

Comprehensive exploration of the translocation, stability and substrate recognition requirements in VIM-2 lactamase.

Chen, J. Z.¹, Fowler, D.M.^{2,3}, and Tokuriki, N.¹

¹Michael Smith Laboratories, University of British Columbia

²Department of Genome Sciences, University of Washington

³Department of Bioengineering, University of Washington

Abstract

Metallo- β -lactamases (MBLs) degrade a broad spectrum of β -lactam antibiotics, and are a major disseminating source for multidrug resistant bacteria. Despite many biochemical studies in diverse MBLs, molecular understanding of the roles of residues in the enzyme's stability and function, and especially substrate specificity, is lacking. Here, we employ deep mutational scanning (DMS) to generate comprehensive single amino acid variant data on a major clinical MBL, VIM-2, by measuring the effect of thousands of VIM-2 mutants on the degradation of three representative classes of β -lactams (ampicillin, cefotaxime, and meropenem) and at two different temperatures (25°C and 37°C). We revealed residues responsible for expression and translocation, and mutations that increase resistance and/or alter substrate specificity. The distribution of specificity-altering mutations unveiled distinct molecular recognition of the three substrates. Moreover, these function-altering mutations are frequently observed among naturally occurring variants, suggesting that the enzymes has continuously evolved to become more potent resistance genes.

Introduction

The rise of drug-resistant bacterial pathogens has been rapid and inevitable following the introduction of novel antibiotics to the clinic. Pathogens often acquired resistances through horizontal gene transfer (HGT) using mobile genetic elements carrying antibiotic resistance genes, such as plasmids or transposable elements¹⁻³. Under constant selection pressure from antibiotic use, these resistance genes continuously evolve to improve their efficacy and alter and broaden their specificity to other classes of antibiotics^{2,4}. Understanding the molecular mechanisms and evolutionary dynamics of antibiotic resistance genes is crucial to finding sustainable solutions against the future dissemination and evolution of antibiotic resistance, such as through predicting future evolution and aiding in antibiotic and inhibitor design^{5,6}.

Metallo- β -lactamases (MBL), or class B β -lactamases, are one of the major sources for the spread of multi-drug resistance bacteria. MBLs are metal dependent hydrolytic enzymes that degrade a broad spectrum of the widely used β -lactam antibiotics, including “last resort” antibiotics such as carbapenems⁷. Plasmid borne MBLs, such as VIM, NDM, IMP, and SPM-types, have been particularly problematic as they can spread to different bacterial pathogens and have no clinically effective inhibitors⁸. All major MBLs have also been undergoing continual evolution; VIM-type MBLs have diversified up to 70 amino acid mutations (26% sequence difference) into over 50 isolated variants, and some variants seem to have developed new substrate specificity⁹⁻¹³. Much effort has been made to characterize the molecular mechanisms and identify key residues in several major MBLs using diverse biochemical and structural approaches¹⁴⁻²⁴. However, the contributions of the majority of residues in these enzymes remains unexplored, and little is known of the molecular mechanisms governing substrate recognition.

One way to resolve these questions is through comprehensive, large-scale characterizations of mutations affecting the MBLs' efficacy and specificity. Deep mutational scanning (DMS) is a recently developed method for the characterization of thousands of mutations within a protein using deep sequencing^{25–27}. The resulting high-resolution and comprehensive mutational datasets provide invaluable information for deciphering a subset of mutations related to monogenetic disease^{28–30}, understanding evolutionary dynamics of proteins—including viral coat, antibiotic resistance genes and hormone receptors^{31–34}—as well as elucidating protein sequence-structure-function relationships^{35–41}. In particular, conducting DMS on a protein of interest under varying conditions—e.g. against different substrates or in different environments—has further unveiled in-depth molecular details of a protein, such as residues contributing substrate specificity^{42,43} and protein–environment interactions^{44–46}.

In this work, we use DMS to characterize the functional behavior of all ~5600 single amino acid variants of VIM-2 against three classes of β -lactam antibiotics(ampicillin, cefotaxime, and meropenem) and at two different temperatures (25°C and 37°C), and gain deep insights into the molecular and evolutionary determinants of VIM-2's behavior. We generate a series of comprehensive and high-quality datasets, and develop a global understanding of VIM-2 by identifying residues that are critical for its function, stability and/or substrate specificity. We also examine VIM-2's signal peptide—an often overlooked feature despite its importance in expression and transport. Moreover, we use the data to assess their resistance characteristics and rationalize evolutionary outcomes of the clinically isolated natural variants of VIM-type genes, revealing that several mutations in the natural variants are functionally beneficial and lead to changes in substrate specificity.

Results and Discussion

Deep mutational scanning of VIM-2 metallo-β-lactamase

DMS was conducted on a library of VIM-2 variants, each encoding a single amino acid substitution. The wild-type (wt) VIM-2 (UniProt ID: A4GRB6) was sub-cloned into a custom pIDR2 vector with a chloramphenicol resistance marker, where VIM-2 expression is controlled by the constitutive ampR promoter (**Supplementary File 1**). We constructed the library of all possible single amino acid variants of wtVIM-2 through PCR based saturation mutagenesis, where each codon position is mutated to an ‘NNN’ codon using restriction-free (RF) cloning (**Fig. 1a**)⁴⁷; there are 5607 possible variants in the library ((20 a.a. + stop codons) × 267 positions). The plasmids of mutagenized codon libraries were pooled into seven groups—six groups of 39 codons (117 bp, 819 variants in each group) and 1 group of 33 (99 bp, 693 variants)—so the mutagenized region of each group can be covered by Illumina NextSeq deep sequencing. We estimated the mutation rate of our library construction by determining the full sequence of 87 variants by Sanger sequencing: only one nucleotide substitution was found outside the intended codon—which corresponds to a mutation rate of 1.4×10^{-5} —while two other variants had an insertion/deletion. Thus, we constructed a high quality variant library, comparable to other libraries constructed and deep sequenced in a similar manner⁴².

Cultures of *E. coli* cells transformed with VIM-2 libraries (each group was treated separately) were subjected to antibiotic selection by incubating the culture at 37°C with LB media in the presence (selected) and the absence (non-selected) of three different classes of β-lactam antibiotics—ampicillin (AMP), a 3rd generation penicillin, cefotaxime (CTX), a 3rd generation cephalosporin and meropenem (MEM), a carbapenem (**Fig. 1b**). To determine the selection conditions, the growth of *E. coli* cells harboring the plasmid encoding wtVIM-2 was examined at

a range of antibiotic concentrations (1.0-1024 µg/mL AMP, 0.625-64 µg/mL CTX, 0.002-2.0 µg/mL MEM) (**Fig. 1c** and **Supplementary Fig. 1**). We chose to test the highest antibiotic concentration where wtVIM-2 can grow almost 100% relative to growth in media without β-lactam antibiotics, and at successive lower concentrations at 8-fold decrements where the range permits; selected conditions are 128, 16 and 2.0 µg/mL of AMP, 4.0 and 0.5 µg/mL CTX, and 0.031 µg/mL MEM (**Fig. 1c**). The selection process for each antibiotic was conducted in duplicate on separate days. After selection, the plasmids were isolated, the mutagenized region of each group was amplified by PCR, and the amplicons were sequenced by the Illumina NextSeq 550 platform. The sequencing reads were error filtered, and the fitness score of each variant relative to wtVIM-2 was calculated using equation (1). (see methods for “deep sequencing and quality control”).

$$fitness\ score = \text{Log}_2 \left(\frac{\frac{frequency\ of\ variant_{Selected}}{frequency\ of\ variant_{Non-selected}}}{\frac{frequency\ of\ wt_{Selected}}{frequency\ of\ wt_{Non-selected}}} \right) \quad (1)$$

Where the frequency of a variant (or wt) is the deep sequencing read count of the variant divided by the total reads in the corresponding sample. Variants with frequencies below the threshold of deep sequencing errors that was estimated from the deep sequencing of wtVIM2 (see methods for “Variant identification and noise filtering”) were excluded during scoring. The non-selected library shows excellent coverage, with 5535 of 5607 (98.7%) variants present after filtering in at least one replicate while 97.8% are observed in both replicates (**Supplementary Table 1**). For selected libraries, we calculate the fitness score for any variants present in at least one non-selected replicate then average the fitness scores between the two selection replicates (see **Supplementary Data 1** for all fitness scores).

Our DMS experiments show high replicability in all conditions tested. The R^2 of a linear regression between variants observed in both replicates is 0.94 for selection with 128 µg/mL AMP,

0.91 for 4.0 $\mu\text{g/mL}$ CTX and 0.85 for 0.031 $\mu\text{g/mL}$ MEM (**Fig. 2** and **Supplementary Fig. 2**). As expected, variants with synonymous mutations have near neutral fitness and variants with nonsense mutations (stop codons) have the lowest fitness scores. At the highest concentration of each antibiotic, variants with stop codons have fitness scores centered around -4 and lower, thus a fitness score of -4 is considered the lowest score cut-off for downstream analyses (**Fig. 2**). Like previous DMS studies with other proteins, the overall fitness distribution of all variants exhibit a bi-modal distribution with a peak at neutral fitness and a long tail stretching toward negative fitness to another peak at the cutoff of -4 (**Fig. 2**)^{32,36,39,48}.

We confirmed the DMS fitness scores reflect the actual resistance level of variants (**Fig. 2** and **Supplementary Data 2**). We isolated 45 unique variants (61 unique codons), and determined the half maximal effective concentration (EC_{50}) of *E. coli* culture harboring each variant for three antibiotics by measuring antibiotic dose-response curves. We fit the relationship using a sigmoidal function and identify a linear range of correlation for fitness scores within -3.1 to 0.3 for AMP (EC_{50} 27–78 $\mu\text{g/mL}$), -2.8 to 0.6 for CTX (EC_{50} 1.4–4.1 $\mu\text{g/mL}$) and -2.6 to 1.8 for MEM (EC_{50} 0.011–0.058 $\mu\text{g/mL}$), which correspond to a 2.8, 2.9 and 5.0-fold range of EC_{50} values for AMP, CTX and MEM, respectively⁴⁹. All fitness scores outside the linear range are still qualitatively consistent with EC_{50} , where higher fitness scores correspond to higher EC_{50} values and lower fitness scores correspond to lower EC_{50} values.

Global view of VIM-2 enzyme characteristics

The fitness scores for variants selected at 128 $\mu\text{g/mL}$ AMP are shown in **Fig. 3** (See **Supplementary Fig. 3-4** for CTX and MEM). At a glance, there are several interesting trends in the DMS data of VIM-2. Variants with Cys mutations are highly deleterious throughout the catalytic domain (positions 27-266). As wtVIM-2 possesses only one Cys for metal binding,

additional Cys may cause the formation of undesired disulfide bonds, leading to misfolding. Pro variants are also often deleterious, as this residue disrupts secondary structures^{32,36,41}. We found 112 positions (42% of all residues) are highly sensitive to mutations, where over 75% of missense variants (excluding synonymous and nonsense mutations) display a fitness score < -2.0. These positions are likely key requirements for catalytic activity or protein stability and folding in the wtVIM-2. Indeed, these positions include all six active-site metal coordinating residues (His114, His116, Asp118, His179, Cys198, His240), as well as 3-4 residues adjacent to each metal binding residue in the amino acid sequence that are likely to play important roles in the metal configuration and enzymatic function. Additionally, 92% of positions with high mutational sensitivity—including all metal binding residues—are located in the core of the protein (accessible surface area, ASA, of residue < 30%) and 63% are almost completely buried (ASA < 5%), congruent with previous findings^{25,32,37,42,46,50} (**Fig. 4a**).

To examine the distribution of fitness effects (DFE) of VIM-2 variants, we classified the 5291 nonsynonymous variants as having negative (<-0.7), neutral (-0.7 to 0.7) or positive (>0.7) fitness by performing Z-tests of each variant's fitness scores against the fitness distribution of 244 synonymous variants (the null model distribution), adjusting for 5% false discovery rate using the Benjamini-Hochberg procedure. The DFE of VIM-2 variants is similar across all selection antibiotics at the highest screening concentration (128 µg/mL AMP, 4.0 µg/mL CTX, 0.31 µg/mL MEM), with ~65% of variants being negative, ~30% being neutral and ~5% being positive (**Supplementary Fig. 5**). The overall DFE also agrees with observations found in DMS of other enzymes, such as *E. coli* amidase, TEM-1 β-lactamase and levoglucosan kinase^{32,36,38,43}.

Next, in order to determine biophysical properties that explain the fitness scores, a linear model was constructed using the fitness score of variants as the response factor and various

parameters such as ASA, $\Delta\Delta G$ predicted by Rosetta, change in amino acid volume and polarity, and the wt and variant amino acid states as predictors (**Table 1**, **Supplementary Tables 2 and 3** for CTX and MEM models, **Supplementary Data 3** for data used in models). We first modeled each predictor alone with the response, then selected predictors that account for at least 10% of variance in the response ($R^2 > 0.10$) and modeled them in combinations of 2 or more as individual terms and as interacting terms. The final model accounted for the greatest amount of variance using the fewest predictors—i.e. optimized for adjusted R^2 —and combined four predictors (ASA, $\Delta\Delta G$, wt and variant amino acid) capable of explaining 55% of the variation in fitness scores (adjusted $R^2 = 0.55$). The ASA alone can explain 21% of fitness score variation (**Table 1**) and ASA of the wt amino acid alone show a significant correlation ($R^2 = 0.50$) to the average fitness scores of the position, with mutations at exposed residues having less deleterious fitness effects on average (**Fig. 4b**). The $\Delta\Delta G$ explains an additional 18% (**Table 1**) and there is overall correlation between $\Delta\Delta G$ and fitness score while individual predictions are relatively scattered, similar to previous findings that compared fitness to Rosetta folding energies or solubility scores (**Fig. 4c**)^{36,38}. Knowing the wt and variant amino acid can further explain another 10% and 5% of variation respectively^{41,51}. Thus, the results indicate structure and biophysical factors can explain the majority of fitness score tendencies.

Codon and amino acid optimization in the signal peptide

The first 26 residues of VIM-2 has been identified as the signal peptide^{14,15,52}, which is a sequence used to translocate the enzyme to the periplasm, then cleaved after transport^{53–56}. Our DMS data supports the known length of the signal peptide as mutations to Cys are much less deleterious before residue 26, suggesting these positions are not incorporated into the mature enzyme in the periplasm. In general, the signal peptide sequence has an amino terminal (N) region (residues 1-7)

with 1 or more positive residues, a hydrophobic (H) region (residues 8-21) and a carboxy terminal (C) region (residues 22-26) that precedes the cleavage site containing a PXAXS motif (**Fig. 5a**)^{53–56}. The signal peptide is conserved at 17 of 25 positions across all VIM variants, and the remaining are binary differences between the conserved sequences of the VIM-1 and VIM-2 clades (**Fig. 5a**, clades are defined by **Supplementary Fig. 6**).

Mutations in the signal peptide are generally tolerated (64% of nonsynonymous mutations are neutral with 128 µg/mL of AMP) and even beneficial (9.5% of mutations are positive, **Fig. 5b**), which is consistent with a previous DMS study with TEM-1³⁶. In the N-terminal region, mutations to Lys3 are especially deleterious, likely due to the importance of a net positive charge in the N-region for efficient translocation^{53,57,58}. In contrast, Lys7 is tolerant to substitution—in fact, half of the natural VIM variants have a Ser at this position—suggesting that Lys3, rather than Lys7, is critical for translocation. In the H-region, residues Val10 through Ile16 are the most sensitive to mutation, especially when changed to a charged amino acid^{36,53}. The C-region is only strongly affected by the mutation of Leu23 to Cys or Trp, while variants at other positions are neutral, including the PXAXS motif.

Interestingly, variants with evolutionarily conserved residues in the signal peptide often do not have the highest fitness, and we note both residue level and codon level dependencies in fitness (**Fig. 5c**). At most positions, a number of variants with nonsynonymous mutations produce higher fitness than wtVIM-2. Furthermore, at some positions, variants with synonymous mutations have significantly higher (Ile16, Ala18, and Ala20) or lower fitness (Lue5 and Thr13) (**Supplementary Fig. 7b, Supplementary Data 4**). Overall, it appears the signal peptide is less than optimal. In terms of codon level fitness among synonymous variants, we find a significant association between a mutated codon's change in mRNA folding energy and fitness within mutants close to the start

codon, supporting previous findings that avoidance of secondary structure near the start codon is favored (**Supplementary Fig. 7c**, **Supplementary Data 5**, see methods for ‘RNA folding energy calculation’)^{59,60}. We also found that 73 of 143 ‘codon dependent variants’—where any pair of synonymous codon scores differ by more than 2.0—are within the signal peptide, further supporting the importance of codon choice near the start of the coding region⁶¹. The less than optimal residue level fitness of wtVIM-2 may be because we employ *E. coli* as a host while natural VIM variants are often found in *Pseudomonas*⁶², and/or the signal peptide is not selected to produce maximum expression in natural environments. It has been shown that different signal peptides produce variable expression levels and translocation rates for a given protein, both of which affect the final resistance, especially in different host organisms^{23,57,58,63–65}. Furthermore, the signal peptide is frequently mutated in naturally occurring VIM-type variants (see section on ‘Natural VIM variation’ below), suggesting changes in the signal peptide sequences may have played significant roles in dissemination of MBL genes to different hosts and adaptation to higher antibiotic concentrations.

Elucidation of the role of residues in the catalytic domain

We sought to further examine the functional and structural roles of residues in the catalytic domain (positions 27-266) of wtVIM-2. We compare fitness scores between selection in 128 µg/mL and 16 µg/mL AMP, as fitness at different AMP concentrations reflect a residue’s degree of involvement in the enzyme’s stability, expression and/or catalytic activity. Selection was also performed at 25°C in addition to 37°C to examine temperature dependent mutational effects, highlighting residues involved in protein folding and stability; in general, lower temperatures are permissive to variants with poor folding and high aggregation propensity while having a uniform effect on catalytic rate. To assess the role of each residue, we classified all positions in the catalytic

domain into four types: *i*) ‘tolerant’, if 75% of variants are neutral even in the most stringent condition with 128 $\mu\text{g/mL}$ AMP at 37°C, *ii*) ‘essential’, if 75% of variants are highly deleterious even in the least stringent condition with 16 $\mu\text{g/mL}$ AMP at 25°C, *iii*) ‘temperature dependent’, if the difference in the fitness score between 25°C and 37°C is more than 2.0 in either 128 or 16 $\mu\text{g/mL}$ AMP, and *iv*) ‘residue dependent’, if variants are temperature independent (similar scores at the two temperature) and exhibit a range of fitness rather than being mostly neutral or negative (**Fig. 6a-b**, **Supplementary Data 6** for all classifications).

As expected, the 55 ‘tolerant’ positions are scattered around the surface of the protein and are mostly solvent exposed (80% of positions have greater than 30% ASA) (**Fig. 6d**). The 20 ‘essential’ residues includes all six metal binding residues and deeply buried residues (70% have less than 5% ASA), which form the central core of the protein (**Fig. 6c**). This core is further expanded into a larger scaffold by the 93 ‘temperature dependent’ positions that are mostly buried in the structure (76% have less than 30% ASA) and largely hydrophobic—75% of the temperature dependent residues are non-polar (A, G, I, L, P, V) or aromatic residues (F, W, Y) compared to 58% for the entire catalytic domain. The 72 ‘residue dependent’ positions tend to be near the surface or at packing interfaces between α -helices and β -sheets (**Fig. 6d**).

The fitness of variants at temperature and residue dependent positions show equally strong association with the Rosetta predicted $\Delta\Delta G$, indicating both classes of residues have contributions to structural packing (**Supplementary Fig. 8a**). However, the location and hydrogen bonding behavior of each class of residues suggest different functional roles. Essential and temperature dependent residues display an enrichment of sidechain-backbone h-bonding relative to the proportion of h-bonding residues in each class (**Supplementary Fig. 8b**, **Supplementary Data 7**), suggesting—when combined with the formation of a hydrophobic core—these residues are

largely involved in protein folding and stability^{39,46,66}. In contrast, ‘residue dependent’ positions are prominent in the two major active site loops (14/23 positions from 60-68 and 201-214) and on packing interfaces on these loops’ distal faces from the active site. The loop holding metal-binding residues His114, His116 and Asp118, and the helix positioning the loop into the active-site (positions 112-129) are also enriched in residue dependent positions (10/15 non-metal-binding positions), suggesting possible effects on metal and substrate positioning. Thus, many of the ‘residue dependent’ positions are likely to be involved in catalysis through direct or indirect substrate interactions, and also affect the overall shape of the active site.

Distinct recognition for different classes of β -lactam substrates

VIM-2 is known for its broad spectrum activity against all classes of β -lactam antibiotics except monobactams, but how residues achieve substrate recognition remains unknown. We examined mutations that alter substrate specificity to identify wt residues responsible for substrate recognition by comparing fitness scores between the three antibiotics (128 μ g/mL AMP, 4.0 μ g/mL CTX, 0.031 μ g/mL MEM). First, we identified 29 ‘globally adaptive’ variants across 10 positions in the catalytic domain that increase resistance (fitness score >1) to all antibiotics (**Supplementary Data 8**). Residues at positions 47, 55, 66, 68 and 205 each give rise to at least three globally adaptive variants (24 total) while 57, 65, 115, 180 and 201 each give rise to one; 9/10 positions with globally adaptive variants are near the active site, having at least one atom within 15 Å of the active site zinc ions (**Fig. 7a**). Next, we compare fitness scores of different antibiotics in pairs, and identified variants with a change in fitness effect classifications (negative, neutral or positive) combined with a 2.0 fitness score difference between antibiotics. We identified 78 specificity altering variants across 25 positions, with 23/25 positions near the active site (**Fig. 7b**, **Supplementary Table 4** and **Supplementary Data 9** for individual specificity variants). We

confirm the specificity by comparing the fitness scores with the $\log_2(EC_{50\text{ var}}/EC_{50\text{ wt}})$ of the variant in the three antibiotics (**Supplementary Fig. 9**). Of the 25 positions, five are shared by both specificity and globally adaptive variants, and specificity changes are enhanced by the positive fitness at three of these positions (68, 201 and 205). However, most changes in specificity are due to decreases of fitness in one or two substrates, and only 3 variants (R205H/I/V) maintain neutral or higher fitness in all antibiotics^{42,43}. When examining the roles of these positions, we find nine ‘residue dependent’ and one ‘tolerant’ position, as expected for positions that interact with substrate rather than folding⁶⁷. However, the other 15 positions are ‘temperature dependent’ positions, suggesting residues that are involved in substrate specificity are also embedded in the protein core.

Interestingly, there is a strong bias in specificity changes depending on mutations and their positions in the active site. In 21 of 25 positions, specificity variants decrease AMP fitness, while in 10 positions variants decrease CTX fitness and variants decrease MEM fitness at only one position (**Fig. 7b**). This bias is maintained at the level of individual specificity variants, where 96% decrease the fitness in either AMP and/or CTX—29 only decrease AMP, 33 only decrease CTX and 10 decrease both—while only 3 variants decrease MEM. The residues specific to AMP—where mutations to the residue decrease AMP fitness, but are neutral for CTX and/or MEM—are spread around the active site, including the two active site loops (60-68 and 201-214) as well as residues in the protein scaffold. In contrast, the residues specific for CTX are restricted to the two active site loops^{12,13,68–71}; mutations in positions 62, 67 and 68 are fully specific to CTX, while positions 202, 205, 210 and 211 all have mixed specificity for CTX and AMP. To visualize residue-substrate interactions, we overlaid AMP, cefuroxime and MEM substrates in the active site of VIM-2 through alignment with VIM-1 and NDM-1 structures crystallized in complex with

these substrates. Some substrate interactions are apparent from proximity, such as the packing of hydrophobic residues in loop 60-68 to the non-polar, aromatic substituents on the AMP (C6) and cefuroxime (C7) that is missing in most carbapenems (**Fig. 7c-d**). However, Glu202 and Arg205 seem to be in better position to interact with the C2 substituent of MEM and are further from AMP or cefuroxime, yet both residues are specific for AMP and CTX. Moreover, many residues in the protein scaffold that are affecting AMP specificity do not directly interact with the substrates. We hypothesize that these distant residues may contribute to solvent related phenomenon—such as displacement of solvent and/or bridging solvent with ligand to affect substrate binding energy^{72,73}—or alter protein dynamics to affect substrate specificity^{22,74–77}.

Although wtVIM-2 degrades all three β -lactams, our observations suggest that the enzyme interacts with each substrate in a different manner. AMP interacts with many residues around the active site and is the most sensitive to mutations, while CTX specificity relies exclusively on interactions with residues in the active site loops. Interestingly, MEM seems to rely on contacts shared with other antibiotics, which suggest that carbapenem resistance of VIM variants may have coevolved with other antibiotics.

Natural VIM variation favors neutral, adaptive and specificity mutations

Currently, 56 unique VIM-type MBL sequences (including wtVIM-2) have been found on plasmids in β -lactam resistant clinical isolates (**Supplementary Table 5**)^{13,62}. The DMS data of VIM-2 enable us to characterize these naturally occurring mutations comprehensively. We classified these sequences into four clades, represented by VIM-1 (between 25-29 mutations from VIM-2 each, 45 unique mutations total), VIM-2 (1-6 mutations, 31 total), VIM-7 (70 mutations), and VIM-13 (32-33 mutations, 34 total) (**Fig. 8a, Supplementary Fig. 6**), with 131 unique point

mutations relative to VIM-2 across 99 positions (**Supplementary Fig. 10, Supplementary Data 10**).

As expected, the fitness distribution of naturally occurring mutations in the VIM-type MBL for all antibiotics (128 $\mu\text{g/mL}$ AMP, 4.0 $\mu\text{g/mL}$ CTX, 0.031 $\mu\text{g/mL}$ MEM) shows enrichment for neutral and positive mutations (**Fig. 8b**). The trend suggests that natural variants have adapted to higher resistance as 31% of all ‘globally adaptive’ catalytic domain variants in the DMS are also naturally occurring mutations, and 17% of all natural mutations have positive fitness effects for at least one antibiotic. At least 66% of variants have neutral fitness effects within each antibiotic. Natural mutations are disfavored in residues crucial for activity or stability (**Fig. 8c**): Of the 99 mutated positions, only a small proportion of ‘essential’ (1/20 positions) and ‘temperature dependent’ (21/93) positions have been mutated while large proportions of the signal peptide (20/26) and ‘tolerant’ positions (32/55) have been mutated. Interestingly, 44% (11/25) of specificity altering positions have been mutated, which suggests that VIM variants may have changed their substrate specificity during evolution (**Fig. 8d**).

However, 10% of mutations are still highly deleterious (fitness score < -2.0) in 128 $\mu\text{g/mL}$ AMP (6.9% for 4.0 $\mu\text{g/mL}$ CTX, 6.1% for 0.031 $\mu\text{g/mL}$ MEM), indicating other factors that affect natural variation (**Fig. 8b**). These 13 deleterious mutations are spread over 11 positions, where one position is in the signal peptide, and 10 are in the catalytic domain. The signal peptide mutation (K3Q) occurs only in VIM-7 and eliminates the Lys3 critical to translocation, but this is likely neutral as VIM-7 has a S6R mutation that replaces the positive charge. Two mutations are only deleterious to AMP (Q60H, A143T), thus altering substrate specificity. Furthermore, we suspect the neutral I185V mutation (all 55 natural variants have Val185 while our wtVIM-2 has Ile185) acts as a global suppressor^{78,79}, and permit the accumulation of four mutations that are deleterious

in all antibiotics (T139A, T139I, V236G and V255A) as I185V is the only other mutation in these natural variants. The remaining 6 deleterious mutations are likely neutralized by specific intramolecular epistasis, or the background dependence of mutational effects^{80–84}, as these mutations occur in natural variants with at least 25 mutations relative to VIM-2. While such epistasis will hamper our ability to perfectly predict the effect of mutations, the VIM-2 dataset presented in this study contributes to further our understanding of MBL evolution, and can help orient our predictions concerning the emergence of future resistance.

Conclusion

In this work, we report the first comprehensive mutational analysis of an enzyme in the MBL, class B β -lactamase family, one of the most important enzyme family underlying the dissemination of multi-drug resistance to pathogens. We uncover a lack of optimization in the signal peptide of VIM-2, that may be due to codon dependent RNA folding or incompatibility with the host translocation system. Such findings highlight the importance of genome and host context in resistance gene compatibility^{23,60,65}. By performing DMS at various conditions, three different antibiotics, and two temperatures, we enhance the understanding of sequence-structure-function relationships by unveiling a set of mutations for protein stability, catalysis and substrate specificity of VIM-2. We find VIM-2's substrate specificity altering residues to be enriched near the active site, which enables us to elucidate the molecular basis of enzyme-substrate interactions. This finding is in contrast to a previous DMS study that tested multiple substrates: specificity-altering mutations of *E. coli* amidase (amiE) were distributed across the entire structure in a global mode of specificity determination⁴³. Thus, it is likely that each enzyme takes different mechanisms for recognizing diverse substrates. The monomeric MBLs have large, solvent-exposed active site clefts to recognize a wide range of substrates, while amiE which has a small, occluded active site

and a hexameric quaternary structure that potentially favors controlling specificity through packing and subunit interactions⁴³. However, determinants of reaction enantioselectivity in 4-OT—another enzyme active as a hexamer—are concentrated in the active site, which further highlights unique behaviors in different enzymes⁸⁵. Regardless, understanding distinct mode of enzyme-substrate interactions will lead to design and development of new antibiotics and inhibitors to re-sensitize these enzymes. Finally, we demonstrated that VIM-type variants have been continuously evolving by enhancing their resistance as well as altering their substrate specificity in nature. The study of natural variation also reinforces the observation that mutations found to be neutral or beneficial in an experimental setting tend to be enriched in nature as well ^{34,66}. It is likely that many new VIM variants will emerge in the future, and our results will provide a valuable basis to predict likely mutations and estimate the resistance of newly found variants.

Methods

Materials

LB Broth, Miller (BP1426), ampicillin sodium salt (BP1760) and cefotaxime sodium salt (BP29511) were purchased from Fisher Scientific. Meropenem trihydrate (M2574) was purchased from Sigma-Aldrich (Millipore sigma). E. cloni® 10G electrocompetent cells (60061) and E. cloni® 10G chemically competent cells (60107) were purchased from Lucigen Corp. The KAPA HiFi PCR Kit (KK2102) was purchased from KAPA Biosystems Inc., the E.Z.N.A.® Cycle Pure Kit was purchased from OMEGA Bio-tek Inc. and the QIAprep Spin Miniprep Kit was purchased from Qiagen. The NextSeq 500/550 High Output Kit (300 cycles) (20024908) was purchased from Illumina Inc.

Generation of a VIM-2 mutagenized library with all possible single amino acid substitutions

The wild-type (wt) VIM-2 gene including its signal peptide sequence from *Pseudomonas aeruginosa* was synthesized (Bio Basic Inc.) and subcloned into an in-house plasmid, pIDR2 (chloramphenicol resistance) (**Supplementary File 1**), under a constitutive AmpR promoter using *Nco* I and *Xho* I restriction enzymes (Fisher Scientific). The ATG codon in the *Nco* I site was used as the start codon. However, the cut site requires an extra G nucleotide to follow the start codon and an additional Gly (GGA codon) residue was inserted into the second position of the VIM-2 sequence; this extra Gly relative to wtVIM-2 will be labelled as G2 to distinguish it from position 2 in the wt sequence. The pIDR2 plasmid containing the wtVIM-2 gene will be referred to as pIDR2-wtVIM-2.

To generate all single amino acid variants, a library of codon mutants was made for each codon (267 positions) in the wtVIM-2 gene using restriction-free cloning (RFC)⁴⁷ (**Supplementary Fig. 11**). We designed a forward primer for each codon position that contains a degenerate ‘NNN’ codon—using a script used to design primers for the PFunkel method⁸⁶—and a single reverse primer, allowing a PCR to amplify part of the wtVIM-2 gene while incorporating the mutation. The PCR reaction to amplify part of the gene was done using a KAPA HiFi PCR Kit (Kapa Biosystems, Inc.) for 30 cycles of amplification each with denaturation at 98°C for 20 s, annealing at 62°C for 15 s and extension at 72°C for 15 s; 1 ng of pIDR2-wtVIM-2 was used as template in a 20 µL reaction, with 1 µL each of the forward and reverse primer (10 µM). The first PCR products were purified using E.Z.N.A.® Cycle Pure Kit (OMEGA Bio-tek, Inc.). Afterwards, 10 µL of the first PCR product was then used as a primer to extend the entire plasmid, where the cycling conditions were identical to the first reaction except the extension time (90 s) and 1 ng of pIDR2-wtVIM-2 was freshly added as the template. Product from the second PCR was treated with *Dpn* I for one hour at 37°C to degrade the original wtVIM-2 plasmids, and then the amplified plasmids were purified and concentrated by the ethanol precipitation method. Subsequently, the purified plasmids were transformed into *E. cloni*® 10G chemically competent cells (Lucigen Corp.) using the supplier’s recommended heat-shock transformation protocol and plated on LB-Cm (containing 25 µg/mL chloramphenicol) agar plates. We then counted the number of colony forming units (CFU) obtained after the transformation for every mutagenesis library. Using CASTER⁸⁷ and GLUE⁸⁸, we conservatively estimated that at least 700 CFU after transformation is needed to achieve 100% coverage of all 64 codon variants per position. If a transformation met the required CFU, all colonies were collected and the plasmids were purified using QIAprep Spin

Miniprep Kit (QIAGEN N.V.), while those that did not were re-transformed or re-cloned until the count was met.

Antibiotic selection of the VIM-2 mutagenized library

Mutant libraries at individual codons were mixed into seven groups of 39 (33 for the last group) consecutive codons (see “Deep sequencing and quality control”). *E. cloni*® 10G electrocompetent cells (Lucigen Corp.) were transformed with 1 ng of the plasmid DNA from each of the seven groups using the supplier’s recommended electro-transformation protocol and grown overnight in 10 mL LB-Cm shaking at 30°C. We plated 1/1000 of the transformed culture on LB-Cm agar plates to estimate total CFU after transformation. Using CASTER and GLUE, it was estimated that 20,000 CFU are needed to fully cover 2496 codon mutants (64 codons × 39 positions) and all groups transformed had at least 100,000 CFU. The transformed libraries were suspended in LB media and preserved in 1 mL aliquots at -80°C in LB with 25% final volume glycerol.

Antibiotic selection was conducted in duplicate on two separate days. For each experiment, the 1 mL glycerol stock from each group was thawed and grown in 10 mL LB-Cm shaking at 30°C for 16 hrs, with optical density at 600 nm (OD600) of the cell culture reaching ~1.5. The cultures were then diluted by 1000 fold into fresh LB-Cm and grown at 37°C for 1.5 hrs. After 1.5 hrs of growth (OD600 of the culture is between 0.01 and 0.02), 960 µL of each culture was directly introduced to 40 µL of the antibiotics at 25× concentration (final concentrations are 128, 16 and 2.0 µg/mL for AMP, 4.0 and 0.5 µg/mL for CTX, and 0.031 µg/mL for MEM, prepared in water) or water (no selection) into the wells of a 2.2mL deep-well 96 well plate, and grown for 6 hrs at 37°C. The cultures were also selected at the same AMP concentrations or grown without selection at 25°C.

A culture of *E. cloni*® 10G electrocompetent cells transformed with pIDR2-wtVIM-2 was also grown for 6 hours at 37°C.

After placing the cultures under selection, antibiotics and DNA from lysed cells were removed by washing the selected cultures three times using a Biomek 3000 (Beckman Coulter Inc.) liquid handling robot. For each wash, the culture was centrifuged at 4000 RPM for 12 min, and the supernatant was removed using the Biomek. Subsequently, 800 µL of fresh LB was manually dispensed into the wells, the plate was sealed with plastic film, and the pellets were resuspended by vortexing. The resuspended cultures were centrifuged again for the next cycle of the wash. After the final wash, all cultures were propagated overnight shaking at 30°C in 1 mL of LB-Cm. Plasmid DNA was purified from the cultures using a QIAprep 96 Turbo Kit (QIAGEN N.V.).

Determination of half maximal effective concentration (EC_{50})

We isolated 12 variants from codon libraries at positions 55, 62, 67, 68 and 11 variants from codon libraries at positions 205, 209, 210, 211 by transforming the libraries into *E. coli*, plating on agar plates, and picking single colonies. The identity of each variant was obtained by Sanger sequencing. The variants are grown into glycerol stocks in a 2.2 mL 96 well deep well plate; 2 single colonies of the VIM-2 WT and empty vector were also placed on this plate as controls.

The variants in the plate were then placed under the same liquid culture selection procedure as the DMS selection experiments (see “Antibiotic selection of the VIM-2 mutagenized library” above), up to the end of the 6 hrs of selection where the cell growth (OD₆₀₀) was measured. The range of selection is 1 – 1024 µg/mL for AMP, 0.0625 – 64 µg/mL for CTX and 0.002 – 2 µg/mL for MEM, separated in 2 fold increments. All variants were also grown without antibiotics as a growth control. We calculate the EC_{50} by fitting **equation (2)** using the “curve_fit” function of the “Scipy.optimize” package.

$$\% \text{ growth} = \text{bottom} + \frac{\text{top} - \text{bottom}}{1 + \left(\frac{EC_{50}}{\text{drug concentration}} \right)^{\text{Hill Coefficient}}} \quad (2)$$

Initial estimates were 100% for top, 0% for bottom, -1.0 for the Hill coefficient and 1.0 for the EC_{50} . In the case where a variant's growth curve did not produce a successful fit based on the initial estimate, only the Hill coefficient and EC_{50} were adjusted until the fit was successful. Variants where the EC_{50} do not appear in the growth curve (stop codons, highly deleterious mutations and empty vector) could not be fitted and were excluded.

The DMS fitness scores (y-values) for each antibiotic were fitted against the EC_{50} (x-values) using a similar sigmoidal curve in **equation (3)** with the “curve_fit” function.

$$\text{DMS fitness score} = \text{bottom} + \frac{\text{top} - \text{bottom}}{1 + \left(\frac{x_0}{EC_{50}} \right)^{\text{Hill Coefficient}}} \quad (3)$$

The initial guesses for top, bottom and Hill coefficient are 2.0, -4.0 and 1.0 respectively for all antibiotics. The initial estimate for x_0 , the inflection point of the curve, was 64 $\mu\text{g/mL}$ for AMP, 4.0 $\mu\text{g/mL}$ for CTX and 0.031 $\mu\text{g/mL}$ for MEM.

The linear region of the sigmoidal curve for the DMS fitness scores was calculated using **equations (4-7)**, based on the final fitted values for each antibiotic⁴⁹.

$$Y_{\text{lower}} = \text{top} + \left(\frac{\text{bottom} - \text{top}}{1 + 1/4.6805} \right) \quad (4) \quad Y_{\text{upper}} = \text{top} + \left(\frac{\text{bottom} - \text{top}}{1 + 4.6805} \right) \quad (5)$$

$$X_{\text{lower}} = x_0 \left(\frac{\text{bottom} - Y_{\text{lower}}}{Y_{\text{lower}} - \text{top}} \right)^{\frac{1}{\text{Hill Coefficient}}} \quad (6) \quad X_{\text{upper}} = x_0 \left(\frac{\text{bottom} - Y_{\text{upper}}}{Y_{\text{upper}} - \text{top}} \right)^{\frac{1}{\text{Hill Coefficient}}} \quad (7)$$

Deep sequencing and quality control

We grouped individual codon libraries into seven groups of 39 consecutive codons (33 for the last group) so that all mutations are within a distance of 117 bp (99 bp for the last group), allowing 150 bp forward and reverse NGS sequencing reads to generate full overlap of each group. PCR amplicons of each library group and wtVIM-2 were generated using primers that flank the 117 bp region of each group, where primers have a Nextera transposase adapter sequence (Illumina, Inc.) in the 5' overhangs. We used KAPA HiFi PCR Kit (Kapa Biosystems, Inc.) for 15 cycles of amplification each with denaturation at 98°C for 20 s, annealing at 65°C for 5 s and no extension time. The amplicons were extended by PCR again to include the sample indices (i7 and i5) and flow cell binding sequence, then sequenced using a NextSeq 550 sequencing system (Illumina, Inc.); all samples were sequenced in the same NextSeq run. The raw sequencing data can be found on the NCBI Sequencing Read Archive (SRA) (BioProject accession: PRJNA606894). Each group under each condition received between 400,000 to 1,000,000 reads, which is at least 160 reads per codon variant on average. There were 5 samples that received as low as 100,000 reads in one of the two replicates, but the correlation of fitness scores between both replicates still showed an R^2 of at least 0.72 and the scores were retained.

To process the NGS data, including merging paired-end reads, quality filtering, variant identification and fitness score calculation, we use a set of in-house Python scripts (<https://github.com/johnchen93/DMS-FastQ-processing>). Paired-end reads in FastQ format were first merged using a Python script, where quality (Q) scores of matching read positions were combined using a posterior probability calculation to obtain posterior quality (Q) scores, measured on the Phred scale for sequencing quality⁸⁹. In the case of a base mismatch between the forward and reverse reads, the base was taken from the read with the higher Q score at the position.

Reads that had more than 20 base mismatches between the forward and reverse reads, or that had any bases with a posterior Q score less than 10 were discarded. It was found that above a Q score of 10 or more, the average proportion of sequencing errors per position stabilized and no sizeable reduction of sequencing errors can be obtained by Q score cut-offs (**Supplementary Fig. 12a**). Usually, 75-85% of all reads passed these filters. Additionally, the expected number of errors per read was calculated from adding the error rates calculated from the posterior Q Scores of every position in the entire read⁸⁹. Reads that had an expected number of error greater than 1 would also be discarded, however no reads exceeded this limit after the previous filters.

Variant identification and noise filtering

Once forward and reverse reads were merged and filtered by read quality, codon mutants were identified and counted, then aggregated into amino acid (or stop codon) variants. Codon mutations were identified by comparing to the wtVIM-2 sequence as a reference. Since we only intended for single codon mutants in the library, any sequence with mutations in more than one codon was discarded, leading to retention of 80-90% of the filtered reads.

To exclude variants that may be due to sequencing errors alone, we estimated the expected frequency of each variant generated by sequencing errors and excluded variants that have less than 2× the expected frequency in the non-selected library. Using sequencing data from wtVIM-2, we calculated the error rates that originates from culture growth, sample preps (PCRs) and sequencing. The error rates at each position was calculated by dividing the errors observed by the total number of reads at that position (**Supplementary Fig. 13d**). The mean of the distribution of the positional errors was used as the estimate for error rates (0.072%) in all positions across the VIM-2 gene. The proportion of each type of nucleotide error (A>T, A>C, A>G, etc.) was calculated to estimate

the likelihood of each type of nucleotide error given a starting nucleotide (**Supplementary Fig. 12b**).

We made the observations that 1) sequencing errors in wt sequences will generate single codon mutants, but errors in single codon variants are most likely to be turned into double codon mutants (**Supplementary Fig. 13a**) 2) ~5-10% of reads in each library group were occupied by wt sequences while other variants are rarely higher than 0.5% (**Supplementary Fig. 13b**) and 3) single nucleotide sequencing errors are the most abundant type of errors affecting up to 10% of all reads, while higher numbers of sequencing errors are nearly negligible (**Supplementary Fig. 13c**). In summary, single nucleotide errors on wt sequences are the main source of single codon variants arising from sequencing errors. Thus, we first calculated the chance of each of the 64 codons to mutate into the 9 adjacent codons by single nucleotide sequencing errors using **Equation (8)** (**Supplementary Fig. 13e**).

$$\begin{aligned} \text{chance of codon with substitution } X \rightarrow Y = \\ \text{error rate per position} \times \text{frequency of substitution } X \rightarrow Y \end{aligned} \quad (8)$$

For example, to gauge how often AAA gets mutated to GAA by chance, we multiplied the per position error rate by the proportion of G mutations when starting from an A, leading to $0.0719\% \times 70.7\% = 0.000719 \times 0.707 = 0.00051$. This means for each 100,000 wt reads that has an AAA codon at a given position, we expect 51 GAA mutants on average that arise by chance at the same position. We calculated expected codon error frequencies from every codon, then summed the expected error frequencies of the codons mutant for each amino acid variant (**Supplementary Fig. 13f**). The error frequency is multiplied by the count of the wt reads in each non-selected library group to arrive at an expected error count for that group. Subsequently, we compared the observed

count of amino acid (or codon) variants in the non-selected library to the expected count from errors alone and we accepted a variant as truly existing if the observed count is at least twice the expected count from errors. In addition, because our filtering method only accounts for the 9 codons with a single mutation relative to the wt codon, we also applied a count cut-off of 5 for all variants to reduce noise by excluding very low count data.

Fitness Score Calculations

The fitness score of each variant was calculated according to **equation (1)** (see main text). To calculate the fitness score of a given amino acid (or codon) variant, the read count of the variant was first normalized to frequencies within the non-selected or selected library group. Variants that exist in the non-selected library but disappear in the selected library were interpreted to have been removed by antibiotic selection, and were given a dummy count of 1 to emulate the minimum frequency observable for that variant. The frequency of the variant after selection was divided by the frequency of the same variant in the library grown without selection for an enrichment ratio; synonymous codon variants were also considered as variants rather than wt during scoring. The variant enrichment ratio was then normalized to the enrichment ratio of the wt. The final score was expressed in Log_2 units, and scores were calculated separately across the 7 groups and separately for each replicate.

When combining all data across the 7 groups, we subtracted the mean fitness scores of all synonymous variants in each group from all variants of that group to center the mean fitness of synonymous variants at a fitness score of 0. To combine scores from replicates, we simply averaged the fitness scores across the two replicates, and take the single score if only one replicate contained the variant above noise in the non-selected library.

Fitness effect classification

To classify each amino acid variant as positive, neutral or negative for each of the three selection antibiotics, we use the score of each variant in a two tailed z-test on a normal distribution (null-model) with the same mean and standard deviation as our synonymous distribution (244 synonymous variants total). The p-values were then FDR corrected to an α of 0.05 using the Benjamini-Hochberg procedure; only missense variants were tested and the total number of tests was 5291. The variants with scores that are significantly different from the synonymous distribution after FDR correction are then classified as ‘positive’ if their score is greater than the synonymous mean and ‘negative’ if the score is less, while the remaining variants are classified as ‘neutral’.

Linear Model of DMS Scores with various predictors

We generated a linear model in R using a combination of terms to try and find properties that best explain the behavior we see in the DMS fitness scores. Using the fitness score as a response, we tried using **1)** wild-type (wt) amino acid **2)** variant (var) amino acid **3)** accessible surface area (ASA) of the residue calculated from the crystal structure of wt VIM-2 (PDB: 4bz3) using ASA view⁹⁰ **4)** change in amino acid volume⁹¹ ($\Delta\text{volume} = \text{volume}_{\text{var}} - \text{volume}_{\text{wt}}$) **5)** change in amino acid polarity (hydrophathy index⁹²) ($\Delta\text{polarity} = \Delta\text{polarity}_{\text{var}} - \Delta\text{polarity}_{\text{wt}}$) **6)** distance of the alpha carbon of each residue in the crystal structure to the active site water held between the Zn ions **7)** Rosetta predicted stability change between the variant and the wt ($\Delta\Delta G = \Delta G_{\text{var}} - \Delta G_{\text{wt}}$) (also see “Rosetta $\Delta\Delta G$ Calculation” below) and **8)** BLOSUM62 score for the substitution from wt to variant. Only variants from positions observable in the crystal structure were modelled (positions 32 to 262), and synonymous variants were excluded.

All parameters were first modelled individually as predictors with DMS fitness score as the response, and the predictors with R^2 higher than 0.10 are then modelled in combinations of two or more until the combination with the least predictors and the highest adjusted R^2 was found. Predictors with R^2 less than 0.10 are also retried in combination with the best predictors when optimizing for adjusted R^2 . Interaction between predictors were tested, but they did not improve adjusted R^2 and were excluded for the sake of simplicity. The relative contribution of each term to the overall adjusted R^2 were calculated using the R package ‘RelImpo’, using the ‘lmg’ method⁹³.

The final equation of the linear model is shown in **equation (9)** where the fitness score of a given variant is the additive combination of the model intercept β_0 and the various properties and the coefficients of the properties (e.g. β_{ASA} and ASA) plus a random error term ϵ .

$$\text{Fitness Score} = \beta_0 + \beta_{ASA} \times ASA + \beta_{\Delta\Delta G} \times \Delta\Delta G + \beta_{wt} \times wt + \beta_{var} \times var + \epsilon \quad (9)$$

The categorical predictors wt and var are simplified in the equation and each is actually a collection of terms in the model, where every amino acid is a single binary term represented by 0 or 1 such that 1 indicates the presence of the amino acid. For example, the variant Q60V has $Q = 1$ for wt and $V = 1$ for var , while all other wt and var amino acids are set to 0. Ala is not present as one of the estimates in either wt or var , because they are used in calculating the intercept; this is the mean fitness of all data points where either $wt = 1$ or $var = 1$ for alanine.

Rosetta $\Delta\Delta G$ calculation

To estimate the effects of each VIM-2 variant on the stability of the protein, we used the Rosetta “ddg_monomer” application to calculate the folding energy of a monomeric protein crystal structure. Rosetta was run on the Compute Canada server Cedar using a Rosetta 3.8 installation. Following the ‘ddg_monomer’ documentation, the VIM-2 structure (PDB: 4bz3) was first

processed using ‘preminimize’ to pre-optimize the packing of the crystal structure and generate a constraints file. Then, all single amino acid variant structures and the wt structure at each position were simulated 50 times each using ‘ddg_monomer’, configured to protocol 16 as specified in Kellog *et al.* 2011⁹⁴ while using Talaris 2014 as the scoring function. We store the simulated structures (variant and wt) as PDB files and scored them using the Rosetta ‘score’ function with Talaris 2014 weights to obtain the predicted ΔG in Rosetta Energy Units (REU). We average the predicted ΔG of all 50 replicates of each variant or wt. The $\Delta\Delta G$ is calculated using **equation (10)** as the difference in average ΔG between variant and wt at the same position.

$$\Delta\Delta G = \Delta G_{\text{var}} - \Delta G_{\text{wt}} \quad (10)$$

RNA folding energy calculation for single codon mutants

The RNA folding energy contribution of the 5’ UTR and signal peptide region is calculated according to a previously described method⁵⁹. The ΔG of folding of the 5’ UTR and signal peptide is calculated using **equation (11)**.

$$\Delta G_{1,118} = \Delta G_{1,841} - \Delta G_{119,841} \quad (11)$$

Each ΔG term is calculated using the NUPACK software package⁹⁵, using the ‘pfunc’ program which calculates the ΔG of all RNA secondary structures from the partition function. All folding energies were calculated using default conditions, with [Na+] = 1 M and T = 37°C, and the results are in units of kcal mol⁻¹. The subscripts for each ΔG term indicates the first and last nucleotide position in the transcript that is used for calculating the ΔG of folding, respectively. Thus, $\Delta G_{1,841}$ represents the calculated folding energy of the full transcript (including the 5’ UTR and coding region, but excluding the 3’ UTR), while $\Delta G_{119,841}$ is the folding energy of the transcript after the signal peptide. The interpretation of $\Delta G_{1,118}$ is that it is the folding energy contribution of

the RNA transcript up to the end of the signal peptide, including energy from non-local interactions with downstream parts of the transcript but excluding folding energy of interactions exclusively within positions downstream of the signal peptide. The DNA sequence of the wtVIM-2 transcript used to calculate the folding energies is shown below, with positions 1-118 italicized and the translated signal peptide region underlined for clarity; the sequence is converted to RNA for calculation.

5'-

*CTGATAAATGCTTCAATAATATTGAAAAAGGAAGCCC**ATGGGATTCAAAC**TTTTGAGTAAGT*
TATTGGTCTATTTGACCGCGTCTATCATGGCTATTGCGAGCCCGCTCGCTTTTTCCGTAGA
TTCTAGCGGAGAATATCCGACAGTCAGCGAAATTCCGGTCGGGGAGGTCCGGCTTTA
CCAGATTGCCGATGGTGTGTTGGTCGCATATCGCAACGCAGTCGTTTGATGGCGCAGT
CTACCCGTCCAATGGTCTCATTGTCCGTGATGGTGATGAGTTGCTTTTGATTGATACA
GCGTGGGGTGCGAAAAACACAGCGGCACTTCTCGCGGAGATTGAGAAGCAAATTGG
ACTTCCTGTAACGCGTGCAGTCTCCACGCACTTTCATGACGACCGCGTCGGCGGGCGT
TGATGTCCTTCGGGCGGCTGGGGTGGCAACGTACGCATCACCGTCGACACGCCGGCT
AGCCGAGGTAGAGGGGAACGAGATTCCCACGCACTCTCTTGAAGGACTTTCATCGA
GCGGGGACGCAGTGCCTTCGGTCCAGTAGAACTCTTCTATCCTGGTGCTGCGCATT
CGACCGACAACCTTAATTGTGTACGTCCCGTCTGCGAGTGTGCTCTATGGTGGTTGTG
CGATTTATGAGTTGTCACGCACGTCTGCGGGGAACGTGGCCGATGCCGATCTGGCTG
AATGGCCACCTCCATTGAGCGGATTCAACAACACTACCCGGAAGCACAGTTCGTC
ATTCCGGGGCACGGCCTGCCGGGCGGTCTTGACTTGCTCAAGCACACAACGAATGTT
GTAAAAGCGCACACAAATCGCTCAGTCGTTGAGTAA-3'

Equation (12) is used to calculate the $\Delta\Delta G$ of folding upon codon mutations in the signal peptide.

$$\Delta\Delta G_{\text{sigpep RNA}} = (\Delta G_{1,118 \text{ var}} - \Delta G_{1,118 \text{ wt}}) / (kT) \quad (12)$$

$\Delta G_{1,118 \text{ wt}}$ is the folding energy of the wtVIM-2 signal peptide sequence shown above, while $\Delta G_{1,118 \text{ var}}$ is the folding energy of the signal peptide sequence with a single codon mutation. The difference in folding energy is normalized to the thermal energy factor kT , where $k = 0.0019872041 \text{ kcal mol}^{-1} \text{ K}^{-1}$ and $T = 310.15 \text{ K}$ (37°C).

Identification of critical residues and temperature dependence

We classify the role of residues in wtVIM-2 by examining the DMS fitness scores of selection conducted at 128 $\mu\text{g/mL}$ and 16 $\mu\text{g/mL}$ AMP at 37°C and 25°C (4 data sets), excluding signal peptide residues 1-26. Residues are classified as ‘essential’ when >75% of variants are below a fitness score of -2.0 when selected at the least stringent condition of 16 $\mu\text{g/mL}$ and 25°C. Residues are classified as ‘tolerant’ when >75% of variants are above a fitness score of -1.0 when selected at the most stringent condition of 128 $\mu\text{g/mL}$ and 37°C. Residues are classified as ‘temperature dependent’ when the fitness score of 25°C selection is higher than 37°C selection of the same AMP concentration by at least 2.0, for two or more variants at either 128 $\mu\text{g/mL}$ or 16 $\mu\text{g/mL}$. The ‘temperature dependent’ classifications overwrite ‘essential’ or ‘tolerant’ classifications (4 occurrences total). Residues that do not fall into the three other classifications are defined as ‘residue dependent’.

Detecting hydrogen bonds in the wtVIM-2 crystal structure

Potential hydrogen bonding pairs in the wtVIM-2 crystal structure (PDB: 5yd7, chain A only) were extracted using the ‘Polarpairs’ script in PyMol (<https://pymolwiki.org/index.php/Polarpairs>). The script filters for pairs of h-bond donor and acceptor atoms within a defined distance and h-bond angle; we set the distance limit to be within 3.6 Å and the h-bond angle to be greater than 63°. The script returns atom indices, which were converted to PDB atom IDs using pymol’s built in

‘id_atom’ function. The atoms were then extracted from the 5yd7 pdb file using the atom IDs, and the atom’s name was used to determine if the h-bond was formed between backbone atoms only (‘N’ and ‘O’ atoms indicate backbone amides and carbonyls, respectively), between sidechain atoms only or between backbone and sidechain atoms. All extracted h-bonds can be found in **Supplementary Data 7**.

Analysis of specificity variants

We define variants with altered specificity by filtering for variants with a change in fitness effect classifications (positive, neutral and negative, defined in “Fitness effect classification”) as well as a fitness score difference of 2.0 between at least 2 of the 3 antibiotics being compared.

Specificity positions are visualized on the wtVIM-2 structure (PDB: 5yd7) using PyMol, with substrates overlaid from aligned MBL homolog structures (AMP - NDM-1(PDB:4hl2), cefuroxime – NDM-1(PDB:4rl0), MEM – VIM-1 (PDB:5n5i)). To overlay the substrates, the 6 metal binding residues (structure positions 114, 116, 118, 179, 198, 240 for VIM-1/VIM-2 and 120, 122, 124, 189, 208, 250 for NDM-1) and the 2 active-site Zn ions were selected from each structure, and the PyMol ‘align’ function was used with the VIM-2 active-site as the target object and the other structure’s active-site as the mobile object. When structures have more than one chain (PDB: 5yd7, 4hl2 and 4rl0), only chain A was used in the alignment. The protein portions of the homolog structures are hidden after alignment, to visualize just the substrates with the wtVIM-2 structure.

Collection of naturally occurring VIM variants

Amino acid sequences of naturally observed VIM variants were extracted by performing a BLASTP search of the NCBI non-redundant protein database ⁹⁶, using the protein sequence of our in-house VIM-2 with Gly removed from the 2nd position, identical to the VIM-2 discovered in

Pseudomonas aeruginosa isolates (UniProt accession: A4GRB6). We retained all BLASTP results with at least 70% identity and >90% query coverage. We also retrieved protein sequences from the Comprehensive Antibiotic Resistance Database (CARD) ⁶². All sequences from both sources were merged and sequences that are exactly identical in length and sequence were combined, while sequences that are less than 250 or greater than 290 residues were excluded. Recombinant variants VIM-12 and VIM-25 were excluded from this analysis, as well as VIM-14 (UniProt accession: Q6GUL7) which is a member of the VIM-1 clade despite being labelled as both VIM-11 and VIM-14 (UniProt accession: A0SWU7) (both are in the VIM-2 clade).

To identify all mutations different between wtVIM-2 and the 55 other variants, a multiple sequence alignment (MSA) was constructed using the MUSCLE method in MEGA 7 (version 7.0.26)⁹⁷. Equivalent positions bearing a different amino acid from VIM-2 were identified as mutations, while deletions are ignored (1 in VIM-1, 1 in VIM-7, 4 in VIM-18). The MSA was also used to generate a maximum likelihood phylogenetic tree using MEGA 7 (default settings). The tree was used to identify separate VIM clades which were labelled using the VIM variant with the lowest number in the clade (VIM-1, VIM-2, VIM-7, VIM-13).

References

1. Wright, G. D. Environmental and clinical antibiotic resistomes, same only different. *Curr. Opin. Microbiol.* **51**, 57–63 (2019).
2. Surette, M. D. & Wright, G. D. Lessons from the Environmental Antibiotic Resistome. *Annu. Rev. Microbiol.* **71**, 309–329 (2017).
3. Codjoe, F. S. & Donkor, E. S. Carbapenem Resistance: A Review. *Med Sci (Basel)* **6**, (2017).

4. Current Epidemiology and Growing Resistance of Gram-Negative Pathogens. *Korean J Intern Med* doi:10.3904/kjim.2012.27.2.128.
5. Crofts, T. S., Gasparrini, A. J. & Dantas, G. Next-generation approaches to understand and combat the antibiotic resistome. *Nat. Rev. Microbiol.* **15**, 422–434 (2017).
6. Tehrani, K. H. M. E. & Martin, N. I. β -lactam/ β -lactamase inhibitor combinations: an update. *Medchemcomm* **9**, 1439–1456 (2018).
7. Bebrone, C. Metallo- β -lactamases (classification, activity, genetic organization, structure, zinc coordination) and their superfamily. *Biochemical Pharmacology* **74**, 1686–1701 (2007).
8. Nordmann, P. & Poirel, L. Emerging carbapenemases in Gram-negative aerobes. *Clinical Microbiology and Infection* **8**, 321–331 (2002).
9. Yan, J.-J. *et al.* Metallo- β -Lactamases in Clinical *Pseudomonas* Isolates in Taiwan and Identification of VIM-3, a Novel Variant of the VIM-2 Enzyme. *Antimicrobial Agents and Chemotherapy* **45**, 2224–2228 (2001).
10. Schneider, I. *et al.* VIM-15 and VIM-16, Two New VIM-2-Like Metallo- β -Lactamases in *Pseudomonas aeruginosa* Isolates from Bulgaria and Germany. *Antimicrobial Agents and Chemotherapy* **52**, 2977–2979 (2008).
11. Bogaerts, P. *et al.* Detection and Characterization of VIM-31, a New Variant of VIM-2 with Tyr224His and His252Arg Mutations, in a Clinical Isolate of *Enterobacter cloacae*. *Antimicrobial Agents and Chemotherapy* **56**, 3283–3287 (2012).
12. Mojica, M. F. *et al.* Exploring the Role of Residue 228 in Substrate and Inhibitor Recognition by VIM Metallo- β -lactamases. *Biochemistry* **54**, 3183–3196 (2015).
13. Martínez-García, L., González-Alba, J. M., Baquero, F., Cantón, R. & Galán, J. C. Ceftazidime Is the Key Diversification and Selection Driver of VIM-Type Carbapenemases. *mBio* **9**, (2018).

14. Lauretti, L. *et al.* Cloning and Characterization of blaVIM, a New Integron-Borne Metallo- β -Lactamase Gene from a Pseudomonas aeruginosa Clinical Isolate. *ANTIMICROB. AGENTS CHEMOTHER.* **43**, 7 (1999).
15. Franceschini, N. *et al.* Purification and Biochemical Characterization of the VIM-1 Metallo- β -Lactamase. *ANTIMICROB. AGENTS CHEMOTHER.* **44**, 5 (2000).
16. Wommer, S. *et al.* Substrate-activated Zinc Binding of Metallo- β -lactamases: PHYSIOLOGICAL IMPORTANCE OF THE MONONUCLEAR ENZYMES. *Journal of Biological Chemistry* **277**, 24142–24147 (2002).
17. Jin, W. *et al.* Comparative Study of the Inhibition of Metallo- β -Lactamases (IMP-1 and VIM-2) by Thiol Compounds That Contain a Hydrophobic Group. **27**, 6 (2004).
18. King, D. & Strynadka, N. Crystal structure of New Delhi metallo- β -lactamase reveals molecular basis for antibiotic resistance. *Protein Science* **20**, 1484–1491 (2011).
19. Baier, F. & Tokuriki, N. Connectivity between Catalytic Landscapes of the Metallo- β -Lactamase Superfamily. *Journal of Molecular Biology* **426**, 2442–2456 (2014).
20. Makena, A. *et al.* Comparison of Verona Integron-Borne Metallo- β -Lactamase (VIM) Variants Reveals Differences in Stability and Inhibition Profiles. *Antimicrob. Agents Chemother.* **60**, 1377–1384 (2015).
21. González, L. J. *et al.* Membrane anchoring stabilizes and favors secretion of New Delhi metallo- β -lactamase. *Nature Chemical Biology* **12**, 516–522 (2016).
22. González, M. M., Abriata, L. A., Tomatis, P. E. & Vila, A. J. Optimization of Conformational Dynamics in an Epistatic Evolutionary Trajectory. *Mol. Biol. Evol.* **33**, 1768–1776 (2016).
23. Socha, R. D., Chen, J. & Tokuriki, N. The Molecular Mechanisms Underlying Hidden Phenotypic Variation among Metallo- β -Lactamases. *Journal of Molecular Biology* **431**, 1172–1185 (2019).

24. Sun, Z., Hu, L., Sankaran, B., Prasad, B. V. V. & Palzkill, T. Differential active site requirements for NDM-1 β -lactamase hydrolysis of carbapenem versus penicillin and cephalosporin antibiotics. *Nat Commun* **9**, 1–14 (2018).
25. Fowler, D. M. *et al.* High-resolution mapping of protein sequence-function relationships. *Nature Methods* **7**, 741–746 (2010).
26. Fowler, D. M., Stephany, J. J. & Fields, S. Measuring the activity of protein variants on a large scale using deep mutational scanning. *Nat Protoc* **9**, 2267–2284 (2014).
27. Rubin, A. F. *et al.* A statistical framework for analyzing deep mutational scanning data. *Genome Biology* **18**, (2017).
28. Starita, L. M. *et al.* Massively Parallel Functional Analysis of BRCA1 RING Domain Variants. *Genetics* **200**, 413–422 (2015).
29. Weile, J. *et al.* A framework for exhaustively mapping functional missense variants. *Mol. Syst. Biol.* **13**, 957 (2017).
30. Matreyek, K. A. *et al.* Multiplex assessment of protein variant abundance by massively parallel sequencing. *Nat. Genet.* **50**, 874–882 (2018).
31. Starr, T. N., Picton, L. K. & Thornton, J. W. Alternative evolutionary histories in the sequence space of an ancient protein. *Nature* **549**, 409–413 (2017).
32. Stiffler, M. A., Hekstra, D. R. & Ranganathan, R. Evolvability as a Function of Purifying Selection in TEM-1 β -Lactamase. *Cell* **160**, 882–892 (2015).
33. Steinberg, B. & Ostermeier, M. Shifting Fitness and Epistatic Landscapes Reflect Trade-offs along an Evolutionary Pathway. *J. Mol. Biol.* **428**, 2730–2743 (2016).
34. Lee, J. M. *et al.* Deep mutational scanning of hemagglutinin helps predict evolutionary fates of human H3N2 influenza variants. *Proc Natl Acad Sci U S A* **115**, E8276–E8285 (2018).

35. Araya, C. L. *et al.* A fundamental protein property, thermodynamic stability, revealed solely from large-scale measurements of protein function. *Proceedings of the National Academy of Sciences* **109**, 16858–16863 (2012).
36. Firnberg, E., Labonte, J. W., Gray, J. J. & Ostermeier, M. A Comprehensive, High-Resolution Map of a Gene's Fitness Landscape. *Molecular Biology and Evolution* **31**, 1581–1592 (2014).
37. Thyagarajan, B. & Bloom, J. D. The inherent mutational tolerance and antigenic evolvability of influenza hemagglutinin. *eLife* **3**, (2014).
38. Klesmith, J. R., Bacik, J.-P., Wrenbeck, E. E., Michalczyk, R. & Whitehead, T. A. Trade-offs between enzyme fitness and solubility illuminated by deep mutational scanning. *Proceedings of the National Academy of Sciences* **114**, 2265–2270 (2017).
39. Roscoe, B. P., Thayer, K. M., Zeldovich, K. B., Fushman, D. & Bolon, D. N. A. Analyses of the Effects of All Ubiquitin Point Mutants on Yeast Growth Rate. *Journal of Molecular Biology* **425**, 1363–1377 (2013).
40. Romero, P. A., Tran, T. M. & Abate, A. R. Dissecting enzyme function with microfluidic-based deep mutational scanning. *Proc. Natl. Acad. Sci. U.S.A.* **112**, 7159–7164 (2015).
41. Gray, V. E., Hause, R. J. & Fowler, D. M. Analysis of Large-Scale Mutagenesis Data To Assess the Impact of Single Amino Acid Substitutions. *Genetics* **207**, 53–61 (2017).
42. Melnikov, A., Rogov, P., Wang, L., Gnirke, A. & Mikkelsen, T. S. Comprehensive mutational scanning of a kinase *in vivo* reveals substrate-dependent fitness landscapes. *Nucleic Acids Research* **42**, e112–e112 (2014).
43. Wrenbeck, E. E., Azouz, L. R. & Whitehead, T. A. Single-mutation fitness landscapes for an enzyme on multiple substrates reveal specificity is globally encoded. *Nature Communications* **8**, 15695 (2017).
44. Mavor, D. *et al.* Determination of ubiquitin fitness landscapes under different chemical stresses in a classroom setting. *Elife* **5**, (2016).

45. Noda-García, L. *et al.* Chance and pleiotropy dominate genetic diversity in complex bacterial environments. *Nat Microbiol* **4**, 1221–1230 (2019).
46. Thompson, S., Zhang, Y., Ingle, C., Reynolds, K. A. & Kortemme, T. Modulating the cellular context broadly reshapes the mutational landscape of a model enzyme. *bioRxiv* 848010 (2019) doi:10.1101/848010.
47. van den Ent, F. & Löwe, J. RF cloning: A restriction-free method for inserting target genes into plasmids. *Journal of Biochemical and Biophysical Methods* **67**, 67–74 (2006).
48. Jacquier, H. *et al.* Capturing the mutational landscape of the beta-lactamase TEM-1. *PNAS* **110**, 13067–13072 (2013).
49. Sebaugh, J. L. & McCray, P. D. Defining the linear portion of a sigmoid-shaped curve: bend points. *Pharmaceutical Statistics* **2**, 167–174 (2003).
50. Kitzman, J. O., Starita, L. M., Lo, R. S., Fields, S. & Shendure, J. Massively parallel single-amino-acid mutagenesis. *Nature Methods* **12**, 203–206 (2015).
51. Gray, V. E., Hause, R. J., Luebeck, J., Shendure, J. & Fowler, D. M. Quantitative Missense Variant Effect Prediction Using Large-Scale Mutagenesis Data. *Cell Syst* **6**, 116-124.e3 (2018).
52. Garcia-Saez, I., Docquier, J.-D., Rossolini, G. M. & Dideberg, O. The Three-Dimensional Structure of VIM-2, a Zn-β-Lactamase from *Pseudomonas aeruginosa* in Its Reduced and Oxidised Form. *Journal of Molecular Biology* **375**, 604–611 (2008).
53. Oliver, D. Protein Secretion in *Escherichia Coli*. *Annual Review of Microbiology* **39**, 615–648 (1985).
54. Pugsley, A. P. The Complete General Secretory Pathway in Gram-Negative Bacteria. *MICROBIOL. REV.* **57**, 59 (1993).
55. Freudl, R. Signal peptides for recombinant protein secretion in bacterial expression systems. *Microbial Cell Factories* **17**, 52 (2018).

56. Paetzel, M. Bacterial Signal Peptidases. in *Bacterial Cell Walls and Membranes* (ed. Kuhn, A.) 187–219 (Springer International Publishing, 2019). doi:10.1007/978-3-030-18768-2_7.
57. Inouye, S. *et al.* Role of positive charge on the amino-terminal region of the signal peptide in protein secretion across the membrane. *PNAS* **79**, 3438–3441 (1982).
58. Iino, T., Takahashi, M. & Sako, T. Role of amino-terminal positive charge on signal peptide in staphylokinase export across the cytoplasmic membrane of *Escherichia coli*. *J. Biol. Chem.* **262**, 7412–7417 (1987).
59. Bhattacharyya, S. *et al.* Accessibility of the Shine-Dalgarno Sequence Dictates N-Terminal Codon Bias in *E. coli*. *Mol. Cell* **70**, 894-905.e5 (2018).
60. Bentele, K., Saffert, P., Rauscher, R., Ignatova, Z. & Blüthgen, N. Efficient translation initiation dictates codon usage at gene start. *Molecular Systems Biology* **9**, 675 (2013).
61. Kelsic, E. D. *et al.* RNA Structural Determinants of Optimal Codons Revealed by MAGE-Seq. *Cell Syst* **3**, 563-571.e6 (2016).
62. Jia, B. *et al.* CARD 2017: expansion and model-centric curation of the comprehensive antibiotic resistance database. *Nucleic Acids Research* **45**, D566–D573 (2017).
63. Brockmeier, U. *et al.* Systematic Screening of All Signal Peptides from *Bacillus subtilis*: A Powerful Strategy in Optimizing Heterologous Protein Secretion in Gram-positive Bacteria. *Journal of Molecular Biology* **362**, 393–402 (2006).
64. Mathiesen, G. *et al.* Genome-wide analysis of signal peptide functionality in *Lactobacillus plantarum* WCFS1. *BMC Genomics* **10**, 425 (2009).
65. Hemmerich, J. *et al.* Use of a Sec signal peptide library from *Bacillus subtilis* for the optimization of cutinase secretion in *Corynebacterium glutamicum*. *Microbial Cell Factories* **15**, 208 (2016).
66. Flynn, J. M. *et al.* Investigating the influence of environment on the evolution of Hsp90 using comprehensive fitness maps. *bioRxiv* 823468 (2019) doi:10.1101/823468.

67. Dellus-Gur, E., Toth-Petroczy, A., Elias, M. & Tawfik, D. S. What Makes a Protein Fold Amenable to Functional Innovation? Fold Polarity and Stability Trade-offs. *Journal of Molecular Biology* **425**, 2609–2621 (2013).
68. Moali, C. *et al.* Analysis of the Importance of the Metallo- β -Lactamase Active Site Loop in Substrate Binding and Catalysis. *Chemistry & Biology* **10**, 319–329 (2003).
69. Merino, M. *et al.* Role of changes in the L3 loop of the active site in the evolution of enzymatic activity of VIM-type metallo- β -lactamases. *Journal of Antimicrobial Chemotherapy* **65**, 1950–1954 (2010).
70. Leiros, H.-K. S. *et al.* His224 Alters the R2 Drug Binding Site and Phe218 Influences the Catalytic Efficiency of the Metallo- β -Lactamase VIM-7. *Antimicrobial Agents and Chemotherapy* **58**, 4826–4836 (2014).
71. Leiros, H.-K. S., Edvardsen, K. S. W., Bjerga, G. E. K. & Samuelsen, Ø. Structural and biochemical characterization of VIM-26 shows that Leu224 has implications for the substrate specificity of VIM metallo- β -lactamases. *The FEBS Journal* **282**, 1031–1042 (2015).
72. Maurer, M. & Oostenbrink, C. Water in protein hydration and ligand recognition. *Journal of Molecular Recognition* **32**, e2810 (2019).
73. Spyarakis, F. *et al.* The Roles of Water in the Protein Matrix: A Largely Untapped Resource for Drug Discovery. *J. Med. Chem.* **60**, 6781–6827 (2017).
74. Petrović, D., Risso, V. A., Kamerlin, S. C. L. & Sanchez-Ruiz, J. M. Conformational dynamics and enzyme evolution. *J R Soc Interface* **15**, (2018).
75. Campbell, E. C. *et al.* Laboratory evolution of protein conformational dynamics. *Current Opinion in Structural Biology* **50**, 49–57 (2018).
76. Campbell, E. *et al.* The role of protein dynamics in the evolution of new enzyme function. *Nat. Chem. Biol.* **12**, 944–950 (2016).

77. Singh, P., Abeyasinghe, T. & Kohen, A. Linking protein motion to enzyme catalysis. *Molecules* **20**, 1192–1209 (2015).
78. Brown, N. G., Pennington, J. M., Huang, W., Ayvaz, T. & Palzkill, T. Multiple global suppressors of protein stability defects facilitate the evolution of extended-spectrum TEM β -lactamases. *J. Mol. Biol.* **404**, 832–846 (2010).
79. Huang, W. & Palzkill, T. A natural polymorphism in beta-lactamase is a global suppressor. *Proc. Natl. Acad. Sci. U.S.A.* **94**, 8801–8806 (1997).
80. Starr, T. N. & Thornton, J. W. Epistasis in protein evolution. *Protein Sci.* **25**, 1204–1218 (2016).
81. Miton, C. M. & Tokuriki, N. How mutational epistasis impairs predictability in protein evolution and design. *Protein Sci.* **25**, 1260–1272 (2016).
82. Breen, M. S., Kemena, C., Vlasov, P. K., Notredame, C. & Kondrashov, F. A. Epistasis as the primary factor in molecular evolution. *Nature* **490**, 535–538 (2012).
83. Sarkisyan, K. S. *et al.* Local fitness landscape of the green fluorescent protein. *Nature* **533**, 397–401 (2016).
84. Pokusaeva, V. O. *et al.* An experimental assay of the interactions of amino acids from orthologous sequences shaping a complex fitness landscape. *PLOS Genetics* **15**, e1008079 (2019).
85. Meer, J.-Y. van der *et al.* Using mutability landscapes of a promiscuous tautomerase to guide the engineering of enantioselective Michaelases. *Nat Commun* **7**, 1–16 (2016).
86. Firnberg, E. & Ostermeier, M. PFunkel: efficient, expansive, user-defined mutagenesis. *PLoS ONE* **7**, e52031 (2012).
87. Reetz, M. T. & Carballeira, J. D. Iterative saturation mutagenesis (ISM) for rapid directed evolution of functional enzymes. *Nat Protoc* **2**, 891–903 (2007).
88. Firth, A. E. & Patrick, W. M. GLUE-IT and PEDEL-AA: new programmes for analyzing protein diversity in randomized libraries. *Nucleic Acids Research* **36**, W281–W285 (2008).

89. Edgar, R. C. & Flyvbjerg, H. Error filtering, pair assembly and error correction for next-generation sequencing reads. *Bioinformatics* **31**, 3476–3482 (2015).
90. Ahmad, S., Gromiha, M., Fawareh, H. & Sarai, A. ASAView: Database and tool for solvent accessibility representation in proteins. *BMC Bioinformatics* **5** (2004).
91. Perkins, S. J. Protein volumes and hydration effects. The calculations of partial specific volumes, neutron scattering matchpoints and 280-nm absorption coefficients for proteins and glycoproteins from amino acid sequences. *European Journal of Biochemistry* **157**, 169–180 (1986).
92. Kyte, J. & Doolittle, R. F. A simple method for displaying the hydropathic character of a protein. *Journal of Molecular Biology* **157**, 105–132 (1982).
93. Groemping, U. Relative Importance for Linear Regression in R: The Package relaimpo. *Journal of Statistical Software* **17**, 1–27 (2006).
94. Kellogg, E. H., Leaver-Fay, A. & Baker, D. Role of conformational sampling in computing mutation-induced changes in protein structure and stability. *Proteins: Structure, Function, and Bioinformatics* **79**, 830–838 (2011).
95. Zadeh, J. N. *et al.* NUPACK: Analysis and design of nucleic acid systems. *Journal of Computational Chemistry* **32**, 170–173 (2011).
96. NCBI Resource Coordinators. Database resources of the National Center for Biotechnology Information. *Nucleic Acids Res.* **46**, D8–D13 (2018).
97. Kumar, S., Stecher, G. & Tamura, K. MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for Bigger Datasets. *Mol. Biol. Evol.* **33**, 1870–1874 (2016).

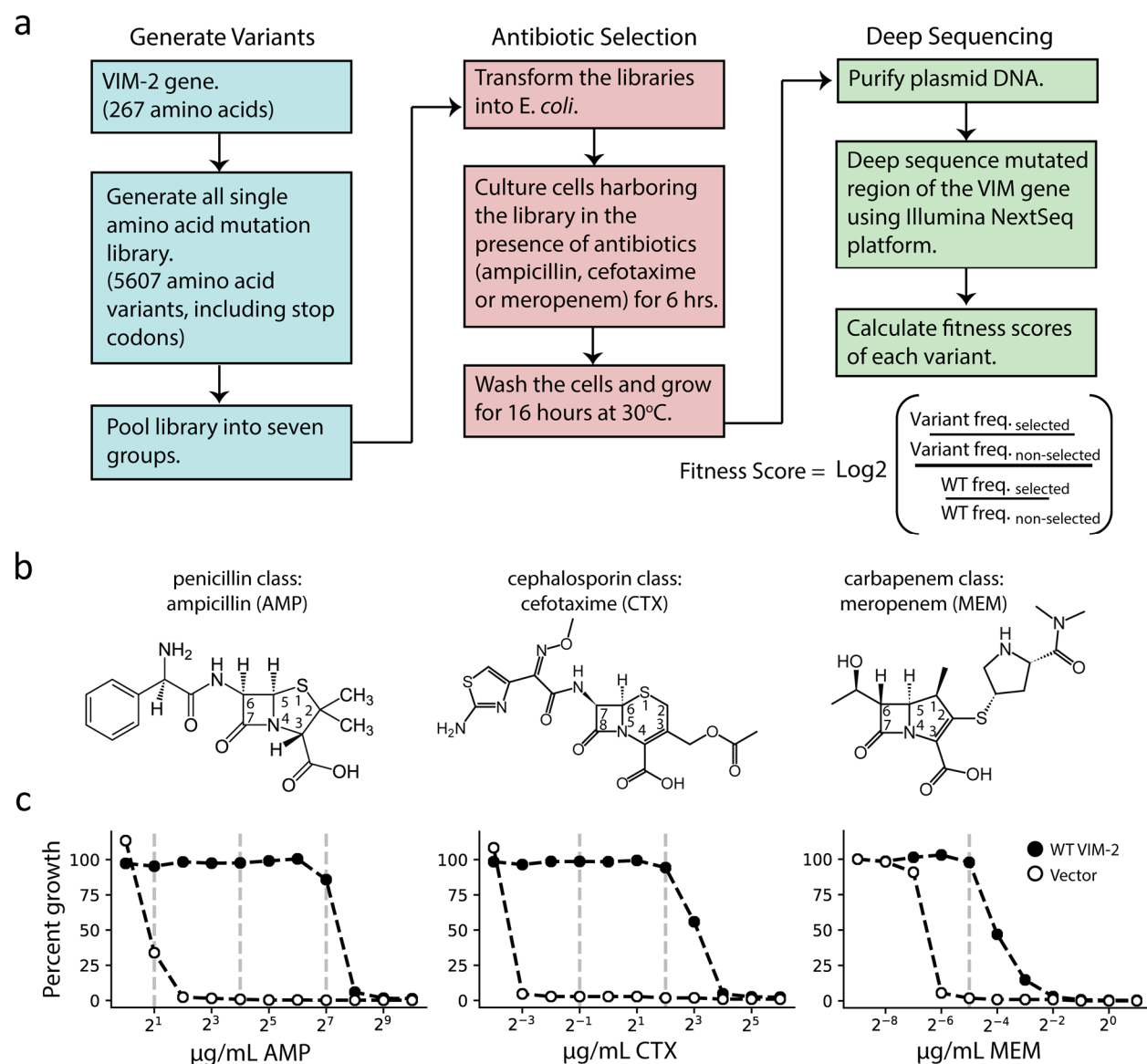


Figure 1. Deep mutational scanning (DMS) overview. (a) The workflow for DMS. All single amino acid variants are first generated using RF cloning, subsequently transformed into *E. coli* and then subject to selection for antibiotic resistance conferred to *E. coli*. The effects of selection (fitness score) were evaluated by deep sequencing and comparing the enrichment of each variant with and without selection. (b) Chemical structures of the antibiotics used in this study. (c) The dose-response growth curve of *E. coli* transformed with wtVIM-2, and an empty vector control for each antibiotic. The vertical dashed lines indicate antibiotic concentrations at which selections were performed in this study.

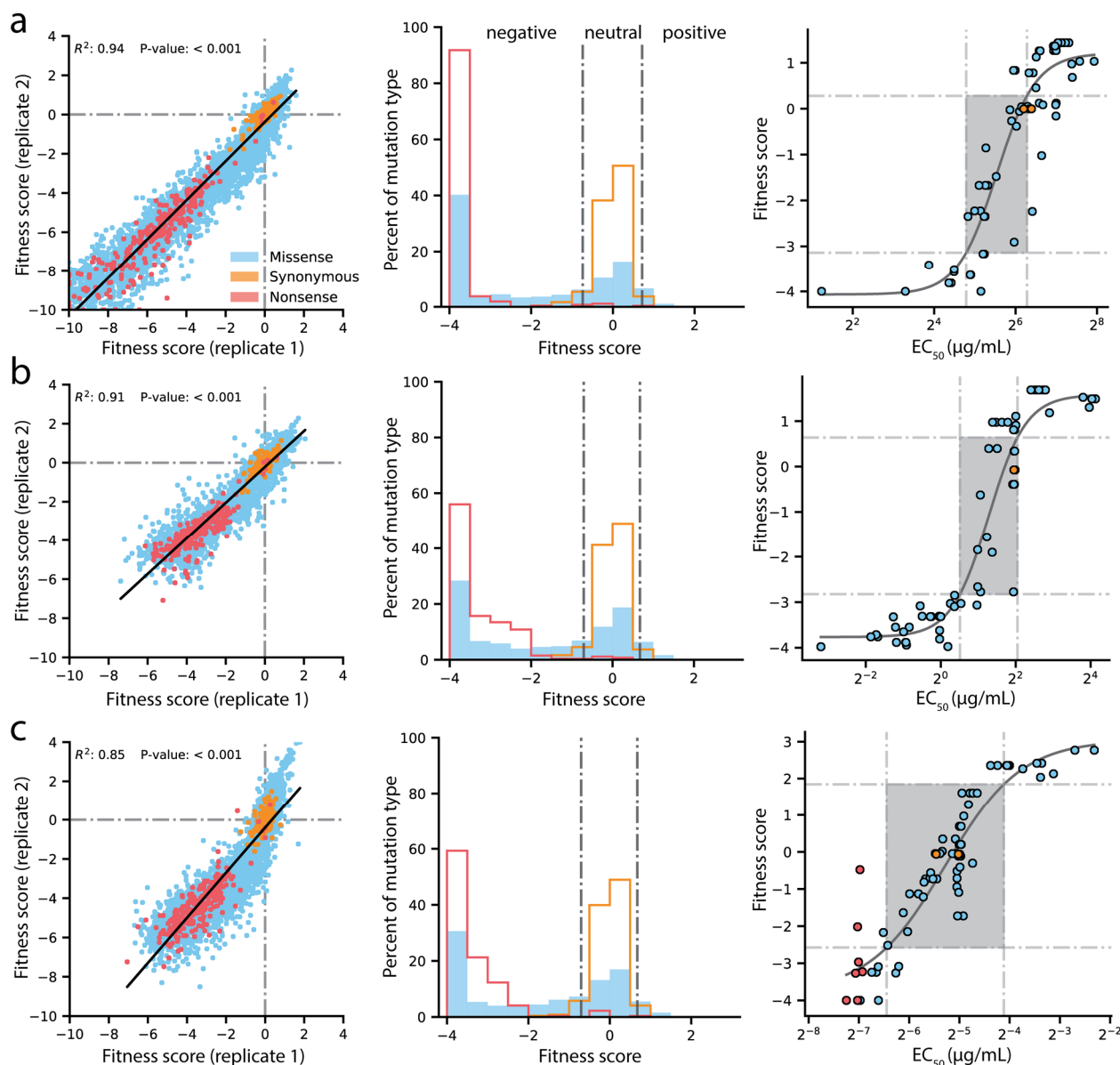


Figure 2. Quality control of DMS data and general mutational properties of VIM-2. In the horizontal panels, data is shown for (a) 128 $\mu\text{g/mL}$ AMP selection, (b) 4.0 $\mu\text{g/mL}$ CTX selection and (c) 0.031 $\mu\text{g/mL}$ MEM selection. For each panel, the left plot shows correlation between fitness scores of all variants in the two replicates of DMS. The middle plot shows distribution of fitness effects for all variants separated into synonymous, missense and nonsense distributions, where the vertical grey lines indicates fitness score cut-offs used to classify fitness effects as positive, neutral or negative. The right plot shows the relationship of DMS fitness scores with antibiotic resistance (EC_{50}) of isolated variants measured in a dose-response curve; variants with resistance lower than the tested range could not be fitted for EC_{50} , leading to 69 individually measured EC_{50} values for AMP, 67 for CTX and 75 for MEM—some points are the same codon or amino acid variant isolated multiple times. The filled rectangle in the background indicates the region of linear association between fitness scores and EC_{50} . The legend in panel a) is shared by all panels.

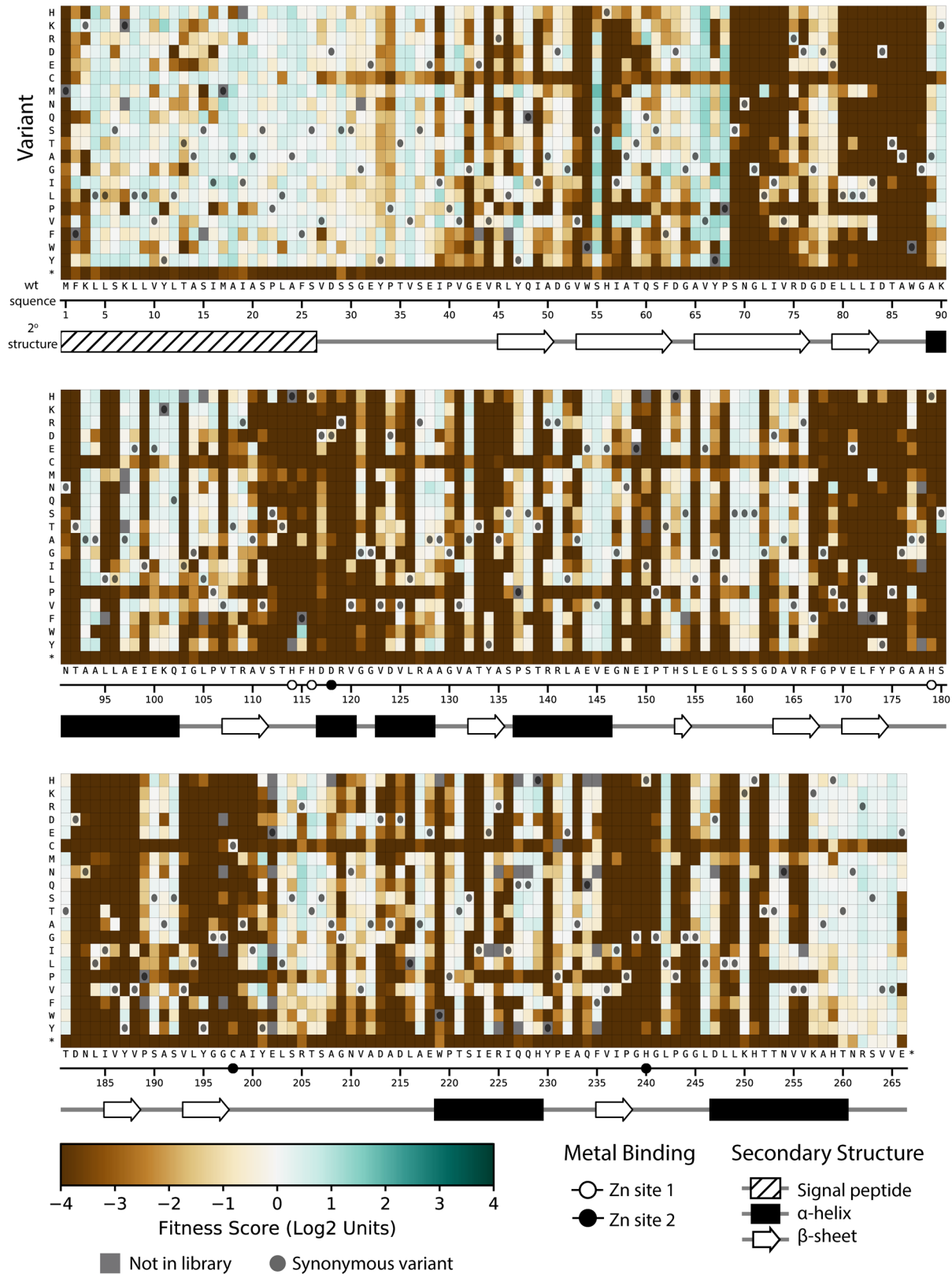


Figure 3. Fitness of all VIM-2 single amino acid variants under 128 µg/mL AMP selection. Each cell in the heat map represents the fitness score of a single amino acid variant. Synonymous variants are indicated by black circles and variants that are not present in the library are in grey. The x-axis under the heatmap indicates the wt residue and position (the 6 active site metal binding residues are highlighted as circles), while the y-axis indicates the variant residue at that position. The secondary structure of the wtVIM-2 crystal structure (PDB: 4bz3) is displayed below the heatmap.

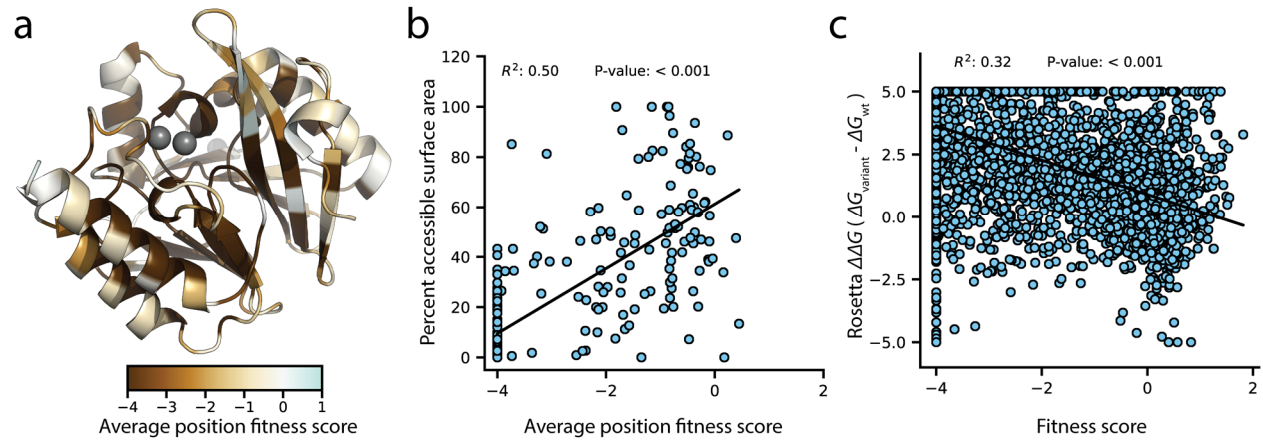


Figure 4. Correlation of fitness with structural attributes. Fitness scores are from DMS at 128 $\mu\text{g/mL}$ AMP selection. (a) Crystal structure of wtVIM-2 (PDB: 4bz3) colored by the average fitness of 20 amino acid mutations at each position. (b) The correlation between accessible surface area and the average fitness of 20 amino acid mutations at each position. (c) The correlation between the changes in folding energy predicted by Rosetta and the DMS fitness scores for all variants.

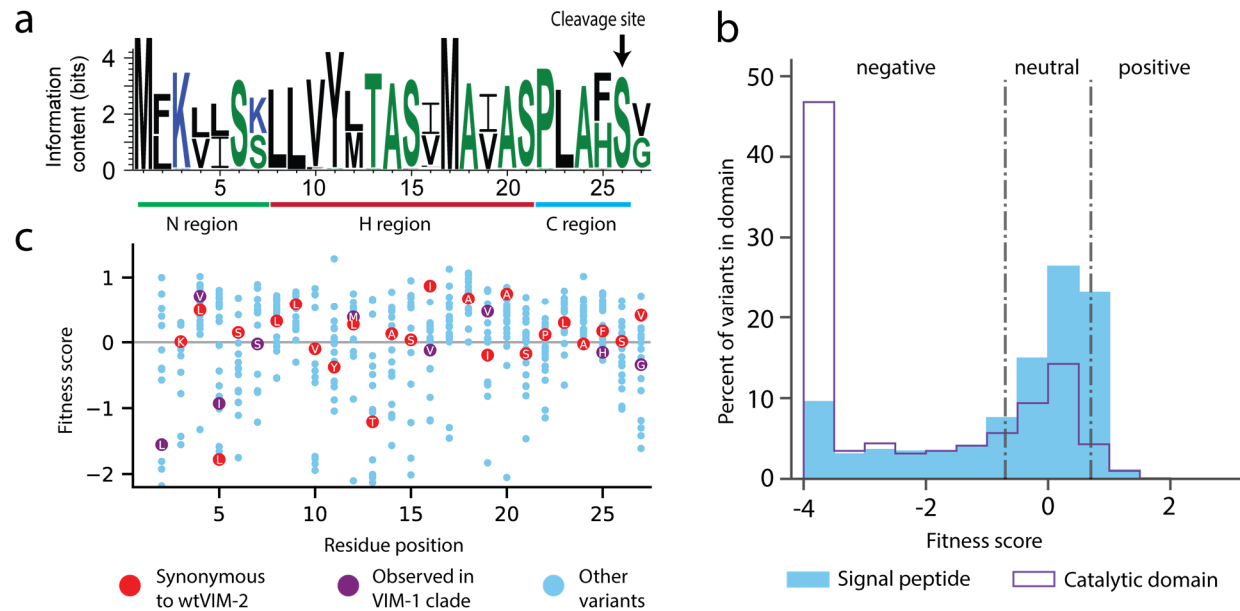


Figure 5. Conservation patterns and fitness scores in the signal peptide. (a) Sequence logo of the signal peptide region aligned across all VIM natural variants generated using WebLogo (<https://weblogo.berkeley.edu/>). Positions with two major naturally occurring residues are conserved differences between the VIM-1 and VIM-2 clades (clades are defined in **Supplementary Fig. 6**). (b) The distribution of fitness effects of all DMS variants, separated into signal peptide variants and catalytic domain variants. (c) DMS fitness scores of all variants at each position of the signal peptide. Synonymous variants of wtVIM-2 and conserved variants observed in the VIM-1 clade are highlighted as labelled circles.

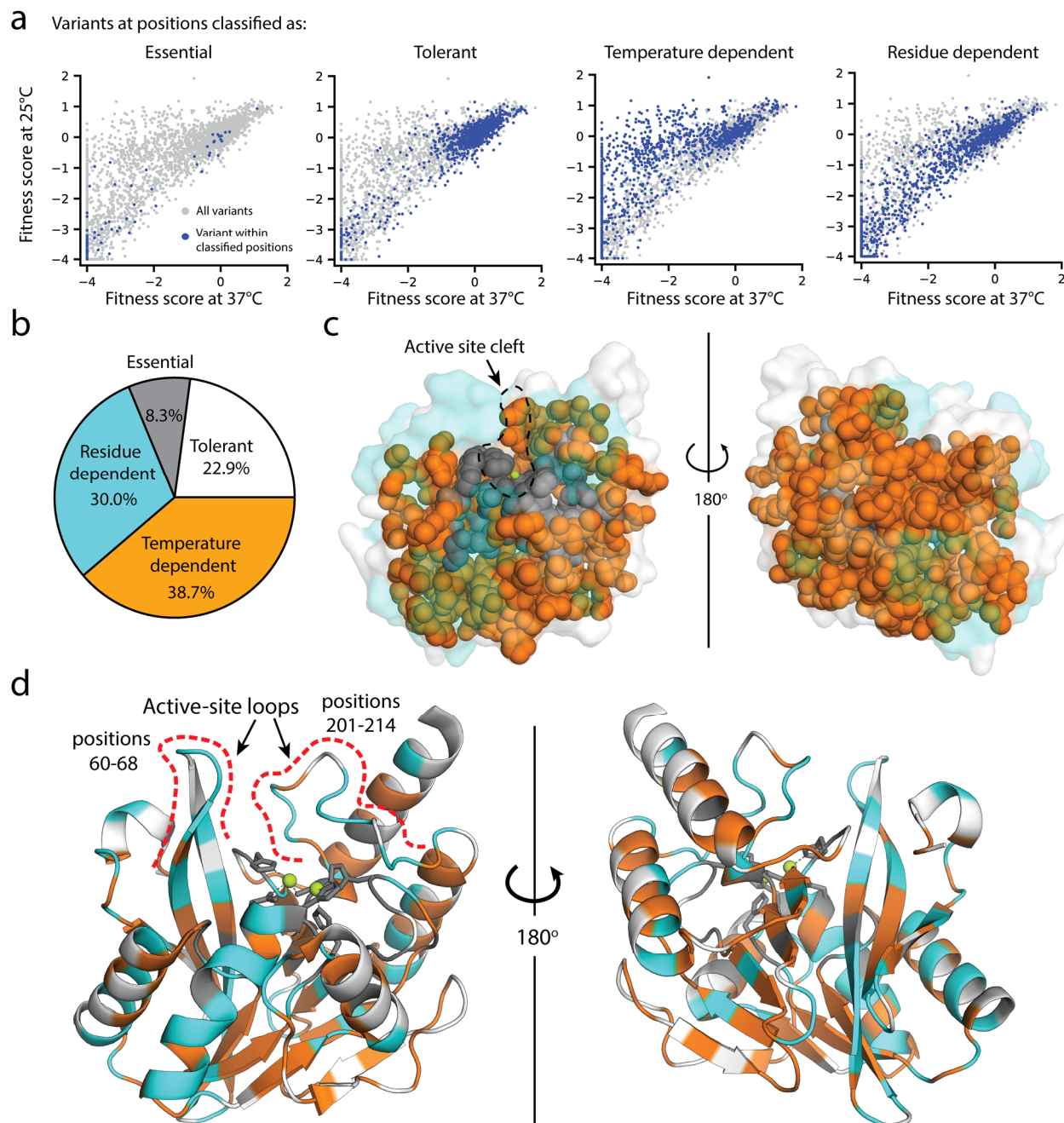


Figure 6. Distribution of mutational tolerance and temperature dependence of wtVIM-2 residues. (a) Scatterplots comparing fitness scores under selection at 25°C and 37°C (128 μ g/mL AMP). Variants within the classified positions are highlighted in dark blue while all variants are plotted in grey for reference. (b) Proportion of residues in the wtVIM-2 catalytic domain that have been classified into each behavioral category. (c-d) The wtVIM-2 crystal structure (PDB: 5yd7) is colored by the behavioral classifications, the active-site zinc ions are colored in green. (c) View of the inner core of wtVIM-2, with essential and temperature dependent residues depicted as spheres. (d) Cartoon representation of wtVIM-2 with metal-binding residues shown as sticks.

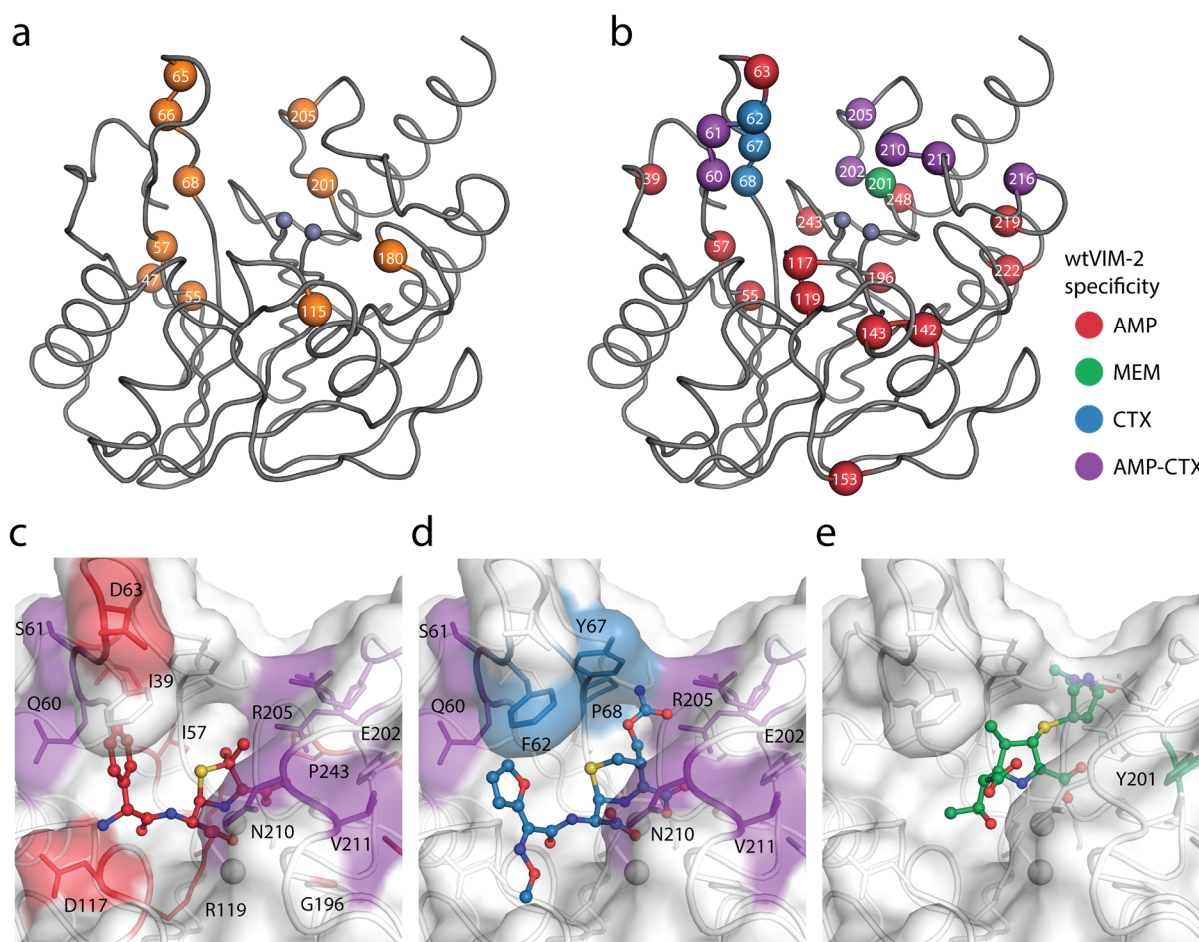


Figure 7. Visualization of VIM-2 specificity determining positions. The wtVIM-2 crystal structure (PDB: 5yd7) is featured in all panels, with the active site Zn ions depicted as grey spheres. **(a)** Positions with at least one globally adaptive mutation are highlighted as orange spheres. **(b)** Positions classified as being responsible for specificity towards certain antibiotics in wtVIM-2 are color coded by antibiotic and highlighted as spheres. **(c-e)**. Close-up views of the specificity residues in the active site with **(c)** hydrolyzed ampicillin (PDB: 4hl2), **(d)** cefuroxime (PDB: 4rl0) and **(e)** meropenem (PDB: 5n5i) from VIM-1 and NDM-1 structures that have been aligned to the wtVIM-2 structure using the active site residues. Residues are colored by the inferred substrate specificity as in **(b)**. Substrates are shown in stick and ball representation.

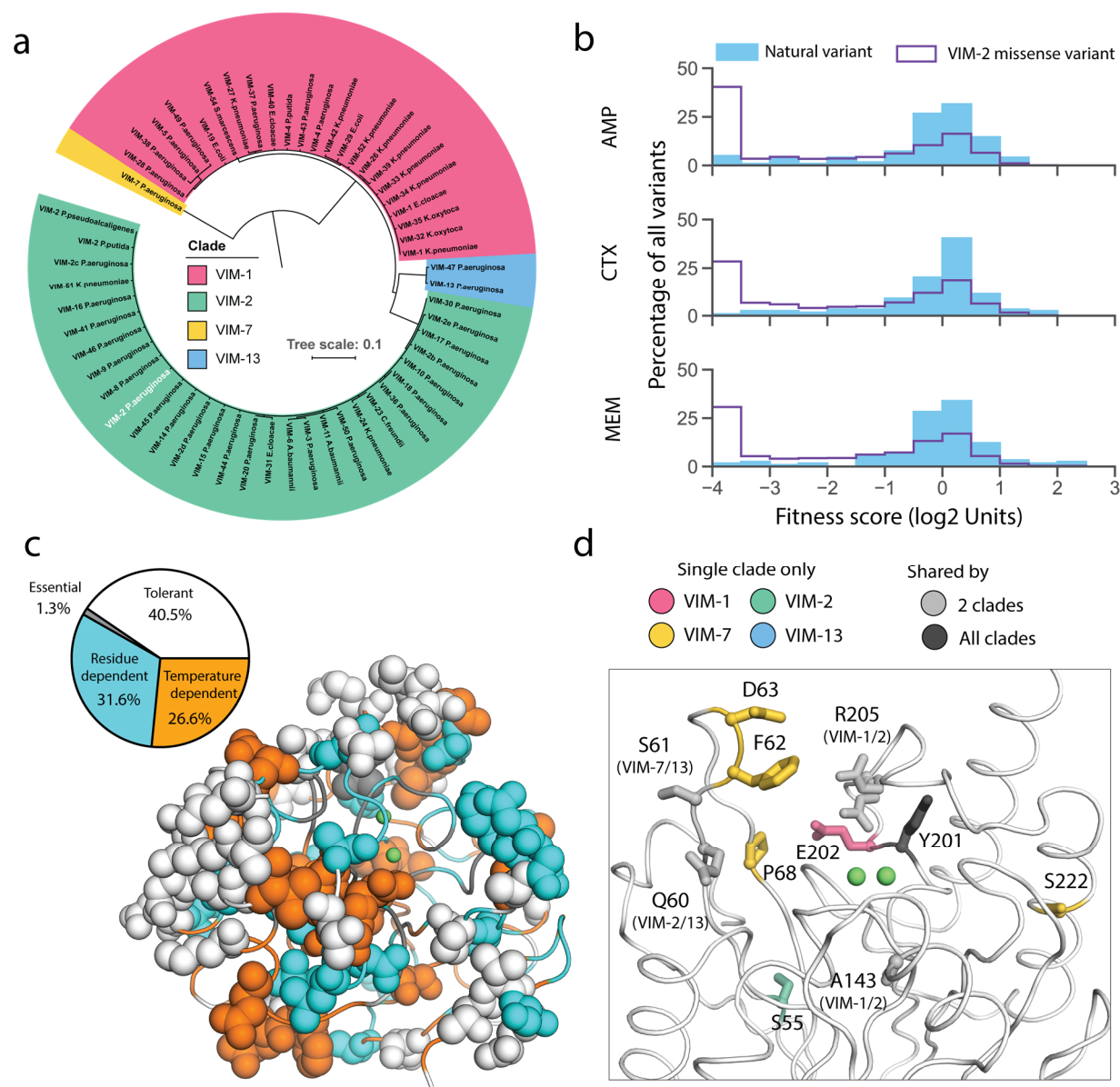


Figure 8. Behavior of natural VIM variants inferred from DMS fitness. (a) Maximum likelihood phylogenetic tree of all natural VIM variants examined in this study, colored by major clades (a larger version of the tree is presented in **Supplementary Fig. 6**). (b) Distribution of fitness for all unique individual mutations found in VIM natural variants compared to all missense variants measured in DMS for all 3 antibiotics. (c) All residues mutated in the natural variants are shown in sphere representation and colored by mutational tolerance and temperature dependence. The pie chart on the right shows the proportion of natural variant positions in each classification. (d) wtVIM-2 residues that are both mutated in at least one natural variant and affect specificity are highlighted as sticks, colored by the clade(s) in which the residue is mutated.

Table 1. Linear model output for DMS fitness scores under 128 µg/mL AMP selection

Predictor ^a		Estimated effect ^b	Adjusted P-value ($\alpha=0.05$) ^c	Variance explained ^d
(Intercept) ^e		-1.73	< 0.001	
ASA ^f		2.83	< 0.001	21%
Rosetta $\Delta\Delta G$ ^g		-0.28	< 0.001	18%
Starting(wt) residue	C	-1.53	< 0.001	10%
	D	-1.33	< 0.001	
	E	-0.53	< 0.001	
	G	-0.92	< 0.001	
	H	-1.09	< 0.001	
	I	-0.50	< 0.001	
	L	-0.43	< 0.001	
	N	-0.62	< 0.001	
	R	-0.35	0.001	
	S	-0.23	0.018	
	V	-0.36	< 0.001	
	W	-1.35	< 0.001	
Variant residue	C	-1.69	< 0.001	5%
	D	-0.40	0.002	
	P	-0.61	< 0.001	
	W	-0.45	< 0.001	

a) Each predictor indicates a class of wtVIM-2 derived values that were used as explanatory variables to model a linear relationship with the observed fitness score.

b) The estimated effect is the predicted change in fitness score away from the intercept with a 1 unit increase in a continuous predictor or a binary change in a categorical predictor.

c) P-values indicates whether a predictor makes a significant contribution to the change fitness score, and are adjusted for a false discovery rate of 5% using the Benjamini-Hochberg procedure.

d) The adjusted R^2 of each predictor when correlated with fitness, which is a measure of how much variation in the fitness score can be explained by each predictor in the linear model.

e) The intercept is the average fitness of all variants where the continuous variable is 0 (ASA and Rosetta $\Delta\Delta G$) and the wt or variant residue is Ala.

f) ASA ranges from 0.0 to 1.0.

g) Rosetta $\Delta\Delta G$ ranges from -5.0 to 5.0 Rosetta energy units