

Structural and molecular rationale for the diversification of resistance mediated by the Antibiotic_NAT family

Peter J. Stogios^{1#}, Emily Bordeleau^{2#}, Zhiyu Xu¹, Tatiana Skarina¹, Elena Evdokimova¹, Sommer Chou², Luke Diorio-Toth³, Alaric W. D'Souza³, Sanket Patel³, Gautam Dantas^{3,4,5,6}, Gerard D. Wright², and Alexei Savchenko^{1,7,8,*}

¹Department of Chemical Engineering and Applied Chemistry, University of Toronto, Toronto, Canada, M5S 3E5.

²David Braley Centre for Antibiotics Discovery, M.G. DeGroot Institute for Infectious Disease Research, Department of Biochemistry and Biomedical Sciences, McMaster University, Hamilton, ON, Canada, L8S 4L8.

³The Edison Family Center for Genome Sciences & Systems Biology, Washington University School of Medicine, St. Louis, MO, USA, 63110.

⁴Department of Pathology and Immunology, Washington University School of Medicine, St. Louis, MO, USA, 63110.

⁵Department of Biomedical Engineering, Washington University in St. Louis, St. Louis, MO, USA, 63130.

⁶Department of Molecular Microbiology, Washington University School of Medicine, St. Louis, MO, USA, 63110.

⁷Department of Microbiology, Immunology and Infectious Diseases, University of Calgary, Calgary, Canada, T2N 4N1.

⁸Center for Structural Genomics of Infectious Diseases (CSGID)

#equal contribution for first authorship

* to whom correspondence should be addressed: Alexei Savchenko, Department of Immunology and Infectious Diseases, Health Research Innovation Centre, 3330 Hospital Drive NW, University of Calgary, Calgary, AB, T2N 4N1, Canada. Tel: +1-403-210-7980; Email: alexei.savchenko@ucalgary.ca.

ABSTRACT

The environmental microbiome harbors a vast repertoire of antibiotic resistance genes (ARGs) which can serve as evolutionary predecessors for ARGs found in pathogenic bacteria, or can be directly mobilized to pathogens in the presence of selection pressures. Thus, ARGs from benign environmental bacteria are an important resource for understanding clinically relevant resistance. Here, we conduct a comprehensive functional analysis of the Antibiotic_NAT family of aminoglycoside acetyltransferases. We determined a pan-family antibiogram of 21 Antibiotic_NAT enzymes, including 8 derived from clinical isolates and 13 from environmental metagenomic samples. We find that environment-derived representatives confer high-level, broad-spectrum resistance, including against the atypical aminoglycoside apramycin, and that a metagenome-derived gene likely is ancestral to an AAC(3) gene found in clinical isolates. Through crystallographic analysis, we rationalize the molecular basis for diversification of substrate specificity across the family. This work provides critical data on the molecular mechanism underpinning resistance to established and emergent aminoglycoside antibiotics and broadens our understanding of ARGs in the environment.

INTRODUCTION

Antibiotic resistance is a global crisis that threatens every class of clinically deployed antibiotic¹. Antibiotic resistance genes (ARGs) isolated from bacteria that cause life-threatening disease can often be traced to environmental microbial communities (reviewed in ²⁻⁴). To understand the sources of antibiotic resistance, the identification of links connecting ARGs in the clinic with those in the environment, characterization of their horizontal transfer, evolution, and biochemical/molecular properties are focuses of continuing research (reviewed in ⁴ and ⁵). For example, the *mcr* family of plasmid-borne colistin resistance genes is thought to originate from chromosomal genes found in various *Moraxella* and *Aeromonas* species ⁶⁻⁸. The family of extended-spectrum β -lactamase *bla*_{CTX-M} genes found on plasmids of Gram-negative pathogens has been traced to the chromosomal genes of various *Kluyvera* species that are only rarely pathogenic ⁹. Given the regular exchange of genetic material harboring ARGs between microbial species, more research is required to understand the breadth and depth of the global resistome, including such aspects as the scope of resistance mechanisms, the specificity and efficiency of ARG products in conferring resistance, and their potential to be mobilized and transferred to pathogens. This comprehensive data is critical for tackling the antibiotic resistance crisis ^{5,10}.

Aminoglycosides (AGs) (**Fig. 1**) are widely used to treat infections caused both by Gram-positive and Gram-negative bacteria due to their broad-spectrum activity ¹¹. Toxicity and resistance are significant problems complicating the use of this class of drugs; nonetheless, they retain value for treating multi-drug and extensively-drug resistant Gram-negative pathogens causing serious infections¹². Canonical AGs are characterized by a core 2-deoxystreptamine ring with substitutions at the 4- and 6- or 4- and 5- positions. Non-canonical AGs possess variations on the 2-deoxystreptamine core such as streptomycin, or apramycin which contains a fourth ring

structure fused to 2-deoxystreptamine. Apramycin is currently used in veterinary medicine^{13,14}, and with the notable exceptions of *aac3-IV* and the emerging resistance determinant *apmA*^{15,16}, few ARGs confer resistance to apramycin, prompting excitement for broader deployment in medicine¹⁷⁻²¹.

AG resistance is primarily conferred by three classes of aminoglycoside-modifying enzymes (AMEs): phosphotransferases (APHs), nucleotidyltransferases (ANTs), and acetyltransferases (AACs)²². AMEs permanently alter the AG substrate, preventing them from binding to their target, the A-site of the 16S rRNA in the bacterial ribosome. AMEs are widely disseminated in pathogens. Current research focuses on their specificity, mechanisms, and inhibition by small molecules to fortify the design of next-generation AG against resistance, as exemplified by the development of plazomicin and apramycin analogs (apralogs,²³⁻²⁵).

Previously, we identified 27 AACs in grassland soils microbial communities using a functional metagenomics (FMG) approach^{26,27}. These AACs belonged to two sequence and structurally distinct acetyltransferase families - GNAT (GCN5-related *N*-acetyltransferase, PFAM clan CL0257) and Antibiotic_NAT (PFAM family PF02522). For GNAT AACs, we showed that many environment derived ARGs, which we called meta-AACs for metagenomic AACs, possess resistance activity, acetylation efficiency, and structural properties comparable to AMEs derived from drug resistant clinical species^{24,27}. Our research established that GNAT meta-AACs include all the qualities necessary to cause high-level resistance if mobilized and transferred to human pathogens.

In contrast to the GNAT family, much less is known about the biochemical, structural, and molecular features of the Antibiotic_NAT family. There are approximately 50 of this family members identified²⁷ and many are highly disseminated in Gram-negative pathogens²⁸,

including AAC(3)-II, AAC(3)-III, and AAC(3)-IVa. The AAC(3)-IIa enzyme possesses narrow AG specificity as it is active only against 4,6-disubstituted compounds, while AAC(3)-IIIa is strongly promiscuous due to its activity against a broad range of 4,5- and 4,6-disubstituted AGs^{29,30}. The AAC(3)-IVa enzyme was also shown to be promiscuous against a broad range of 4,5- and 4,6-disubstituted AGs as well as against apramycin³¹. There have been no studies describing the enzymatic characteristics of environment-derived members of this family and no comprehensive family-wide analysis to understand their diversification of structure and function.

Several members of the Antibiotic_NAT family have been structurally characterized, including AAC(3)-IIIb and AAC(3)-VIa^{30,32} (note: these were erroneously assigned as members of the GNAT family of AAC enzymes in these publications). Other structurally characterized members of this family include FrbF from *Streptomyces rubellomurinus*³³, YokD from *Bacillus subtilis*, and BA2930 from *Bacillus anthracis*³⁴, none of which possess activity against AGs.

Here, we report a comprehensive structural and functional analysis of the aminoglycoside-resistance spectrum conferred by Antibiotic_NAT family enzymes through characterization of 13 environment-derived enzymes and 8 derived from clinical isolates. This analysis shows that many confer high-level, broad-spectrum aminoglycoside resistance, and five environment-derived enzymes confer apramycin resistance. Crystallographic analysis of various family members, including meta-AAC0038, AAC(3)-IVa, AAC(3)-IIb, and AAC(3)-Xa, allowed the construction of a molecular model explaining the diversification of substrate specificity in this ARG family.

RESULTS

The Antibiotic_NAT family sequences branch into four distinct clades, with all but one including environment-derived members

Identification of new members of the Antibiotic_NAT family through antibiotic selections of soil metagenomic libraries²⁶ prompted a revisit of the sequence diversity of this family. Comparative sequence analysis of the family, including these 14 enzymes derived from environmental microbial communities, 12 Antibiotic_NAT enzymes originating from pathogenic strains, and 25 additional representatives identified by BLAST searches of Genbank, confirmed the presence of conserved sequence motifs typical of Antibiotic_NAT enzymes (**Supplementary Fig. 1**). This analysis also identified highly variable regions that correspond to residues 62-95, 110-117, 127-142, and 190-212 in meta-AAC0038, along with a variable C-terminal region (**Supplementary Fig. 1**). The T χ Φ H Φ AE (where Φ =a hydrophobic residue) sequence motif was previously proposed to contain key catalytic residues of this family^{30,32,33}. The glutamate residue in this motif interacts with the histidine serving to increase the basicity of latter residue. The histidine extracts a proton from the AG 3-*N*-amine group, activating it for nucleophilic attack on the acetyl-CoA carbonyl group^{30,32,33}. The threonine in this motif is thought to stabilize the tetrahedral intermediate. Similar sequence signatures (residues Thr165-Glu171 in meta-AAC0038, **Supplementary Fig. 1**) were identified in all analyzed members of this family, with His and Glu (His168 and Glu171 in meta-AAC0038) along with two glycine residues (Gly122 and Gly158 in metaAAC0038), completely conserved. This motif's threonine (Thr165 in meta-AAC0038) is also conserved in all but one of the analyzed sequences where it is substituted by a chemically similar serine (**Supplementary Fig. 1**)^{30,32,33}.

Bayesian reconstruction of the phylogeny of the Antibiotic_NAT family revealed four main clades (Groups 1 through 4, **Fig. 2**). Enzymes identified by our metagenomic sampling

were distributed among all the clades except for Group 2, which exclusively contains sequences derived from *Actinomycetes*. Several meta-AACs such as meta-AAC0038, meta-AAC0016, and meta-AAC0043 appear to be paralogs of AAC(3)-III, AAC(3)-IVa, and AAC(3)-IIa, respectively.

Pan-family antimicrobial susceptibility testing aligns substrate specificity with phylogeny

To comprehensively characterize the spectrum and degree of resistance conferred by Antibiotic_NAT family members, we tested the antimicrobial susceptibility of *E. coli* individually harboring the 21 different Antibiotic_NAT genes on a low-copy plasmid³⁵. The results (**Fig. 2** and **Supplementary Table 1**) show that the spectrum and degree of AG resistance correlate with the phylogenetic clustering. Group 1 members including AAC(3)-IVa and four meta-AACs confer the broadest spectrum and highest degree of resistance to 4,6- and 4,5-disubstituted AGs, consistent with previous studies on AAC(3)-IVa³¹, and confer high-level resistance to apramycin. We found that the Group 2 member AAC(3)-Xa, derived from an *Actinomycetes*, is limited in its AG specificity to the 4,6-disubstituted AGs kanamycin and tobramycin; the only other Group 2 member tested in our host *E. coli* was AAC(3)-IXa did not convey any detectable AG resistance. Group 3 enzymes including AAC(3)-IIIb and four meta-AACs confer resistance to 4,6- and 4,5-disubstituted AGs, consistent with previous data reported for AAC(3)-III enzymes^{29,30}; meta-AAC0038 is the lone member of this family that confers resistance to apramycin. Group 4 members are restricted in activity to 4,6-disubstituted AGs, including AAC(3)-IIb/IIc and six meta-AAC enzymes, which is reflective of reports on the resistance profile of AAC(3)-VIa^{32,36}; AAC(3)-IIIb also confers low-level apramycin resistance.

Notably, each meta-AAC confers AG resistance, with many demonstrating broad-spectrum and high-level resistance, including against apramycin (meta-AAC0016, meta-AAC0018, meta-AAC0033, meta-AAC0030, and meta-AAC0038).

Crystal structures of meta-AAC0038, AAC(3)-IVa, AAC(3)-IIb, and AAC(3)-Xa enzymes show that the variation in the minor subdomain is responsible for diversity in activity against AGs

We undertook a structural genomics campaign to understand the structural basis of the evident diversification of substrate specificity across the Antibiotic_NAT family, with particular interest in the broadly active Group 1 and meta-AAC enzymes. We solved crystal structures of the AAC(3)-IVa, AAC(3)-IIb, AAC(3)-Xa, and meta-AAC0038 enzymes, including ligand-bound states of AAC(3)-IVa and meta-AAC0038. Crystallographic statistics for all determined structures are shown in **Supplementary Table 2**.

The fold typical of the Antibiotic_NAT family is evident in all structures, composed of 13 α -helices and 8 β -strands (**Fig. 3A**), and determined structures superpose with pairwise RMSD's 0.8-1.0 Å between 197 to 266 matching C α atoms. Notably, the primary sequence most conserved across the family representatives (**Supplementary Fig. 1**) belong to what we defined as a major subdomain in the Antibiotic_NAT fold (**Figure 3B**). In contrast, the variable sequence regions identified by our comparative analysis (see above) constitute a minor subdomain (**Fig. 3B**). According to this distinction, the major subdomain is centered on a 7-stranded antiparallel β -sheet with a bundle of 5 α -helices arranged on one face of the sheet, with a second bundle of 4 α -helices arranged on the other face of the sheet. The minor subdomain is characterized by four main structural variations that are subfamily-specific, which we called inserts 1 to 4. Insert 1 (**Fig. 3B**) forms an extended loop structure of variable length, while adopting a helical structure

in AAC(3)-Xa, meta-AAC0038, and AAC(3)-IIb but not in AAC(3)-IVa. Insert 2 forms a short turn between two α -helices, which most closely impacts the AG binding site. Insert 3 forms a two-stranded antiparallel β -sheet while corresponding to a short α -helix found only in AAC(3)-IIb structure. Finally, insert 4 is a C-terminal extension to the major subdomain unique to AAC(3)-IVa and forms an α -helix and a C3H1 Zn²⁺ binding site. Altogether, this global structural analysis reflects that the minor domain is the principal source of structural diversity among members of this family.

A negatively charged cleft is formed in the region between the minor and major subdomains in each structure, with the deepest section formed primarily by the minor subdomain. As will be discussed in detail later, this cleft harbors the AG binding site.

The Antibiotic_NAT enzymes also diversify in their oligomerization state. The meta-AAC0038 adopt dimeric structure with buried surface of $\sim 900 \text{ \AA}^2$ per subunit (**Fig. 3**). This enzyme also forms dimer in solution according to the size exclusion chromatography (not shown). In contrast, the AAC(3)-Xa enzyme exists as a monomer in solution despite forming a dimer in crystal lattice (**Fig. 3**). AAC(3)-IVa also adopted a dimeric structure (**Fig. 3**) both in crystal and in solution, in line with previous reports on its oligomeric state³¹, but the arrangement of the two chains in this enzyme differed from that of the meta-AAC0038 dimer. The buried surface area between subunits of the AAC(3)-IVa dimer ($\sim 650 \text{ \AA}^2$) was formed nearly exclusively through interactions between the major subdomains of the two monomers of this enzyme. Finally, AAC(3)-IIb was monomeric both in the crystal structure or in solution (not shown).

Structural analysis of the group 1 enzyme AAC(3)-IVa suggests a mechanism for broad specificity against AG substrates

To understand the structural basis of the highly promiscuous nature of group 1 Antibiotic_NAT enzymes, we pursued structural characterization of AAC(3)-IVa representative of this clade in complex with AG substrates. To increase the chances of capturing substrate-bound enzyme complex we used catalytically impaired His154Ala mutant of AAC(3)-IVa.

Using this strategy, we were able to determine the crystal structures of AAC(3)-IVa enzyme in complex with gentamicin or apramycin to 2.6 and 2.8 Å, respectively. In both complex structures, the electron density corresponding to the AG molecule localized to the cleft between major and minor subdomains of the enzyme. Most of the AG substrate interactions with the protein are mediated by amino acid sidechains from the minor subdomain (**Fig. 4A, 4B, 4C**). For the AG substrate in both structures, the 3-*N* group is positioned close to residue 154 and proximal to the presumed location of the thiol of CoA. We observe similar substrate orientation in the crystal structures of meta-AAC0038 enzyme complexes, described below, suggesting a common active site topology for this family.

In the complex structures, the gentamicin molecule spans across enzyme's minor subdomain while the apramycin molecule is twisted nearly 90° relative to gentamicin. This difference is reflected in the rotation of the 2-deoxystreptamine rings of each compound being rotated (**Fig. 4B**). The 2-deoxystreptamine/II ring of apramycin stacks against the sidechain of Trp63, and its rotation positioned the central and III rings more into the minor subdomain cleft and towards Asp67. Notably, these two residues are contributed from the much shorter hairpin connecting the α4 and α5 helices compared to the equivalent region in the other enzymes we crystallized. Additionally, Glu185 appears to be a critical residue for interactions with

gentamicin and apramycin as it positions the 2-deoxystreptamine ring for modification through interactions with the 1-*N* of gentamicin or the 5-hydroxyl of apramycin. Interestingly, Cys190, which is just N-terminal to the Zn²⁺ binding site, interacts with the 3-*N* of gentamicin. Finally, the C-terminal extension of AAC(3)-IVa corresponding to residues 236-257 contributes to the interactions with both gentamicin and apramycin via Glu249 side chain.

We identified a Zn²⁺ ion binding site in the C-terminal extension of AAC(3)-IVa structure. This feature may be of only structural significance since neither this ion nor the sidechains of its cysteine and histidine ligands formed any interactions with the AGs. The binding of Zn²⁺ could stabilize this region and allow for orientation of the Glu249 residue for AG recognition. The Zn²⁺-binding residues are fully conserved across Antibiotic_NAT Group 1 representatives.

The analysis of the AAC(3)-IVa•gentamicin complex allowed to propose a mechanism for this enzyme's ability to recognize 4,5-disubstituted AGs. In the complex structure, gentamicin's 5-OH pointed out of enzyme's active site. If similarly oriented, 4,5-disubstituted AGs would not cause a steric clash with this enzyme's active site. Collectively, these observations show that AGs can adopt multiple bound orientations facilitated by the dramatic structural changes in the minor subdomain of AAC(3)-IVa, thereby supporting broad substrate specificity for AG modification.

The meta-AAC0038 enzyme active site's molecular architecture allows for activity against 4,5 and 4,6-disubstituted AGs

Our data presented above demonstrated that the environmental metagenome-derived meta-AAC0038 enzyme can confer high and broad resistance to AGs including to the atypical

AG - apramycin when expressed in *E. coli*. Using the catalytically inactive His168Ala mutant of this enzyme, we were able to determine the crystal structures of ternary meta-AAC0038^{H168A}•apramycin•CoA and the binary meta-AAC0038^{H168A}•acetyl-CoA complexes.

In line with the previously discussed Antibiotic_NAT enzyme structures, meta-AAC0038 accommodated the substrates in the negatively charged cleft formed by the minor subdomain, with the 3-*N* group of apramycin located within 2.6 Å of the sulfhydryl group of CoA (**Fig. 4C**). Notably, the I and III rings of apramycin were positioned out from the active site cleft and did not form interactions with the enzyme except for hydrogen bonds with the Asp94 and Asp162 sidechains. The ability to retain this AG molecule in the active site via very few contacts could explain the activity of meta-AAC0038 on this substrate resulting in the low-level resistance to apramycin which was not detected for the other representatives of Group 3 Antibiotic_NAT enzymes.

AAC(3)-IIIb, another group 3 enzyme, has been previously characterized in detail for its interactions with 4,6- and 4,5-disubstituted AGs³⁰. The meta-AAC0038 and AAC(3)-IIIb structures superimpose with RMSD 0.54 Å across 219 C α atoms, share all the minor subdomain structural elements, and show complete conservation of AG binding residues (**Fig. 4C**). However, the position corresponding to Glu223 in AAC(3)-IIIb is occupied by Asp213 in meta-AAC0038. Glu223 is positioned at the ring I binding site of apramycin, which may impact the ability of AAC(3)-IIIb to accommodate this AG as a substrate.

The group 4 enzyme AAC(3)-Iib harbors a restricted active site

The crystal structure of AAC(3)-Iib represents the first molecular image of enzymes with AAC(3)-II activity. Its structure superimposes with RMSD 0.7 Å over 221 C α atoms with the

previously characterized AAC(3)-VIa structure³², consistent with our phylogenetic analysis placing both these enzymes in the group 4 of the Antibiotic_NAT family. Similarly to the AAC(3)-VIa enzyme³², the minor subdomain loop of AAC(3)-IIb contains the conserved Asn208, which is predicted to clash with substituents at position 5 of the AG substrate, thereby explaining the lack of activity toward 4,5-disubstituted AGs. Other notable amino acids in the active site of AAC(3)-IIa that may restrict the size and positioning of AG substrates include Tyr66, positioned near the binding location of the double prime ring (**Fig. 1**), and Phe97, positioned near the central 2-deoxystreptamine ring. Altogether, AAC(3)-IIb – like AAC(3)-VIa – harbors a more restricted active site, consistent with its limited AG specificity.

AAC(3)-Xa also harbors a restricted AG binding site

As indicated by our AG susceptibility testing, the activity of AAC(3)-Xa is limited to tobramycin and kanamycin (**Fig. 2**). To rationalize this strict specificity, we modeled the position of kanamycin into the active site of the apoenzyme structure based on the position of gentamicin bound to AAC(3)-IVa. This analysis suggested that gentamicin would not be accommodated due to the Tyr79 and Asp130 residues, which would clash with the 4'-OH group or the methylated 3'-amine of the corresponding AG substrate, respectively. This model also provides a hypothesis for the inability of this enzyme to confer resistance to 4,5-disubstituted AGs, as the 5-substituents would clash with Glu220 of the enzyme. Based on comparative analysis of the AAC(3)-Xa and AAC(3)-IVa•apramycin complex structures, Tyr79 would also introduce a steric clash with this AG in the AAC(3)-Xa active site. Notably, Tyr79, Asp130, Glu220, and adjacent active site residues are highly conserved in Antibiotic_NAT Group 2 (**Supplementary Fig. 1**), suggesting these are critical determinants for restricting the specificity of these enzymes.

Genetic elements adjacent to meta-AACs suggest possible mobilization mechanisms

To investigate the potential for lateral transfer of meta-AACs, we searched for mobile genetic elements (MGEs) on the AAC-encoding contigs. Of the genes recovered through FMG, only one - meta-AAC0043 - is syntenic with multiple MGEs. This sequence is co-localized on our phylogeny (**Fig. 1**) with AAC(3)-Ile, suggesting a close evolutionary relationship. This finding is in line with the observation that all 28 gentamicin-selected FMG contigs annotated with a gene encoding an AAC(3)-II family enzyme were syntenic with at least one MGE. Worryingly, this contig shows extremely high similarity to sequences found in both chromosomes and plasmids of pathogens like *E. coli*, *K. pneumoniae*, *C. freundii*, and *V. cholerae* (**Fig. 5**). Taken together, our analysis demonstrates that representatives of Antibiotic_NAT family encoded by environmental microbiome can be directly mobilized across taxonomic boundaries to convey resistance in clinically important bacterial species.

DISCUSSION

The realization that environmental microbial communities are important reservoirs of ARGs provides keys to understanding the emergence of antibiotic resistance in pathogenic species. For most ARG families, the evolution, transferability, and molecular/structural basis for the activity of their environmental relatives has not been well characterized. Given that antibiotic use in agricultural and other anthropogenic settings represents a significant proportion of global antibiotic deployment, it is vital to understand the scope and breadth of resistance in the broader global resistome, which may select for the evolution and transfer of ARGs. This knowledge is

critical to protecting the potency of our current antibiotic arsenal and designing antibiotics that are less susceptible to ARGs.

The Antibiotic_NAT family has remained understudied compared to other AME families, including AACs from the GNAT family, APHs, and ANTs. The AAC(3)-II and -III representatives of this family are widespread in pathogens, but the molecular basis for AG recognition has only recently been investigated in detail³⁰. There remain significant gaps in our understanding of the structure and function of these AMEs, from both pathogens and environmental microbes.

In this study, we follow on our previous identification of multiple Antibiotic_NAT family members in soil-derived metagenomic libraries²⁶ through detailed structural and functional analysis. Phylogenetic reconstruction of this family rationalized the substrate specificity profiles into four main clades represented by the AAC(3)-IV, AAC(3)-VII/VIII/IX/X, AAC(3)-III, and AAC(3)-II/IV enzymes. We conclusively show that this division is reflected in differences in activity against AG substrates and in structural diversification localized to the minor subdomain of the Antibiotic_NAT fold.

Our structural data includes the crystal structure of AAC(3)-IVa enzyme which is the first molecular image of Group 1 Antibiotic_NAT enzymes. Our extensive structural and functional characterization demonstrates that this enzyme mediates broad-spectrum AG resistance, including to 4,5-, 4,6-disubstituted AGs and the atypical AG apramycin by evolving a more spacious active site. This is achieved by a C-terminal extension and modifications the structure and residue composition of the α 4- α 5 hairpin of the minor subdomain of the enzyme which allow for broad spectrum of AG recognition. The role of the Zn²⁺-binding site in the mechanism of action of AAC(3)-IVa and Group 1 enzymes is the subject of ongoing investigation. After the

structures of AAC(3)-IVa•gentamicin and AAC(3)-IVa•apramycin were publicly available in the PDB, another group performed structure-guided mutagenesis on the enzyme³⁷. This analysis confirmed the Glu185 and Asp187 residues' important roles for interactions with AG substrates, and the role of Asp67 residue in specificity for gentamicin recognition. This group also generated a double mutant Cys247Ser/Cys250Ser, which abrogated resistance to both gentamicin and apramycin, suggesting that Zn²⁺-binding is necessary for substrate recognition. However, since no evidence for the effect of these two mutations on overall stability of this enzyme was provided, the direct effect of Zn²⁺ binding on interaction with AG substrates remains unclear.

According to our sequence analysis the Group 1 members meta-AAC0022, meta-AAC0033, meta-AAC0016, and meta-AAC0018 also share the C-terminal extension, the Zn²⁺-binding residues, and the shorter sequence corresponding to the α 4- α 5 hairpin. Thus, we hypothesize that these enzymes are likely to be active against the wide range of AGs including apramycin.

Antibiotic_NAT Group 3 members showed a high degree of promiscuity, including activity toward the 4,5- and 4,6-disubstituted AGs. Notably, the meta-AAC0038 enzyme was also active against apramycin which inspired our structural analysis of this activity. According to our meta-AAC0038-apramycin complex structure, binding of apramycin to this enzyme differed from its interactions to AAC(3)-IVa. Meta-AAC0038 demonstrated activity analogous to AAC(3)-IIIb and AAC(3)-IIIc enzymes, which belonged to the same clade. Other environment-derived members, including meta-AAC0008, meta-AAC0030, and meta-AAC0071, were similarly active against 4,5- and 4,6-disubstituted AGs.

Representatives of Antibiotic_NAT Groups 2 and 4 were the most restricted in their specificity, and this was reflected in more constrained and smaller active sites, as revealed by the

structures of AAC(3)-IIb and AAC(3)-Xa. The environment derived enzymes of Group 4, including meta-AAC0032, meta-AAC0029, meta-AAC0034, meta-AAC0035, and meta-AAC0043, likewise conferred resistance only to kanamycin and tobramycin. The crystal structure of AAC(3)-IIb features an active site highly like that in AAC(3)-VIa, consistent with the 4,6-disubstituted specificity of Group 4 enzymes.

Additionally, our study expanded the repertoire of AMEs active against apramycin to include six environment-derived enzymes, with the Group 1 members meta-AAC0016, meta-AAC0018, meta-AAC0033, and meta-AAC0022 conferring high-level apramycin resistance. The presence of these enzymes in environmental microbial species may be provoked by widespread apramycin use in agriculture settings. As apramycin is deployed in the clinic, it is important to be mindful of possible further dissemination of these ARGs.

Our analysis of lateral gene transfer signatures in the genetic vicinity of meta-AAC genes indicates that these genes show low potential for mobilization, for the most part, with the notable exception of meta-AAC0043. This conclusion is corroborated by the separation of meta-AAC and AAC(3) sequences in each group within our phylogenetic reconstruction, except for the close clustering of meta-AAC0008 with AAC(3)-IIIa (67% identical at the protein level) and meta-AAC0043 with AAC(3)-IIc (96% identical). While no MGEs were identified in the contig containing the meta-AAC0008 gene, multiple MGEs were present in the contig harboring meta-AAC0043. This proximity strongly suggests that meta-AAC0043 has mobilized into pathogens, manifesting in the enzyme AAC(3)-Ile, conferring resistance to 4,6-disubstituted AGs. This precedent suggests that with further FMG sampling, additional meta-AAC genes may be identified which represent environmental sources of clinically relevant Antibiotic_NAT genes.

The metagenomic, structural, and functional data presented in this study establishes key molecular insights into the molecular basis for AG recognition by all four clades of the Antibiotic_NAT family. This provides deeper understanding of primary sequence signatures important for AG resistance profile conferred by corresponding enzymes. Our observation that environmental members of this family can confer broad, high-level AG resistance and have already mobilized into pathogenic species warrants surveillance and FMG sampling to detect new connections between ARGs in the clinic and the environment.

METHODS

Sequence analysis and phylogenetic reconstruction

Previously identified members of the Antibiotic_NAT family from functional selections of soil metagenomes²⁶ were aligned with clinically isolated AAC(3) enzyme sequences and homologs in Genbank identified by BLAST. Sequence alignment was performed using the Clustal Omega server (EMBL-EBI). The phylogenetic reconstruction was generated from the sequence alignment by MrBayes³⁸ and visualized by using FigTree v1.4.2.

Antibiotic susceptibility testing

Environmental and clinical Antibiotic_NAT sequences were cloned into the low copy plasmid pGDP3. Expression levels of each gene were controlled by the strong, constitutive promoter P_{bla}. Aminoglycoside susceptibility testing was completed in triplicate with our hyperpermeable, efflux-deficient strain *E. coli* BW25113 $\Delta tolC\Delta bamB$ following the Clinical and Laboratory Standards Institute (CLSI) protocols for the microbroth dilution method³⁹. *E. coli*

was cultured in a cation-adjusted Mueller Hinton broth (CAMHB) arrayed in a 96-well format.

The plates were incubated for 18 hrs at 37 °C.

Protein Purification

E.coli BL21(DE3) Gold was used for *meta-AAC0038* and *aac(3)-IVa* overexpression. 3 mL overnight culture was diluted into 1 L LB media containing selection antibiotic ampicillin and grew at 37°C with shaking. The cell culture was induced with IPTG at 17 °C once the OD₆₀₀ reached 0.6-0.8. Cell pellets were collected by centrifugation at 7000 g. Ni-NTA affinity chromatography was used for protein purification. Cells were resuspended in binding buffer [100 mM HEPES pH 7.5, 500 mM NaCl, 5 mM imidazole, and 5% glycerol (v/v)], lysed with a sonicator. Then insoluble cell debris was removed by centrifugation at 30000 g. The soluble cell lysate fraction was loaded on a 4 mL Ni-NTA column (QIAGEN) pre-equilibrated with binding buffer, washed with 250 mL washing buffer [100 mM HEPES pH 7.5, 500 mM NaCl, 30 mM imidazole, and 5% glycerol (v/v)], and N-terminal His₆-tagged protein was eluted with elution buffer [100 mM HEPES pH 7.5, 500 mM NaCl, 250 mM imidazole and 5% glycerol (v/v)]. The His₆-tagged proteins were then subjected to overnight TEV cleavage using 50µg of TEV per mg of His₆-tagged protein in binding buffer and dialyzed overnight against the binding buffer. The His-tag and TEV were removed by rerunning the protein over the Ni-NTA column. The tag-free protein was then dialyzed in crystallization buffer (50 mM HEPES pH 7.5, 500 mM NaCl) overnight, and the purity of the protein was analyzed by SDS-polyacrylamide gel electrophoresis.

Crystallization and Structure Determination

The meta-AAC0038 apoenzyme crystal was grown at room temperature using the vapor diffusion sitting drop method solution containing 20 mg/mL protein, 2.5 M ammonium sulfate, 0.1 M Bis-Tris propane pH 7, and 10 mM gentamicin. For the AG-bound structures of meta-AAC0038 and AAC(3)-IVa, we utilized the catalytically inactive mutants His168Ala and His154Ala. The meta-AAC0038^{H168A}-apramycin-CoA complex was co-crystallized from solution containing 20 mg/mL protein, 20% PEG 3350, 50 mM ADA pH 7, 10 mM apramycin. The AAC(3)-IVa apoenzyme was crystallized as selenomethionine-derivative from a solution containing 30 mg/mL protein, 0.2 M magnesium chloride, 0.1 M Tris pH 8.8, and 25% PEG3350. The AAC(3)-IVa^{H154A}-apramycin complex was co-crystallized from a solution containing 0.1 M Hepes pH 7.6, 30% PEG 1K, and 2.5 mM apramycin; the AAC(3)-IVa^{H154A}-apramycin complex was co-crystallized from a solution containing 0.1 M Hepes pH 7.5, 30% PEG 1K and 1 mM gentamicin.

Diffraction data at 100K were collected at a home source Rigaku Micromax 007-HF/R-Axis IV system, at beamline 21-ID-G of the Life Sciences Collaborative Access Team at the Advanced Photon Source (MAR CCD detector with 300 mm plate), or beamline 19-ID of the Structural Biology Center of the Advanced Photon Source, Argonne National Laboratory. All diffraction data were processed using HKL3000⁴⁰. For meta-AAC0038, the apoenzyme structure was solved by Molecular Replacement (MR), using the structure of YokD³⁴ and the CCP4 online server Balbes program. The apramycin complex structure was used solved by MR using the apoenzyme model. For AAC(3)-IVa, the apoenzyme structure was solved by MR using the structure of FrbF (PDB 3SMA)³³ and the CCP4 online server MoRDa program, and the AG bound structures were solved by MR using the apoenzyme model.

All model building and refinement were performed using Phenix.refine⁴¹ and Coot⁴².

Atomic coordinates have been deposited in the Protein Data Bank with accession codes 5HT0, 6MMZ, 6MN0, 7KES, 6MN3, 6MN4, 6MN5, 7LAO, and 7LAP. Dimerization interfaces were determined using the PDBePISA server⁴³. Structural homologs were identified in the PDB using the Dali-lite server⁴⁴ or the PDBeFold server⁴⁵.

REFERENCES

1. World Health Organization (WHO). Antimicrobial resistance: global report on surveillance. (2014).
2. Surette, M. D. & Wright, G. D. Lessons from the environmental antibiotic resistome. *Annual Review of Microbiology* **71**, 309–329 (2017).
3. Finley, R. L. *et al.* The scourge of antibiotic resistance: The important role of the environment. *Clinical Infectious Diseases* **57**, 704–710 (2013).
4. Peterson, E. & Kaur, P. Antibiotic resistance mechanisms in bacteria: relationships between resistance determinants of antibiotic producers, environmental bacteria, and clinical pathogens. *Frontiers in microbiology* **9**, 243 (2018).
5. Crofts, T. S., Gasparini, A. J. & Dantas, G. Next-generation approaches to understand and combat the antibiotic resistome. *Nature reviews Microbiology* **67**, 7–434 (2017).
6. Stogios, P. J. *et al.* Substrate recognition by a colistin resistance enzyme from *Moraxella catarrhalis*. *ACS Chemical Biology* **13**, 1322–1332 (2018).
7. Kieffer, N., Nordmann, P. & Poirel, L. *Moraxella* species as potential sources of MCR-like polymyxin resistance determinants. *Antimicrobial agents and chemotherapy* **61**, e00129–17 (2017).
8. W, Y. *et al.* Novel plasmid-mediated colistin resistance gene *mcr-3* in *Escherichia coli*. *mBio* **8**, (2017).
9. Zhao, W. H. & Hu, Z. Q. Epidemiology and genetics of CTX-M extended-spectrum β -lactamases in Gram-negative bacteria. *Critical Reviews in Microbiology* **39**, 79–101 (2013).
10. Wright, G. D. Solving the antibiotic crisis. *ACS Infectious Diseases* **1**, 80–84 (2015).
11. Fosso, M. Y., Li, Y. & Garneau-Tsodikova, S. New trends in aminoglycosides use. *MedChemComm* **5**, 1075–1091 (2014).
12. Jackson, J., Chen, C. & Buising, K. Aminoglycosides: how should we use them in the 21st century? *Current Opinion in Infectious Diseases* **26**, 516–525 (2013).
13. Zhang, A. *et al.* Characterization of resistance patterns and detection of apramycin resistance genes in *Escherichia coli* isolated from chicken feces and houseflies after apramycin administration. *Frontiers in microbiology* **9**, 261 (2018).
14. KM, B., TS, E., TR, C., KJ, G. & DJ, N. Characterization of antimicrobial resistant *Salmonella* Kinshasa from dairy calves in Texas. *Letters in applied microbiology* **38**, 140–145 (2004).

15. Magalhaes, M. L. B. & Blanchard, J. S. The kinetic mechanism of AAC(3)-IV aminoglycoside acetyltransferase from *Escherichia coli*. *Biochemistry* **44**, 16275–16283 (2005).
16. Kadlec, K., Fessler, A. T., Hauschild, T. & Schwarz, S. Novel and uncommon antimicrobial resistance genes in livestock-associated methicillin-resistant *Staphylococcus aureus*. *Clinical microbiology and infection* □: the official publication of the European Society of Clinical Microbiology and Infectious Diseases **18**, 745–755 (2012).
17. AD, K. *et al.* In vitro apramycin activity against multidrug-resistant *Acinetobacter baumannii* and *Pseudomonas aeruginosa*. *Diagnostic microbiology and infectious disease* **88**, 188–191 (2017).
18. DM, L. *et al.* Activity of aminoglycosides, including ACHN-490, against carbapenem-resistant Enterobacteriaceae isolates. *The Journal of antimicrobial chemotherapy* **66**, 48–53 (2011).
19. Juhas, M. *et al.* In vitro activity of apramycin against multidrug-, carbapenem- and aminoglycoside-resistant Enterobacteriaceae and *Acinetobacter baumannii*. *Journal of Antimicrobial Chemotherapy* **45**, 568 (2019).
20. M, H. *et al.* In vitro activity of apramycin against carbapenem-resistant and hypervirulent *Klebsiella pneumoniae* isolates. *Frontiers in microbiology* **11**, (2020).
21. S, R. *et al.* Evaluation of apramycin against spectinomycin-resistant and -susceptible strains of *Neisseria gonorrhoeae*. *The Journal of antimicrobial chemotherapy* **74**, 1311–1316 (2019).
22. Ramirez, M. S. & Tolmasky, M. E. Aminoglycoside modifying enzymes. *Drug resistance updates* □: reviews and commentaries in antimicrobial and anticancer chemotherapy **13**, 151–171 (2010).
23. Cox, G. *et al.* Plazomicin retains antibiotic activity against most aminoglycoside modifying enzymes. *ACS Infectious Diseases* **4**, (2018).
24. Becker, B. & Cooper, M. A. Aminoglycoside antibiotics in the 21st century. *ACS chemical biology* **8**, 105–115 (2013).
25. JCK, Q. *et al.* Apralogs: Apramycin 5- O-glycosides and ethers with improved antibacterial activity and ribosomal selectivity and reduced susceptibility to the aminoacyltransferase (3)-IV resistance determinant. *Journal of the American Chemical Society* **142**, 530–544 (2020).
26. Forsberg, K. J. *et al.* Bacterial phylogeny structures soil resistomes across habitats. *Nature* **509**, 612–616 (2014).
27. Xu, Z. *et al.* Structural and functional survey of environmental aminoglycoside acetyltransferases reveals functionality of resistance enzymes. *ACS Infectious Diseases* **3**, (2017).
28. Vakulenko, S. B. & Mobashery, S. Versatility of aminoglycosides and prospects for their future. *Clinical Microbiology Reviews* **16**, 430–450 (2003).
29. Norris, A. L. & Serpersu, E. H. Ligand promiscuity through the eyes of the aminoglycoside N3 acetyltransferase IIa. *Protein science* □: a publication of the Protein Society **22**, 916–928 (2013).
30. Kumar, P., Selvaraj, B., Serpersu, E. H. & Cuneo, M. J. Encoding of promiscuity in an aminoglycoside acetyltransferase. *Journal of medicinal chemistry* **61**, 10218–10227 (2018).

31. Magalhães, M. L. B. & Blanchard, J. S. The kinetic mechanism of AAC3-IV aminoglycoside acetyltransferase from *Escherichia coli*. *Biochemistry* **44**, 16275–16283 (2005).
32. Kumar, P., Serpersu, E. H. & Cuneo, M. J. A low-barrier hydrogen bond mediates antibiotic resistance in a noncanonical catalytic triad. *Science Advances* **4**, (2018).
33. Bae, B., Cobb, R. E., DeSieno, M. A., Zhao, H. & Nair, S. K. New N-acetyltransferase fold in the structure and mechanism of the phosphonate biosynthetic enzyme FrbF. *Journal of Biological Chemistry* **286**, 36132–36141 (2011).
34. Klimecka, M. M. *et al.* Structural analysis of a putative aminoglycoside N-acetyltransferase from *Bacillus anthracis*. *Journal of molecular biology* **410**, 411–423 (2011).
35. Cox, G. *et al.* A common platform for antibiotic dereplication and adjuvant discovery. *Cell chemical biology* **24**, 98–109 (2017).
36. P, K. *et al.* Low-barrier and canonical hydrogen bonds modulate activity and specificity of a catalytic triad. *Angewandte Chemie (International ed. in English)* **58**, 16260–16266 (2019).
37. Plattner, M., Gysin, M., Haldimann, K., Becker, K. & Hobbie, S. N. Epidemiologic, phenotypic, and structural characterization of aminoglycoside-resistance gene *Aac(3)-IV*. *International Journal of Molecular Sciences* **21**, 1–11 (2020).
38. Ronquist, F. & Huelsenbeck, J. P. MrBayes 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics (Oxford, England)* **19**, 1572–1574 (2003).
39. M07: Dilution AST for Aerobically Grown Bacteria - CLSI. <https://clsi.org/standards/products/microbiology/documents/m07/>.
40. Minor, W., Cymborowski, M., Otwinowski, Z. & Chruszcz, M. HKL-3000: the integration of data reduction and structure solution--from diffraction images to an initial model in minutes. *Acta crystallographica Section D, Biological crystallography* **62**, 859–866 (2006).
41. Adams, P. D. *et al.* PHENIX: A comprehensive Python-based system for macromolecular structure solution. *Acta Crystallographica Section D: Biological Crystallography* **66**, 213–221 (2010).
42. Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. Features and development of Coot. *Acta Crystallographica Section D: Biological Crystallography* **66**, 486–501 (2010).
43. Krissinel, E. & Henrick, K. Inference of macromolecular assemblies from crystalline state. *Journal of molecular biology* **372**, 774–797 (2007).
44. L, H. & P, R. Dali server: conservation mapping in 3D. *Nucleic acids research* **38**, (2010).
45. E, K. & K, H. Secondary-structure matching (SSM), a new tool for fast protein structure alignment in three dimensions. *Acta crystallographica. Section D, Biological crystallography* **60**, 2256–2268 (2004).

ACKNOWLEDGEMENTS

This work made use of the Centre for Microbial Chemical Biology core facility at McMaster

University for antibiotic susceptibility testing. We thank Rosa Di Leo for cloning, and Zdzislaw

Wawrzak and Karolina Michalska, Argonne National Laboratory, for x-ray diffraction data collection. This project has been funded in whole or in part with Federal funds from the National Institute of Allergy and Infectious Diseases, National Institutes of Health (NIH), Department of Health and Human Services, under Contract numbers HHSN272201200026C and HHSN272201700060C. This research was supported by the Ontario Research Fund Research Excellence (ORF-RE) Grant No. RE07-048 (to G.D.W. and A.S.) and a Canadian Institutes of Health Research grant (FRN-148463) and a Canada Research Chair to G.D.W. This work is supported in part by an award to G. Dantas through the National Institute of Allergy and Infectious Diseases of the National Institutes of Health (U01 AI123394). A.W. D'Souza received support from the National Research Service Award-Medical Scientist grant to Washington University (T32 GM007200), and the Institutional Program Unifying Population and Laboratory-Based Sciences Burroughs Welcome Fund grant to Washington University. The content is solely the responsibility of the authors and does not necessarily represent the official views of the funding agencies.

AUTHOR CONTRIBUTIONS

PJS solved all crystal structures, completed sequence and structural analyses, prepared figures and wrote the manuscript. EB completed all antibiotic susceptibility testing, prepared figures and wrote the manuscript. ZX performed protein purification and characterization. TS performed protein purification and crystallization. EE performed protein purification and crystallization. SC performed cloning. LTD performed FMG, sequence analysis and prepared figures. AWD and SP performed FMG. GD, GDW and AS oversaw the work and wrote the manuscript.

COMPETING INTERESTS STATEMENT

All authors do not have competing interests.

FIGURE LEGENDS

Figure 1. Chemical structures of aminoglycosides.

The 3-amino group is highlighted in red.

Figure 2. Family-wide antibiotic susceptibility mapped onto phylogenetic reconstruction of Antibiotic_NAT family. The four main groups are separately colored. Sequence names are only shown for meta-AAC, clinical isolates of AAC(3) enzymes, and outgroup members with Antibiotic_NAT fold but with no activity against aminoglycosides (FrbF, YokD, and BA2930); other sequences not labeled are hits from BLAST search of NCBI nr database. Node labels are Bayesian probability values. Right side represents a heatmap of AG susceptibility (fold change MIC relative to control strain containing no resistance element), with key shown at bottom right, full data in **Supplementary Table 1**.

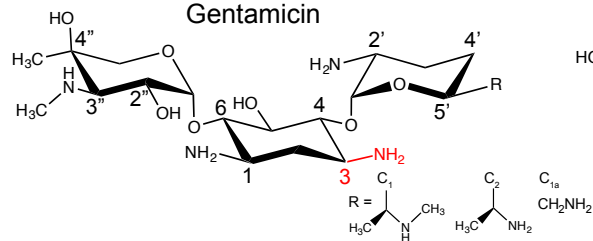
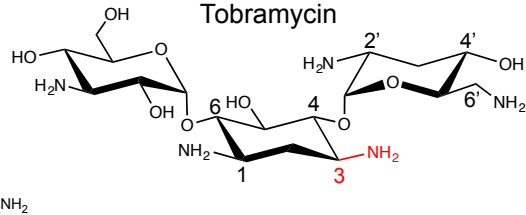
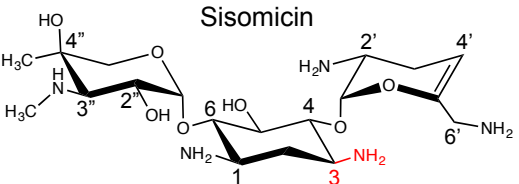
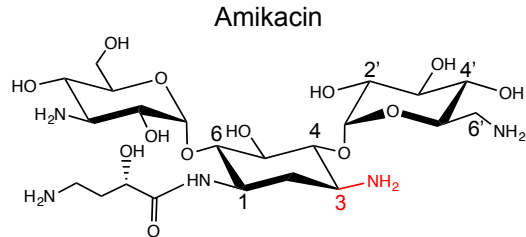
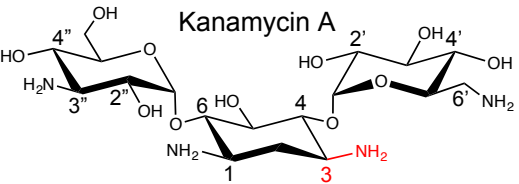
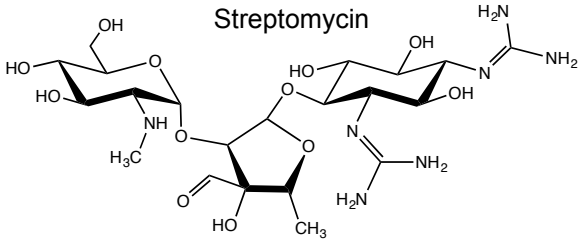
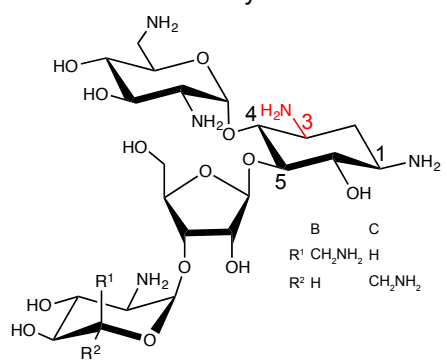
Figure 3. a) Structures of AAC(3)-IVa, AAC(3)-Xa, meta-AAC0038, and AAC(3)-IIb. As representatives of groups 1 through 4, respectively. The conserved major subdomain of the Antibiotic_NAT fold is colored in cyan; the variable minor subdomain is colored in dark blue. Zn²⁺ ion bound to AAC(3)-IVa is shown as a dark grey sphere. Ligands bound to AAC(3)-IVa

and meta-AAC0038 are shown in sticks and labeled. b) Schematic of structural variations in the minor subdomain as insertions or extensions to the major subdomain, numbered 1 through 4.

Figure 4. Details of molecular recognition of aminoglycosides by meta-AAC0038 and

AAC(3)-IVa. a) From solved crystal structures, active sites of AAC(3)-IVa^{H154A} and gentamicin, b) AAC(3)-IVa^{H154A} and apramycin, c) meta-AAC0038^{H168A} and apramycin and CoA. Dashes indicate hydrogen bonds. Since each protein was crystallized with inactive mutants, His168Ala of His154A mutations for meta-AAC0038 and AAC(3)-IVa, respectively, these sidechains shown in this figure are from the apoenzyme structures and indicated with asterisks. Residues colored in dark and light blue are from the major and minor subdomains of the two enzymes, respectively. Acetylation sites (3-*N* groups) are labeled with red arrows.

Figure 5: Synteny of meta-AAC0043 with mobile genetic elements. The contig containing meta-AAC0043 was queried against the NCBI nucleotide database and filtered for highly similar sequences, revealing the presence of similar sequences in a hugely diverse set of taxa. A representative set of similar genomic segments are shown, with grey bars indicating blastn percent identity $\geq 99.5\%$. Many of these matches are from plasmid sequences, and almost all of them contain ORFs annotated as MGEs (e.g., transposons, insertion sequences, etc.). FMG = functional metagenomic library.

Gentamicin**Tobramycin****Sisomicin****Amikacin****Kanamycin A****Streptomycin****Neomycin****Apramycin**