest decrease in sensitivity to these antibiotics due to impaired efflux and outer membrane composition in this strain, verifying that the enzymatic activity observed in vitro has minimal impact on resistance phenotype (Supplementary Table 6).

MphB is another homolog also reported to have a narrow substrate range, modifying erythromycin but not azithromycin. In contrast to previous reports, we found that MphB confers resistance to all macrolides tested (Supplementary Tables 2 and 3), and inactivates both telithromycin and azithromycin.
Moreover, MphB modifies the desosamine 2'OH of erythromycin (Supplementary Fig. 6). As expected, the structures of MphB (from our study) and MphH showed that the enzymes adopt the bi-lobe kinase fold reminiscent of APH(2′′′) enzyme crystal structures, with an N-terminal domain dominated by β sheets and a C-terminal domain.

**Table 1.** Antibiotic susceptibility of *mphl* and two library mutants with expanded resistance to C3 cladinose macrolides expressed in *E. coli* BW25113 ΔbamBΔtolC

| Antibiotic       | Empty vector | mphl | azi_1.T | ery_aph |
|------------------|--------------|------|---------|---------|
| Erythromycin     | 0.5          | 1    | 16      | 16      |
| Clarithromycin   | 0.5          | 1    | 4       | 4       |
| Azithromycin     | 0.063        | 0.25 | 2       | 2–4     |
| Telithromycin    | 0.13         | >16  | >16     | >16     |
| Spiramycin       | 2            | 256  | 256     | 256     |
| Tylosin          | 2            | 1024 | 1024    | 1024    |
| Jasomycin        | 2–4          | >64  | >64     | >64     |
| Kanamycin        | 1            | 1    | 1       | 1       |

While preparing this manuscript, crystal structures of MphA and MphB were reported in complex with various macrolides, which we have included in our analysis of substrate specificity. We also attempted to crystallize Mphl and MphK as apoenzymes, with several GTP analogs, and with spiramycin, tylosin, and telithromycin, but we were unsuccessful.

As expected, the structures of MphB (from our study) and MphH showed that the enzymes adopt the bi-lobe kinase fold reminiscent of APH(2′′′) enzyme crystal structures, with an N-terminal domain dominated by β sheets and a C-terminal domain.
comprised mainly of α-helices, linked via a β-hairpin insertion (residues 100–110) (Fig. 4a; Supplementary Fig. 7a). The asymmetric unit of the MphH azithromycin crystal contains two copies of the complex ("complex A" and "complex B"), each showing subtle but important conformational differences (Fig. 4b). In both complexes, the azithromycin molecule is bound in a deep, electronegative cleft with one face formed by the end of C-terminal domain (i.e., helices α10 and α11), with the other face formed by the apposition of the β-hairpin, the highly conserved region of the C-terminal domain containing the catalytic residues, and the nucleotide-positioning loop (NPL; residues 29–33) of the N-terminal domain (Fig. 4a). However, it was evident that the "complex B" encloses the bound macrolide more intimately and more appropriately for catalysis (Fig. 4b) and this complex dominates our analysis.

We compared the overall structure of MphH with MphA (5IGI45) and MphB (from ref. 45, 5IGV) to identify possible conformational changes associated with macrolide binding. While the MphA, MphB, and MphH structures are very similar, we observed structural distinctions in all three enzymes that localized to the (a) NPL, (b) β-hairpin insert, and (c) multiple residues in the macrolide-binding cleft (Supplementary Fig. 7a). As we observed in MphH, the NPL, and β-hairpin regions of both MphA and MphB adopt different conformations with macrolide binding. Consistent with the flexibility of MphH in our crystal structure, these observations highlight that conformational changes in these key regions of Mph enzymes play a role in macrolide substrate binding.

We compared the positioning of bound macrolides among the Mph enzymes to gain better insight into recognition of the C3 cladinose ring. The position of azithromycin in the macrolide-binding cleft was compared with that of erythromycin and telithromycin (Fig. 4c–e). We observed that the binding orientations of azithromycin, erythromycin, and telithromycin are similar in MphH, MphA, and MphB, indicating that these macrolides are recognized in a similar manner by these enzymes. However, the clustering of residues around the C3 cladinose ring is different among these macrolides, with azithromycin having a more extended interaction with the enzyme compared to erythromycin and telithromycin (Fig. 4c–e).

Fig. 6 Kinetic characterization of MphI and mutants with improved activity towards C3 cladinose macrolides. Steady-state enzyme kinetics for MphI library mutants (a, b, c) and intermediate Ery_1 mutants (d, e, f). a K_m, b k_cat, and c k_cat/K_m values for MphI and library mutants are presented in clustered bar graph format. d K_m, e k_cat, and f k_cat/K_m values for intermediate Ery_1 mutants are also presented. Error bars represent the standard error of the mean (s.e. m.). The y-axis for c and f are presented in log10 format.
Substrate selectivity is an emergent property in Mphs differently upon binding to individual Mph enzymes. This analysis reveals that the C3 cladinose ring is positioned occupies deeper regions of the binding clefts of MphA or MphB. Consequently, the C3 cladinose ring of azithromycin is positioned closer to the nucleotide binding region as compared to its the azithromycin molecule bound to MphH is positioned ~3 Å binding pockets of MphA, MphB, and MphH are similar in that rings are oriented appropriately for catalysis, and the ‘‘hairpin’’ regions. However, the azithromycin molecule bound to MphH is positioned ~3Å closer to the nucleotide binding region as compared to its position in either the MphA or MphB active sites (Supplementary Fig. 7a). Consequently, the C3 cladinose ring of azithromycin occupies deeper regions of the binding clefts of MphA or MphB. This analysis reveals that the C3 cladinose ring is positioned differently upon binding to individual Mph enzymes.

**Substrate selectivity is an emergent property in Mphs.** Since structural analysis demonstrated that the C3 cladinose occupies slightly different positions in the Mpha, MphB, and MphH active sites, we speculated that MphI and MphK may have a large residue facing toward the C3 cladinose that would sterically hinder binding of these macrolides. We therefore compared the sequence composition of the six residues that approach the cladinose in the Mph structures (i.e., in the MphH azithromycin co-structure—Glu196, Phe229, Phe265, Ala268, Gly271, Tyr272) across a multiple sequence alignment of all known Mphs and their close homologs (>80% identity). We found no significant correlations between residue type and antibiotic substrate specificity (Fig. 4b; Supplementary Fig. 7b, Supplementary Fig. 8). Ultimately, this analysis of the Mph azithromycin molecular contacts shows that the physicochemical properties of residues in proximity with the C3 cladinose ring are not conserved even in wide substrate range Mphs. Therefore, C3 cladinose macrolide specificity does not appear to be directed by nearby residues.

Since Mph substrate specificity is a phylogenetic trait, we reasoned that reconstructing the evolutionary path of functional divergence could identify the molecular determinants influencing C3 cladinose phosphorylation. To this end, we reconstructed ancestral amino acid sequences31 using extant enzymes based on a multiple sequence alignment and phylogeny of related Mph homologs (>80% identity) (Fig. 5a). Both Mph phylogenies (Figs. 2 and 5a) independently indicate MphI and MphK lost activity towards C3 cladinose macrolides, increasing our confidence in the branching order of Mph sequences. Therefore,
the equivalent residues different between Ancestor 1 (MphI and MphK) and Ancestor 2 (broad substrate range) may be determinants of substrate specificity. We mapped these positions onto the MphH-azithromycin crystal structure and found four common sites near the macrolide-binding pocket, thus raising the possibility of their involvement in increased activity towards C3 cladinose macrolides (Fig. 5b–d). According to the MphH-azithromycin structure, one of these sites, MphH_L31, interacts with azithromycin; MphH_T160 faces toward the desosamine and C6–C7 of the macrolactone ring. MphHA200 (MphH_D201) is on the short helix α8 that approaches MphH_L190, which forms a key hydrophobic packing interaction with the macrolactone ring close to the methyl on C13 and the ethyl group on C14. MphH_A190 (MphH_T166) is in α6 that lines the macrolide-binding pocket and lies within 6 Å of the cladinose. The final site, MphH_Y101 (MphH_S111) is more peripheral to the active site and is in the β hairpin but faces away from the macrolide-binding pocket.

We independently mutated each of these positions in MphI to the corresponding residue in the ancestral sequence: none of these variants conferred resistance to C3 cladinose macrolides (Supplementary Table 9). Since engineering new substrate specificity (gain of function) into enzymes is often complex and mutations may require a specific constellation of multiple residues to result in a functional change, we created an MphI library with mutations at all four positions, in combination, using combinatorial site-saturation mutagenesis. This enabled effective sampling all 160,000 possible amino acid substitutions at each of the four positions identified in ancestral reconstruction, in combination (1.05 million codon combinations). Using the powerful positive selection for antibiotic resistance to screen this mutant library, we were able to identify gain of function azithromycin or erythromycin resistance. We identified three azithromycin-resistant variants (Azi_1.1, Azi_1.2, Azi_1.3) with identical amino acid sequences, but different codon combinations, and one erythromycin-resistant variant (Ery_1) (Fig. 5c, Table 1; Supplementary Table 10). All variants with an expanded substrate range were identical at three residue positions: Ala40, Glu166, and Leu211 (mutants of MphI Met40, Thr166, Pro211). This result was revealed a very limited number of routes to increase substrate range.

**MphI substrate selectivity is correlated with catalytic rate.** We purified Azi_1.1 and Ery_1, and verified azithromycin and telithromycin phosphorylation in an in vitro enzyme assay (Supplementary Fig. 5). Mutations in other antibiotic modifying enzymes are known to alter regiospecificity, and the cladinose hydroxyl of azithromycin is a possible phosphate acceptor. We unambiguously identified the desosamine 2'-OH as the phosphorylation site of Azi_1.1 and Ery_1 using tandem mass spectrometry (Supplementary Fig. 6). Using steady-state kinetics, we demonstrated that the k_{cat}/K_m of Azi_1.1 and Ery_1 for C3 cladinose macrolides increased 35–86 fold and 26–46 fold, respectively (Fig. 6a–c; Supplementary Tables 11 and 12). Rather than a significant decrease in K_m for C3 cladinose macrolides, the increase in activity towards these substrates was primarily driven by an increase in k_{cat}. The K_m decreased by about half, while the k_{cat} value increased 16× (erythromycin), 22–25× (azithromycin), and 48–51× (roxithromycin). Our results suggest that the four mutations in Azi_1.1 and Ery_1 increase the rate of product formation, and do not significantly increase catalytically productive C3 cladinose macrolide-binding affinity.

**Expanded substrate specificity requires binary mutations.** Ancestral sequence reconstruction successfully identified residues implicated in substrate specificity, but the molecular roles of these residues remains unclear. To further explore these, we probed each position's contribution by systematically reversing each Ery_1 (MphI_M40A/S111R/T166E/P211L) mutation to wild-type sequence. MphI_M40A/T166E/P211L and MphI_M40A/S111R/P211L conferred resistance to C3 cladinose macrolides and retained the same resistance phenotype as Ery_1 (Table 2). Similarly, MphI_S111R/T166E had the same resistance phenotype as MphI. These results demonstrate that positions 111 and 166 are not involved in expanding substrate specificity, consistent with their location more distal to the macrolide-binding site. In contrast, MphI_M111R/T166E/P211L and MphI_M40A/S111R/T166E lost the ability to confer C3 cladinose macrolide resistance, indicating that a binary combination of both mutations is required. MphI_M40A/P211L had the same resistance phenotype as Ery_1 (MphI_M40A/S111R/T166E/P211L), further supporting the role of both mutations. We confirmed these results using steady-state kinetics of MphI_M40A/S111R/T166E/P211L, MphI_M40A/S111R/T166E and MphI_S111R/T166E (Fig. 6d–f; Supplementary Tables 13–15). Each MphI variant had k_{cat}/K_m values similar to wild-type MphI. These results indicate that together residues Met40 and Pro211, play key roles in macrolide substrate specificity of Mphs. Met40 and Pro211 are located in the nucleotide-positioning loop and in the highly structurally conserved region near GDP-coordinating/catalytic residues, respectively.

**Discussion**

New generations of macrolide antibiotics are being developed to bypass resistance, which necessitates understanding the evolutionary landscape of Mphs. In this study, we find that two non-obvious mutations in tandem are required for expanding the resistance phenotype of MphI to include C3 cladinose macrolides by increasing the catalytic rate. These results are consistent with a plastic MphI active site that accommodates a wide selection of macrolides, but selectively phosphorylates macrolides without a cladinose. Narrow range Mphs such as MphI lost the ability to modify C3 cladinose macrolides, likely because of lack of selection by diverse macrolide structures in the local environment accompanied by neutral drift.

Amino acid residues positioned near, but not in, substrate-binding sites can impact function by modulating protein dynamics. For example, combinations of mutations located outside the β-lactam binding pocket of β-lactamases expand substrate range by increasing conformational flexibility. Introduction of each mutation individually does not expand the substrate range of β-lactamases. In our study, M40A and P211L are located outside the MphI macrolide-binding pocket and both mutations are required to confer resistance to C3 cladinose macrolides. Altogether, the binary M40A and P211L mutations may increase the conformation flexibility of the macrolide-binding pocket, increasing the phosphorylation rate of C3 cladinose macrolides.

Understanding the evolution of resistance enzymes is of considerable interest because of their wide functional and genetic diversity, and their biological importance in medicine and microbial ecology. Combining random mutagenesis strategies with antibiotic selection is powerful for interrogating the genetic-phenotypic continuum of resistance enzymes. Indeed, many of the strategies for protein engineering and synthetic biology were first developed using resistance enzymes, and provide a strong foundation for biochemical investigations of resistance. Our approach builds on previous strategies by identifying target positions using phylogenetic and evolutionary principals, and sampling random genetic diversity at each of the four positions simultaneously. The
strategy we report here offers a single step for efficient site-saturation combinatorial mutagenesis of multiple positions using Gibson Assembly and productive sampling of the mutant library (>1 million variants) using positive antibiotic selection. Our approach successfully identified candidate positions involved in Mph substrate selectivity that were unidentifiable from Mph–macrolide co-structures reported here or by Feng et al.19. Despite the macroalctone ring and dimethylamino sugar shared by all macrolides, the M40A and P211L mutations in Mph specifically increase phosphorylation rate of those with a C3 cladinose without affecting catalysis of descladinose macrolides. In the MphH-azithromycin co-structure, the equivalent MphH positions Leu31 and Ala200 do not interact with the C3 cladinose. Our study illustrates the utility of evolutionary guided mutagenesis for studying non-obvious molecular determinants of substrate selectivity.

Ancestral sequence reconstruction has proven to be a valuable tool in drug discovery19, protein engineering46,48, and molecular evolution studies47,50,66. We believe our approach will be widely applicable for uncovering the molecular basis of functional divergence and antibiotic resistance enzymes. New generations of antibiotic scaffolds are being designed to bypass resistance by evolving the potential for evolving resistance is high. Use of ancestral antibiotics17,55. Our study suggests that macrolide binding may be a precursor to phenotypic resistance in Mphs. Therefore, next generation macrolide antibiotics that bypass Mph-mediated resistance should also be investigated for active site binding, and whether homologs from environmental bacteria confer resistance to the new antibiotics, the potential for evolving resistance is high. Use of ancestral sequence reconstruction combined with structural and biochemical studies, as performed here, will be useful for predicting the possibility and feasibility of expanded resistance phenotypes to semi-synthetic antibiotics (Fig. 7).

Recent platforms that synthesize diverse macrolides will enable the discovery of newer generations of resistance-proof antibiotics17,55. Our study suggest that macrolide binding may be a precursor to phenotypic resistance in Mphs. Therefore, next generation macrolide antibiotics that bypass Mph-mediated resistance should also be investigated for active site binding, and whether homologs from environmental bacteria confer resistance to Mphs.

Methods

Antibiotics and reagents. All buffers and salts were purchased from BioShop (Burlington, ON, Canada) unless otherwise specified. Macrolide antibiotics, GTP, pyruvate kinase: lactate dehydrogenase (PK4LDH), phosphopolypruvate, and NADH were purchased from Sigma-Aldrich (Oakville, ON, Canada). Molecular biology kits and organic solvents were purchased from Fisher Scientific (Ottawa, ON, Canada). Telithromycin was purified from the drug formulation Ketek (400 mg, Sanofi-Aventis US)36.

Phylogenetic reconstruction of macrolide kinases. Diverse Mph sequences were identified by BLASTp56 by querying the Mph sequences in Supplementary Table 1 with an e-value cutoff of 1e–30, and the following flags: -max_target_seqs 500, and -max指望 ‘QUERY NOT partial’. Note that Mph enzymes were renamed for consistency with established nomenclature56 and to avoid multiple distinct names (<80% amino acid identity) with the same name. A hit was defined as having at least 50% sequence identity and an alignment length >80% of the query length. Sequences were clustered at 80% identity using the cluster smallmem algorithm of uclust56, and aligned with MAFFT (L-INS-i method)58. APH(2’)- enzymes are homologs of MphB17, and therefore APH(2’)-Ila, APH(2’)-Ile, and APH(2’)-I (Ile) were used as outgroups for rooting the tree. The alignment was then weighted by alignment confidence using the transitive consistency score (TCS) function of T-coffee72. RAxML was used to generate a maximum likelihood phylogeny using the transitive consistency score (TCS) function of T-coffee72. RAxML was used to generate a maximum likelihood phylogeny using the rapid bootstrap72. RAxML was used to generate a maximum likelihood phylogeny using the rapid bootstrap72. RAxML was used to generate a maximum likelihood phylogeny using the rapid bootstrap72. RAxML was used to generate a maximum likelihood phylogeny using the rapid bootstrap72.
Ancestral reconstruction of ancestral Mph sequences. Close Mph homologs were identified with ncbi-blast-2.23.1 + BLASTP68 by querying the Mph sequences in Supplementary Table 1, e-value cutoff of 1e-50, -max_target_seq 500, and –entrez_query ‘NOT partial’. A close homolog was defined as having >80% identity and an alignment length of ≥85% of the query length. Sequences were aligned using MAFFT (L-INS-i) and the BLOSUM45 substitution matrix. The alignment was weighted using the t-coffee TCS function. A maximum likelihood tree was generated with RAxML with the JTT model (PROTGAMAAAUTO) and rapid bootstrap analysis of 100 replicates. Reconstructed ancestral sequences were inferred using an Empirical Bayes approach (PAML 4.9a) with the Jones amino acid replacement matrix. For ancestral reconstruction, the RAxML generated tree was rooted at the point of functional divergence and used the unweighted multiple sequence alignment. The ancestral sequences of each Mph subtype (i.e., those that phosphorylate C3 cladocine macrolide efficiently and those that do it poorly) were contrasted, and these positions were mapped onto the MphH crystal structure to identify residues near the macrolide-binding pocket. The alignment for ancestral reconstruction was also used to identify amino acid conservation for sequence logos.

Construction of Mph mutant library. We constructed the Mph mutant library by simultaneously introducing ‘NNK’ codons at amino acid positions 40, 111, 166, and 211 using a hierarchical approach based on the Gibson Assembly method66. This method assembled 4 PCR fragments; 3 sections of mphi and another containing part of the gene and the vector. Oligonucleotides were designed to amplify 93–358 (f1) and 471–633 (f2), and incorporate ‘NNK’ codons at the target positions with an additional 25 bp on each end for Gibson Assembly. Additionally, 334–495 (f3) was used as the vector. In total, 634–827 (f4) (Supplementary Fig. 9). PCR fragments were amplified with Phusion Polymerase (Thermo Scientific) and gel purified with 2% low melting point agarose (Thermo Scientific). Fragments f1 (266 bp), f2 (162 bp), f3 (188 bp), and f4 backbone (3636 bp) were assembled in a 20 μL Gibson Assembly reaction (performed in triplicate) and incubated at 50 °C for 1 h. The Gibson reaction contained: 100 ng f1 backbone; 20 ng f2, 12 ng f3, 14 ng f4, 0.004 U T5 exonuclease (New England Biolabs), 0.025 U Phusion polymerase, 1 U Taq ligase (New England Biolabs), 0.2 mM each dNTP (FroggBio), 1 mM NAD (Sigma), 100 mM Tris- HCl pH 7.5 (Bioshop), 10 mM MgCl2 (Fisher), 10 mM DTT (Bioshop), and 125 mg mL−1 PEG-8000 (Sigma). Reactions were purified using PCR purification kit (Thermo Scientific) and eluted in 20 μL water. A volume of 1 μL was used to transform three individual E. coli Electromax electrocompetent cell (Invitrogen) reactions. Cells were recovered in a total of 3 mL Super Optimal broth with Catabolite repression (SOC) (Invitrogen) for 1 h at 37 °C with shaking. Library titer was estimated by plating 103 and 104 dilutions and determining the CFU mL−1 of transformation. Plasmids from 10 random colonies were isolated, sent for Sanger sequencing (Mobix, McMaster University), and verified to contain a unique missense mutation at each of the four target positions. In total, the library contained 3.9 million variants. The rest of the transformations were used to inoculate 100 mL LB-Lennox with 100 μg mL−1 carbenicillin. The reaction was allowed to incubate with shaking for 16 h at 37 °C. 50 mL was centrifuged at 4000 × g for 1 h at 4 °C, resuspended in 40 mL of fresh LB-Lennox and 250 μL of transformation. Plasmids from 10 random colonies were isolated, sent for Sanger sequencing (Mobix, McMaster University). For protein purification experiments Mph mutants were subcloned into pET22b.

Site-directed mutagenesis. Oligonucleotides in Supplementary Table 16 were used for site-directed mutagenesis in the following 50 μL reaction; 10 ng plasmid, 0.5 μM primers, 200 μM dNTPs, 1× HF buffer, 1 U Phusion polymerase. The following thermocycler conditions were used; initial denaturation 98 °C 2 min, 98 ° C 10 s, 55 °C 20 s, 72 °C 3.75 min for 16 cycles, and 7 min final extension at 72 °C. An aliquot of 2 μL of FastDigest DpnI (Thermo Scientific) was added and incubated at 37 °C for 1 h to remove template DNA. Reactions were cleaned up with PCR purification, and 5 μL was used to transform chemically competent E. coli TOP10. Plasmids were isolated and sent for Sanger sequencing to confirm the desired mutations. Wild-type Mph was used for initial protein preparations with which to reconstruct the E. coli MphI 1 to 1 by matching the unique 4 letter ecf tag and a TEV protease cleavage site. Se-Met-substituted versions of MphB and MphH were expressed using the standard M9 high yield growth procedure according to the manufacturer’s instructions (Shanghai Medicolon). Crystalization was performed at room temperature using the sitting drop method and 2 μL protein or protein:ligand mixture plus 2 μL reservoir solution. The MphH concentration was 10 mg mL−1 and the MphB concentration was 90 mg mL−1. Se-Met MphB (apoenzyme) was crystallized with reservoir solution 0.2 M calcium acetate, 0.1 M sodium cacodylate pH 6.5, 18% (w/v) PEG 8 K. Native MphH (apoenzyme) was crystallized with reservoir solution 2 M ammonium sulfate, 0.1 M Hepes pH 7.5, 2% (w/v) PEG 400 and 1% trehalose. Se-Met MphH-GDP complex was crystallized with reservoir solution 2 M ammonium sulfate, 0.1 M Hepes pH 7.5, 2% (w/v) PEG400, 1% trehalose, 5 mM azithromycin and 5 mM GMPPCP. Native MphHazithromycin complex was crystallized with reservoir solution 2 M ammonium sulfate, 5 mM azithromycin and 5% DMSO. All crystals were cryoprotected in paratone oil prior to data collection. Diffraction data for MphH (apoenzyme) was collected at 100 K using a Rigaku Micromax 007-HF rotating anode and a Rigaku R-AXIS IV++ detector. Diffraction data for MphB (Se-Met, apoenzyme), MphHGDP (Se-Met) and MphHazithromycin (native) were collected at the Advanced Photon Source, Argonne National Laboratory, Life Sciences Collaborative Access Team beamlines 21-ID-F or 21-ID-G at the selenium absorption edge. Data was processed by HKL-300061 or XDS62 and Aimless63. SAD structure determination was completed using Phenix.autosolv64 and Molecular replacement for native structures was compiled using Phenix Phaser. Refinement was completed with Phenix.refine and Coot65. All B-factors were refined as isotropic with TLS parameterization. All geometry was verified using Phenix validation tools and the wwPDB server. Electron density maps are in Supplementary Figure 16.

Code availability. The custom python code used to survey taxonomic distribution of Mphs is available from the corresponding author upon request.

Data availability. The crystal structures of Mph enzymes were deposited into the PDB; MphB Se-Met apoenzyme (5UXA), MphH native apoenzyme (5UXB), MphH (native) GDP + Pi complex (5UXC), and MphH (Se-Met) azithromycin complex (5UXD). Protein accession numbers for all Mph enzymes included in this study are listed in Supplementary Table 1. Other data are available from the corresponding author upon reasonable request.

Received: 7 September 2017 Accepted: 19 December 2017 Published online: 09 January 2018

References
1. Perry, J. A. & Wright, G. D. Forces shaping the antibiotic resistance. BioEssays 36, 1179–1184 (2014).

2. Davies, J. & Davies, D. Origins and evolution of antibiotic resistance. Microbiol. Mol. Biol. Rev. 74, 417–433 (2010).
3. Jia, B. et al. CARD 2017: expansion and model-centric curation of the comprehensive antibiotic resistance database. Nucleic Acids Res. 45, D566–D573 (2017).

4. Gibson, M. K., Forsberg, K. J. & Dantas, G. Improved annotation of antibiotic resistance determinants reveals microbial resistomes cluster by ecology. ISME J. 9, 207–216 (2015).

5. Cao, M. S., Gasparini, A. J. & Dantas, G. Next-generation approaches to understand and combat the antibiotic resistance. Nat. Rev. Microbiol. 15, 422–434 (2017).

6. Bobicsek, A. et al. Fluoroquinolone-modifying enzyme: a new adaptation of a chloramphenicol acetyltransferase. Nat. Med. 12, 83–88 (2006).

7. Chanal, C. et al. Nucleotide sequences of CAZ-2, CAZ-6, and CAZ-7 beta lactamase genes. Antimicrob. Agents Chemother. 36, 1817–1820 (1992).

8. Henguell, C., Chanal, C., Siriot, D., Labia, R. & Siriot, J. Molecular characterization of nine different types of mutants among 107 inhibitor-
resistant TEM beta-lactamases from clinical isolates of Escherichia coli. Antimicrob. Agents Chemother. 39, 427–430 (1995).

9. Casin, I. et al. Antibiotic resistance in clinical isolates of Enterobacter cloaceae and Citrobacter freundii. Antimicrob. Agents Chemother. 47, 1747–1751 (2003).

10. Silver, L. L. Challenges of antibacterial discovery. Clin. Microbiol. Rev. 24, 71–109 (2011).

11. Sonnenburg, D. A., Munck, C., Toft-Kehler, R. V. & Andersson, D. I. Prediction of antibiotic resistance: time for a new preclinical paradigm? Nat. Rev. Microbiol. 15, 689–696 (2017).

12. Bonnefoy, A., Girard, A. M., Agouridas, C. & Chantot, J. F. Ketolides lack inducibility properties of MLS(B) resistance phenotype. J. Antimicrob. Chemother. 40, 85–90 (1997).

13. Armstrong, E. S. & Miller, G. H. Combating evolution with intelligent design: the neoglycoside ACHN-490. Curr. Opin. Microbiol. 13, 565–573 (2010).

14. Papp-Wallace, K. M., Endimiani, A., Taracila, M. A. & Bonomo, R. A. Synthesis and structure-activity relationship of novel macrolide antibiotics. Nat. Commun. 9, 1462 (2018).

15. Sum, P. E. & Petersen, P. Forty years of beta-lactam research. FEMS Microbiol. Rev. 41, 533–556 (2017).

16. Rolinson, G. N. Macrolide resistance mechanisms in Streptococcus pneumoniae and Staphylococcus aureus. J. Antimicrob. Chemother. 55, 4934–4960 (2010).

17. Sum, P. E. & Petersen, P. Synthesis and structure-activity relationship of novel glycyclines: a new class of antibiotics leading to the discovery of GAR-936. Bioorg. Med. Chem. Lett. 9, 1459–1462 (1999).

18. Rolinson, G. N. Forty years of beta-lactam research. J. Antimicrob. Chemother. 41, 589–603 (1998).

19. Fernandes, P., Martens, E. & Pereira, D. Nature nurtures the design of new semisynthetic macrolides. J. Antibiot. 52, 523–533 (1999).

20. Hicks, L. A. et al. US outpatient antibiotic prescribing variation according to geography, patient population, and provider specialty in 2011. Clin. Infect. Dis. 60, 1308–1316 (2015).

21. Gomes, C. et al. Macrolide resistance mechanisms in Enterobacteriaceae: focus on azithromycin. Crit. Rev. Microbiol. 43, 1–30 (2017).

22. Kanez, K. et al. The general mode of translation inhibition by macrolide antibiotics. Proc. Natl Acad. Sci. USA 111, 15958–15963 (2014).

23. Dunkle, J. A., Xiong, L., Mankin, A. S. & Cate, J. H. D. Structures of the Escherichia coli ribosome with antibiotics bound near the peptidyl transferase center explain spectra of drug action. Proc. Natl Acad. Sci. USA 107, 17152–17157 (2010).

24. U, W., Blaha, G., Moore, P. B. & Steitz, T. A. Structures of MLSB resistance bound to mutated large ribosomal subunits provide a structural explanation for resistance. Cell 121, 257–270 (2005).

25. Bailey, M., Chettiath, T. & Mankin, A. S. Induction of erm(C) expression by noninducing antibiotics. Antimicrob. Agents Chemother. 52, 866–874 (2008).

26. Wolter, N. et al. Telithromycin resistance in Streptococcus pneumoniae is conferred by a deletion in the leader sequence of the ketolide resistance gene. Antimicrob. Agents Chemother. 48, 270–276 (2004).

27. Bhuller, K. et al. Antibiotic resistance is prevalent in an isolated cave microbiome. PLoS ONE 7, e34953 (2012).

28. Ohara, K., Kanda, T. & Kono, M. Structure of a phosphorylated derivative of oleandomycin, obtained by reaction of oleandomycin with an extract of an aerobic bacterium. FEMS Microbiol. Lett. 293, 147–159 (2008).

29. Bhatt, K. et al. Antibiotic resistance is prevalent in an isolated cave microbiome. PLoS ONE 7, e34953 (2012).

30. Park, B. & Min, Y. Inducible expression of erm(B) by the ketolides telithromycin and cethromycin. Int. J. Antimicrob. Agents 46, 226–227 (2015).

31. Roberts, M. C. Update on macrolide–lincosamide–streptranslamin, ketolide, and oxazolidinone resistance genes. FEMS Microbiol. Lett. 1573 (2015).

32. Matsuoka, M., Inoue, M., Endo, Y. & Nakajima, Y. Characteristic expression of three genes, msr(A), mph(C) and erm(Y), that confer resistance to macrolide antibiotics on Staphylococcus aureus. FEMS Microbiol. Lett. 144, 197–202 (1996).

33. Matsuo, M., Inoue, M., Endo, Y. & Nakajima, Y. Characteristic expression of three genes, msr(A), mph(C) and erm(Y), that confer resistance to macrolide antibiotics on Staphylococcus aureus. FEMS Microbiol. Lett. 144, 197–202 (1996).

34. Matsuo, M., Inoue, M., Endo, Y. & Nakajima, Y. Characteristic expression of three genes, msr(A), mph(C) and erm(Y), that confer resistance to macrolide antibiotics on Staphylococcus aureus. FEMS Microbiol. Lett. 144, 197–202 (1996).

35. Matsuo, M., Inoue, M., Endo, Y. & Nakajima, Y. Characteristic expression of three genes, msr(A), mph(C) and erm(Y), that confer resistance to macrolide antibiotics on Staphylococcus aureus. FEMS Microbiol. Lett. 144, 197–202 (1996).

36. Matsuo, M., Inoue, M., Endo, Y. & Nakajima, Y. Characteristic expression of three genes, msr(A), mph(C) and erm(Y), that confer resistance to macrolide antibiotics on Staphylococcus aureus. FEMS Microbiol. Lett. 144, 197–202 (1996).

37. Wang, C. et al. Functional characterization and phylogenetic analysis of acquired and intrinsic macrolide phosphotransferases in the Bacillus cereus genome. Environ. Microbiol. 17, 1573–1582 (2015).

38. Pavlovski, A. C. et al. A diverse intrinsic antibiotic resistance from a cave bacterium. Nat. Commun. 7, 13803 (2016).
65. Bratulic, S., Gerber, F. & Wagner, A. Mistranslation drives the evolution of robustness in TEM-1 beta-lactamase. Proc. Natl Acad. Sci. USA 112, 12758–12763 (2015).

66. Gaucher, E. A., Thomson, J. M., Burgan, M. F. & Benner, S. A. Inferring the palaeoenvironment of ancient bacteria on the basis of resurrected proteins. Nature 425, 285–288 (2003).

67. Pawlowski, A. C., Johnson, J. W. & Wright, G. D. Evolving medicinal chemistry strategies in antibiotic discovery. Curr. Opin. Biotechnol. 42, 108–117 (2016).

68. Camacho, C. et al. BLAST+: architecture and applications. BMC Bioinformatics 10, 421 (2009).

69. Edgar, R. C. Search and clustering orders of magnitude faster than BLAST. Bioinformatics 26, 2460–2461 (2010).

70. Katoh, K. & Standley, D. M. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. Mol. Biol. Evol. 30, 772–780 (2013).

71. Shakya, T. et al. A small molecule discrimination map of the antibiotic resistance kinome. Chem. Biol. 18, 1591–4601 (2011).

72. Chang, J. M., Di Tommaso, P. & Notredame, C. TCS: a new multiple sequence alignment reliability measure to estimate alignment accuracy and improve phylogenetic tree reconstruction. Mol. Biol. Evol. 31, 1625–1637 (2014).

73. Stamatakis, A. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. Bioinformatics 30, 1312–1313 (2014).

74. Clinical and Laboratory Standards Institute. Methods for dilution antimicrobial susceptibility testing of bacteria that grow aerobically. (CLSI, Wayne, PA, USA, 2012).

75. Gibson, D. G. et al. Enzymatic assembly of DNA molecules up to several hundred kilobases. Nat. Methods 6, 343–345 (2009).

76. Minor, W., Cymborowski, M., Otwinowski, Z. & Chruszcz, M. HKL-3000: the HKL-3000. Acta Crystallogr. D Biol. Crystallogr. 62, 859–866 (2006).

77. Kabasch, W. XDS. Acta Crystallogr. D Biol. Crystallogr. 66, 125–132 (2010).

78. Winn, M. D. et al. Overview of the CCP4 suite and current developments. Acta Crystallogr. D Biol. Crystallogr. 67, 235–242 (2011).

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

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

Acknowledgments

We thank Nicholas Waglechner (McMaster) for useful discussions on phylogenetics, Mph surveillance, and ancestral sequence reconstruction. We also thank Murray Junop (Western University) for helpful discussion on protein engineering experiments. This research was funded by a Canadian Institutes of Health Research grant (FRN-148463) and by a Canada Research Chair in Antibiotic Biochemistry (to G.D.W.) and through funds from the National Institute of Allergy and Infectious Diseases, National Institutes of Health, Department of Health and Human Services, under Contract No. HHSN27201200026C.

Author contributions

A.C.P. and G.D.W. designed research. A.C.P., P.J.S., and K.K. performed research and analyzed data. A.C.P. performed the following experiments; antibiotic susceptibility testing, phylogenetics, enzyme kinetics, ancestral sequence reconstruction, mass spectrometry, site-saturation combinatorial mutagenesis, and in vitro enzyme assays. P.G.S. performed crystal structure determination of MphH. K.K. performed tandem mass spectrometry experiments. T.S. and E.E. performed protein crystallization. A.S. supervised protein crystallization and contributed to interpreting mechanism of expanded substrate range. A.C.P., P.J.S., and G.D.W. wrote the manuscript with input from all authors.

Additional information

Supplementary Information accompanies this paper at https://doi.org/10.1038/s41467-017-02680-0.

Competing interests: The authors declare no competing financial interests.

Reprints and permission information is available online at http://npg.nature.com/reprintsandpermissions/

Publisher’s note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.