Supporting Information for :

Capturing the mutational landscape of the beta-lactamase TEM-1

Hervé Jacquier*, André Birgy, Hervé Le Nagard, Yves Mechulam, Emmanuelle Schmitt, Jérémy Glodt, Beatrice Bercot, Emmanuelle Petit, Julie Poulain, Guilène Barnaud, Pierre-Alexis Gros, Olivier Tenaillon*

*To whom correspondence should be addressed. E-mail: olivier.tenaillon@inserm.fr or herve.jacquier@lrb.aphp.fr

Supporting Methods

Plasmids construction

Template plasmid pUCHJ. The TEM-1 β-lactamase gene was cloned into a pUC19 plasmid (New England Biolabs, Ipswich, USA) modified by introducing new restriction sites (*Nco*I and *Not*I,) flanking the start and stop codons of TEM-1's open reading frame.

Cloning plasmid pBRHJ. The mutagenized amplicons were cloned into the previous modified plasmid containing the pMB1 origin of replication from pBR322 and gentamicin resistance gene (*aac*C4) at the *Xba*I site.

Mutagenesis. The mutagenesis was performed using GeneMorph® II Random Mutagenesis Kit (Stratagene, La Jolla, USA) to obtain an average of one mutation per gene, according to the manufacturer's instructions: template plasmid pUCHJ (2.4 µg) and the primers TEMF1 and TEMR1 (Table S1) were used for 25 PCR cycles. The resulting PCR products were digested with *Nco*I and *Not*I (New England Biolabs) and subcloned into the cloning vector pBRHJ. The ligation products were transformed into ElectroMax DH10B-T1 Phage Resistant *Escherichia coli* Competent Cells (Invitrogen, Fisher Scientific, France) and plated on Luria–Bertani agar supplemented with gentamicin (20 mg/L). The number of mutations per clone followed a Poisson-like distribution, with an average of 1.93. However, a higher variance (2.4) was observed, as expected in the cases of random PCR mutagenesis (1). The mutational bias was compatible with the one described in the mutagenesis kit.
The single mutants library will be available upon request.

Sequencing. A total of 10,368 randomly picked clones were stored into 384-well microplates (Molecular Devices, Sunnyvale, USA). Plasmid DNA was amplified with the ϕ29 DNA

polymerase by rolling circle amplification (GE Healthcare, UK). Sequencing was done on a ABI sequencer with BigDye terminator cycle sequencing kit (Applied Biosystems, Foster City, USA) using primers TMF and TMS.

We developed a software using Phred and Phrap for automated mutations detections. Effective quality control was performed immediately after sequence production: sequences with coverage <1X and low quality sequences were excluded. All sequences (ca 10% of the library) with barely quality information were controlled using Staden software (http://staden.sourceforge.net/)

Minimum Inhibitory Concentration (MIC) measurements. The library was grown 18 hours at 37°C on 384-well microplates (Abgene, Epsom, United Kingdom) containing 50 µl LB broth supplemented with gentamicin (10 mg/L). Each clone was plated on MH agar plates containing a growing concentration of amoxicillin (0, 12.5, 25, 50, 100, 250, 500, 1000, 2000 and 4000 mg/L) using a polypropylene pin sterile replicator (Scinomix, Earth City, USA). After 18 hours of incubation at 37°C, the MIC was defined as the first concentration of amoxicillin inhibiting the growth of bacteria. Each plate was pictured and we built an image-analysis platform to detect the growth of each spot and to measure the size of each bacterial spot (Fig. S1). A threshold size was used to eliminate the presence of a single colony due to some mutants emerging on the plate. These data were subsequently analyzed with Microsoft Excel to determine MIC of each clone. These experiments were performed in triplicate.

We estimated that we have two sources of noise in MIC measurements. The first one is due to the limitations of MIC estimation by agar dilution. Indeed, it is known that results obtained with MIC by agar dilution could vary in an order of one dilution. The second source of noise relies on the existence of mutations outside of the sequenced region that may occur in the cloning process, such as mutations affecting plasmid copy number or the promoter of TEM-1. Using the wild-type MICs replicates, we estimated that about 4% of clones had a MIC shifted due to this factor, resulting in MIC measurement changes of about 4% (on a log2 scale).

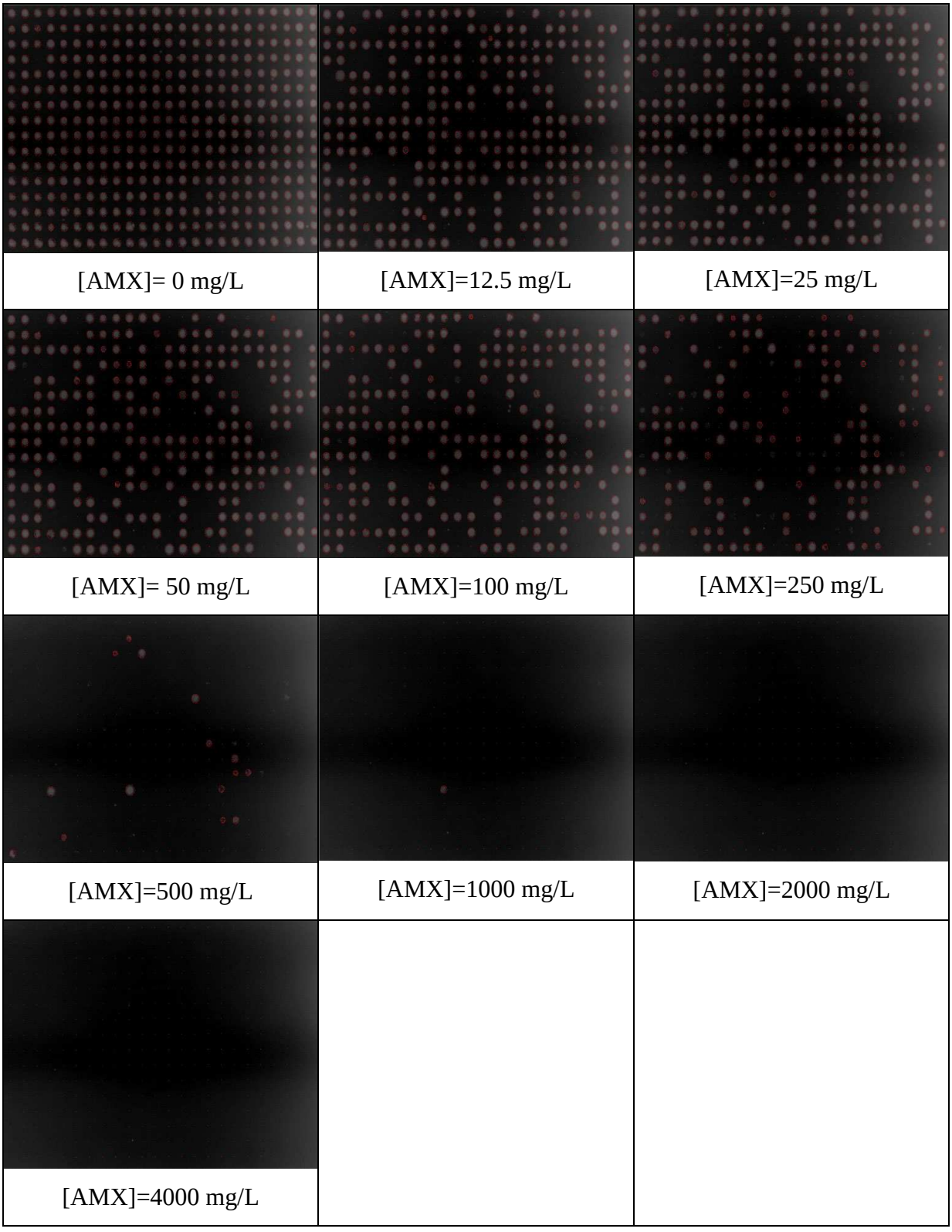


Fig. S1 : Example of data used to infer MIC from pictures. White spot corresponds to the bacterial growth and the red circle to the detection of the spot by our image-analysis platform.

Table S1: Primers for mutagenic PCR and sequencing

TEMF1	5'- CTTTTCTACGGGGTCTGACG-3'
TEMR1	5'- TTGAATTCGCCTCGTGATACGCCTATTT-3'
TMF	5'- TTATCAAAAAGGATCTTCACCTAG-3'
TMS	5'- TACATTCAAATATGTATCCGCTC-3'

Growth curves. All mutants harbouring a single non-synonymous mutation were transferred with a robot (EpMotion, Eppendorf) in ten 96-well plates filled with LB + 40% glycerol and stored at -80°C. Each plate included 7 blanks and 7 random clones having no mutations. From glycerol plates, using a 96 pins polypropylene replicator, a 96 deep-well plate was inoculated with 200µl of MH medium with gentamicin (1 mg/L) and a plastic bead in each well. Oxygen permeable sealing was used and the 18h incubation at 37°C was performed in an incubator assuring 75% humidity. These plates were then diluted with pin replicators twice (about >10,000 fold dilution) into fresh MH broth with either 6 or 100 mg/L of amoxicillin. 50µl of sterile mineral oil was added to each well. Growth curves were then produced with a Tecan infinite 96-well plate reader at an optical density of 600.

Using R, the growth curves were smoothed, the minimum value of each curve was subtracted and a constant set to 0.01, to 0.05 was added. The addition of the constant guarantees that the maximum growth rate is estimated at optical densities that is far enough from the measurement noise of the blanks. All results were validated with different values of this constant and were similar, but the 0.05 value gave smaller coefficient of variation among growth rates of non-mutated clones, so all results presented use that value. The transformed values were further log transformed and the maximum value of the derivative of the logarithm was used as the Maximum Growth Rate (MGR). Using the no-mutation clones, a plate effect was deduced and all values were corrected by that effect.

MGR per hour smaller than 0.18 were counted as 0, and MGR that occurred after 35,000 seconds were counted as growth of mutants and were also counted as 0.

Growth Curves and MIC. MIC is an integrated measure with many components. It measures the efficiency of TEM-1, its effective concentration, but also the costs associated with its production, its potential aggregation effects on growth… In addition, MIC may measure the evolvability of the enzyme, as in the growth process in the presence of the antibiotic some resistant mutant may be selected for. To limit this effect that may affect our understanding of

mutation effect, we choose to measure MIC on plates. On plates, the growth of few mutants is made clear by the presence of independent colonies rather than the full growth of a spot. Our software included a threshold size of detection that eliminated some of these mutants growth. In liquid, if the OD is measured only after 18h, a single mutant can take over and it is not possible to see the contribution of mutants to the MIC. However, if the full growth curve is analyzed, we can identify late growth and discard it. We can use these data to compare MIC on plate and MIC in liquid and find discrepancies. Using a low concentration of amoxicillin, we found that 11 out of 757 mutants had an MIC greater or equal to 25 mg/L and could not grow at a concentration of 6 mg/L in liquid. 9 of these showed delayed growth. So the error is about 1.4%. When we looked at a higher concentration of amoxicillin (100 mg/L), we found that 12% of mutants that were supposed to grow at that concentration according to their MIC (MIC>500 mg/L) had delayed growth which suggested that the MIC might have been overestimated by 1 unit in about 10% of cases. We therefore performed all analysis excluding the mutants showing delayed growth on one or the other concentration and recovered results qualitatively similar to those reported in the main text. Globally, the omission of the late growth mutants improved by 1 to 2 % the variance explained by the different models. We therefore think that the evolvability component of MIC is a negligible factor in our analysis.

Enzyme assays on cell extracts. TEM-1 and its variants were grown overnight in 96-deep-well plates containing MH supplemented with gentamicin, as described above. A volume of 45 µL of culture was added to 155 µL of lysis buffer inside microtiter plates, to obtain a lysate containing phosphate buffer (K_2HPO_4 50 mM, EDTA 1 mM,pH 7.8), Cell Culture Lysis Reagent 1X (Luciferase Assay System®, Promega, USA) supplemented with 1.5 mg/mL lysozyme and 3 mg/mL bovine serum albumin. Plates were then incubated at 22°C for 30 minutes. Lysates were diluted 1:100 in microtiter plates containing 190 µL of 100 mM potassium phosphate buffer (pH 7.25) supplemented with chromogenic beta-lactam nitrocefin (50 µg/mL). $O.D_{486}$ was measured for 3 h at 3min intervals (25 °C) using a Tecan infinite 96-well plate reader. Initial velocities (V_0) were determined at linear phases of nitrocefin hydrolysis using R. Concomitantly; overnight cultures were diluted 1:3 to measure OD_{610} to normalize V_0 by biomass.

Correlation between estimated initial velocity V_0 and MIC. We estimated the initial velocity on nitrocefine substrate for 474 single mutants having various MIC. Initial velocity was estimated as the slope of the changes in OD through time at 486 nm. The assay was

performed in duplicate in 96 well plates and 9 controls without mutations were present on each plate to measure a plate effect. We also measured the OD_{610} of each culture before lysis and potentially controlled V_0 for the initial OD of the culture. Both log transformed OD corrected and non-OD corrected V_0 correlated significantly with MIC (r=0.5 without OD correction and r=0.4 with OD correction).

Correlations and linear models. All statistical analysis were performed using the statistics software R (version 2.15.1), with the cor.test function for simple correlations and the lm function for multiple regression analysis. Rgl package was used for 3D structure visualization, plotting amino acid as a sphere centered on the first carbon of the amino-acid.

Protein purification. Genes for TEM-1 and its variants were cloned into pET36b using *Nde*I and *Not*I restriction sites. The ligation products were transformed into BL21 (DE3) *Escherichia coli* Competent Cells (Invitrogen, Fisher Scientific, France), and plated on Luria–Bertani agar supplemented with kanamycin (30 mg/L). The transformants were cultured in 200 mL 2xYT broth containing kanamycin (30 mg/L) at 37 °C for 18 hours. Overexpression of the beta-lactamase was induced with 1 mM of IPTG at 22 °C for 5 hours. Cells were harvested by centrifugation (7,000 g for five minutes at 4 °C). The cell pellet was suspended in 10 mL HEPES buffer (10 mM, pH 8.0), then disrupted by ultrasonic treatment. The extract was clarified by centrifugation at 11,000 g for ten minutes at 4 °C. The extract was then loaded onto an anion-exchange column (Q Sepharose FF, GE Healthcare; 5cm x 10mm) pre-equilibrated in 10 mM HEPES buffer (pH 8.0) and eluted with a linear 0–400 mM NaCl gradient (2.5 ml/min; 5mM/min). Active fractions were pooled and applied at 0.4 ml/min onto a Superdex 75 column (GE Healthcare; 30cm x 10 mm) equilibrated in 10 mM HEPES (pH 8.0) containing 50 mM NaCl. The enzyme was more than 95% homogeneous as judged by Coomassie blue staining after SDS-PAGE, except for A36D and L250Q mutants for which we could obtained around 80% purity.

Thermal denaturation of proteins. Enzyme denaturation (0.5 µM in 100 mM potassium phosphate buffer, pH 7) was obtained by raising the temperature from 25 to 80°C at 1.5 °C/min ramping rates, using a Jasco FP-8300 fluorescence spectrophotometer. Changes in tryptophan fluorescence (280 nm excitation, 340 nm emission) were followed. The collected spectroscopic data were corrected to eliminate temperature dependent changes in tryptophan

fluorescence of the native and unfolded states (as determined from the linear dependences of fluorescence at the beginning and end of experiments), as to determine the midpoint temperature of transition (Tm).

Enzyme assays. Enzyme assays were performed in 100 mM potassium phosphate, 1mM EDTA buffer (pH 7.0). Initial velocity was measure spectrophotometrically at 486 nm using the chromogenic substrate nitrocefin at 50 µM (Calbiochem, La Jolla, CA), using extinction coefficient ε= 20,500 M^{-1} cm^{-1}. Initial velocity was measured between 27°C and 67°C with 5°C ramp, after pre-incubation of the enzyme for 5 minutes at the temperature to be assayed. Finally, $T_{1/2}$ was calculated as the temperature for which a decrease of 50% in initial velocity is observed.

Supplementary Text

Synonymous mutations. We used Kolmogorov Smirnov test (KS test) to compare distribution of mutants MICs. The distribution of MICs of synonymous mutants was similar to that of wild-type (KS test p=0.96). The one of mutants having a single amino acid change was similar to the one of mutants having one amino acid changes in combination with some or no synonymous mutations (KS test p=1) (Fig. S2).

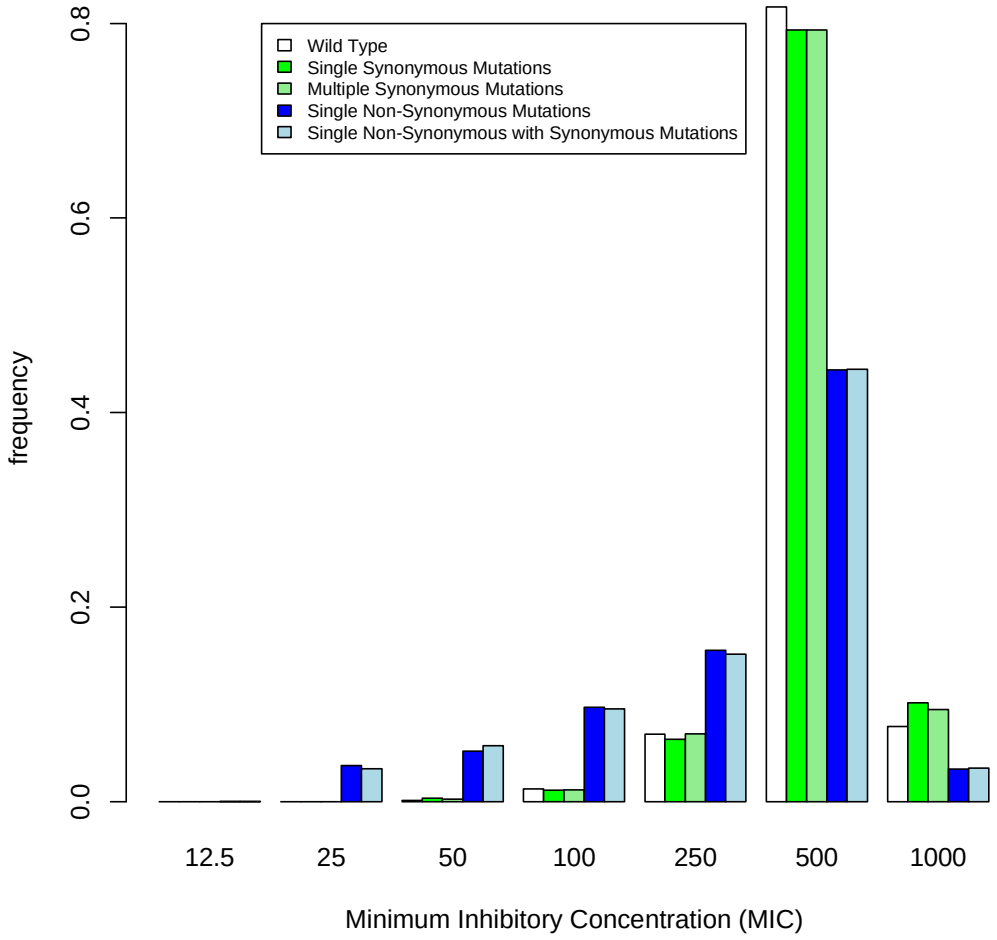


Fig. S2: Distribution of MICs, in mg/L, for wild-type (white, n=1594), and mutants bearing: a single synonymous mutation (green, n=590), multiple synonymous mutations (light green, n=750), one non-synonymous mutation alone (blue, n=1402) or in combination with some synonymous mutations (light blue, n=2313). The effect of synonymous mutations does not differ from the measurement error on the wild type, therefore synonymous mutations can be ignored, and only non-synonymous mutations eventually in combination with synonymous mutations can be used.

Active site. Based on the extensive knowledge accumulated on beta-lactamase TEM-1, an extended active site, composed of the pocket receiving the substrate was identified (2). The distribution of the effects of mutations in that active site was drastically different from the rest of the protein (KS test $p<1e-7$), with a larger fraction of damaging mutations, as expected. Yet the active site being only a small fraction of the protein, the overall distribution of mutations with and without the active site was marginal (KS test $p=0.60$) (Fig. S3).

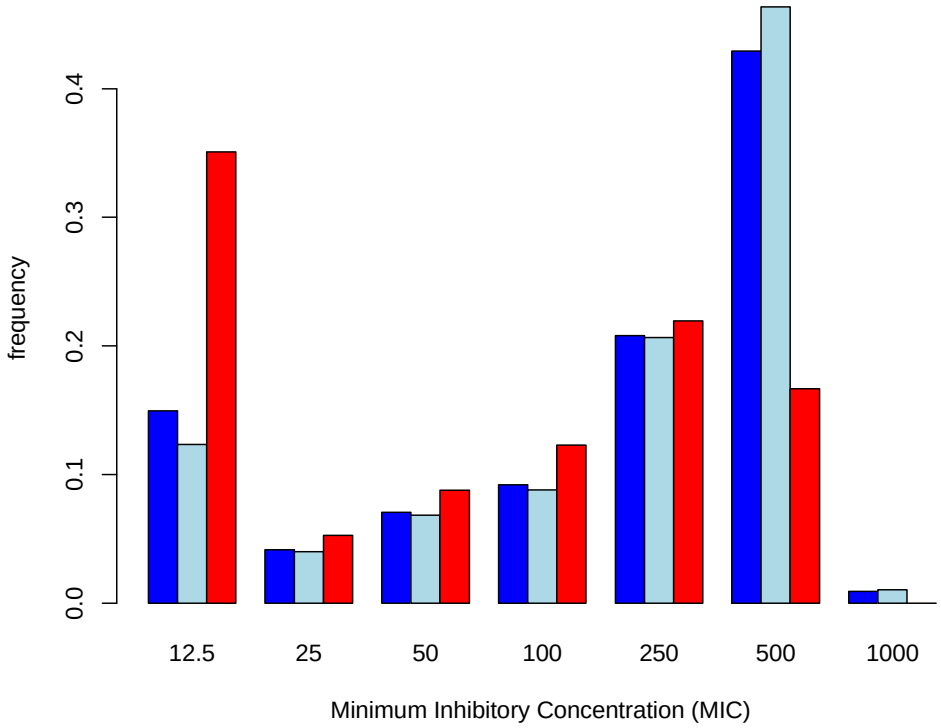


Fig. S3: Distribution of MICs, in mg/L, for amino acid changes in the whole enzyme (blue, n=990) the enzyme without the extended active site (light blue, n=876) and the active site (red, n=114)

Fitting the distribution of mutants MICs in TEM-1 with continuous distributions. To fit the semi-quantitative values of MICs (in mg/L) with some distributions, we first estimated the noise in the measure by fitting a normal distribution on the MIC scores of the wild type (n=1594) and found a standard deviation of 0.37. Then, for each continuous distribution, 1 million points were sampled and a measurement noise added to each. This distribution was binned in MIC classes corresponding to the experimental measures. The likelihood of the observed distribution of MIC scores was then computed on the binned distribution. The optim function in R software was used to find the set of parameters that maximized the likelihood (Table S2, Fig. S4).

The biophysical model used is derived from Wylie and Shakhnovich (3). It assumes that:

- MIC is directly proportional to the fraction of protein folded in the native state.
- the fraction of folded protein is directly related to the protein stability, more precisely to the protein free energy, ΔG, which reflects the difference between the entropy of the native fold and the cumulative entropy of other folds.
- the effect on stability of mutants or $\Delta\Delta G$, is drawn in a shifted normal distribution.

Therefore, to fit the distribution of mutants, the model has 3 parameters:

- M, a scaling parameter, which associates a maximum MIC to a genotype with 100% of folded proteins.
- ΔG_{μ}, a parameter describing the average free energy of mutants (μ in text).
- σ, describing the standard deviation of the ΔG_{μ} (or $\Delta\Delta G$) of mutants.

Hence, the model can be written as:

$$MIC\ score\ (TEM1\ mutants) = M + log2\left(\frac{1}{1 + e^{\frac{\Delta G\mu + \mathcal{N}(0,\sigma)}{kT}}}\right) + error$$

Where k, is the Boltzmann constant, T the temperature (kT=0.62 kcal/mol), and $\mathcal{N}(0,\sigma)$ a normal distribution with mean 0, and standard deviation σ. Note that the error is distributed as $\mathcal{N}(0,0.37)$.

Distribution	Number of parameters (+ lethal)	Parameters values (+ lethal when they improve the model)	Log Likelihood (+ lethals)
Exponential	2 (3)	max= 0.08 (-0.19) rate= 0.45 (0.91) lethals= 0 (0.117)	-1664 (-1649)
Gamma	3	max= -0.36 shape=0.26 rate=0.11	-1581
Frechet	3	max=-0.12 shape=0.75 scale=0.48	-1598
Beta	4	max= -0.34 range=34.04 shape1=0.31 shape2=4.80	-1581
Biophysics (see text)	3	M= -0.36 σ, sd of mutant's ΔG= 2.49 ΔG_{μ} mutant's mean ΔG =-0.62	**-1576**

Table S2: Likelihood of the distribution of MIC scores ($\log_2$(MIC/500)) with different models. For each model an additional parameter, a fraction of inactivating mutations or lethal mutations, was tested and is represented in parenthesis only when it improved the model.

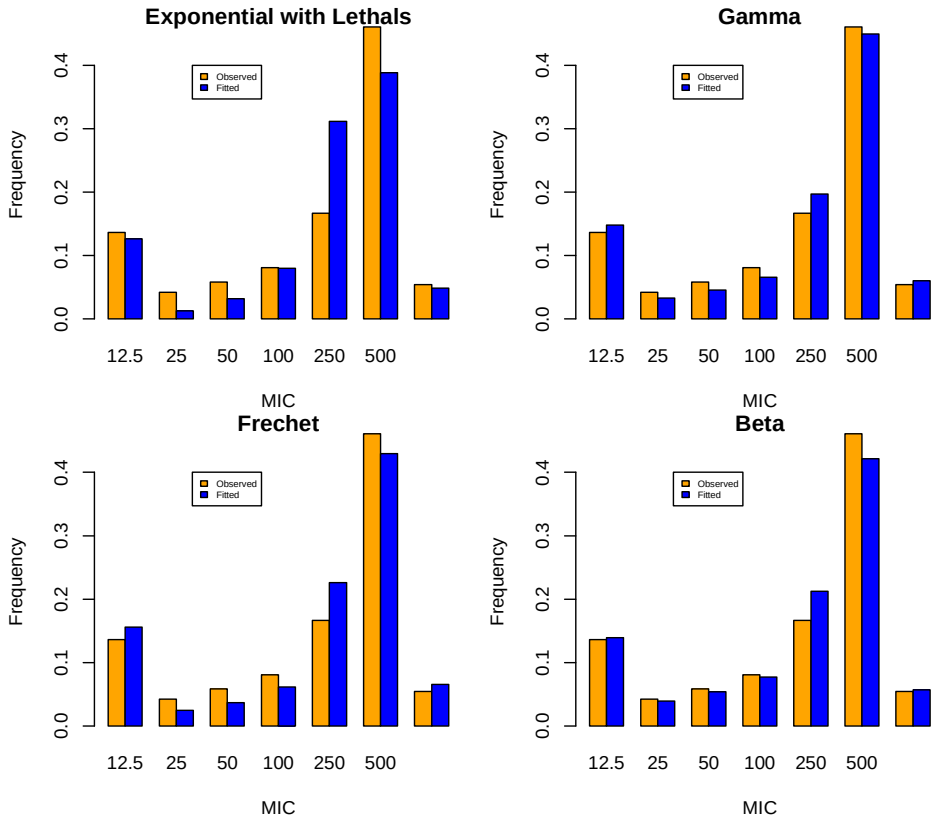


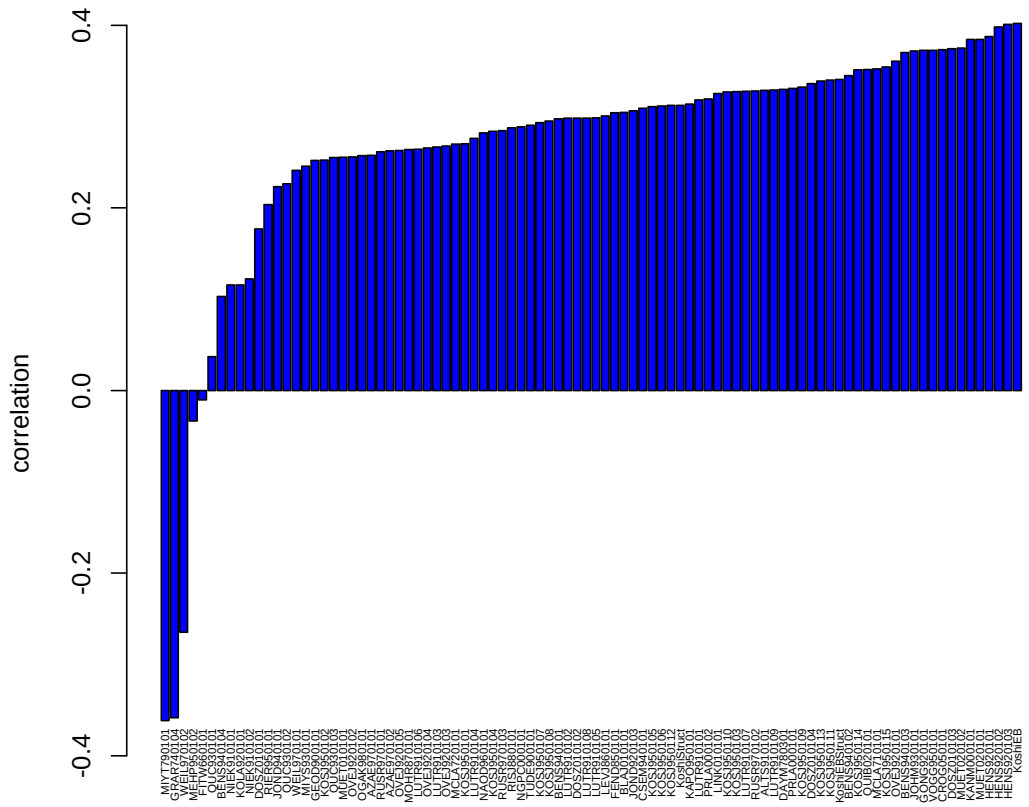
Fig. S4: Visual representation of the quality of the fit provided by the different classical distributions. The fitted and observed data are represented by blue and orange bars respectively. MICs are expressed in mg/L.

Amino acid matrices. Matrices were downloaded from http://www.genome.jp/aaindex/ (4). They are either amino acid transition penalty, reflecting the cost to attribute to an amino acid change in a protein alignment, or amino acid biophysical distance. In the first case a low score reflects a low probability mutation or in functional terms a costly mutation, in the latter case a low score reflects a high similarity and therefore a high probability mutation. Hence depending on the type of matrices positive of negative correlations are expected.

First for each matrix, we computed a correlation C1, between the MIC scores of each mutant and the score of its underlying mutation in the matrix (Fig. S5A). Three additional correlations were derived using the Koshi-Goldstein matrices (5) that are conditional on accessibility and/or secondary structure. Using the accessibility and secondary structure data, three penalty scores were attributed to each mutation, one taking into account accessibility (KoshiEB), one the secondary structures (KoshiStruct), and the last one both criteria (KoshiEBStruct).

Using all mutants from TEM-1, we could also compute matrix associating to each type of amino acid change an average MIC score change. A correlation, C2, between this matrix and the ones previously described was then computed (Fig. S5B).

A



B

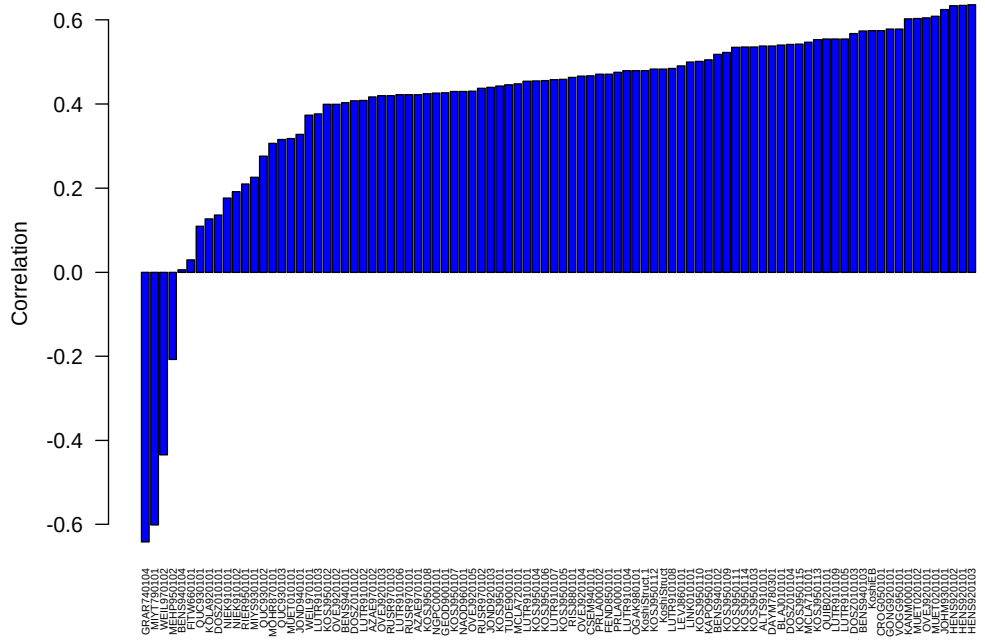


Fig. S5: A) Sorted correlations, C1, between different amino acid transition penalty (positive correlations) or amino acid distances (negative correlations) and the MIC score of mutants. Matrices names come from AAindex

database. B) Sorted correlations, C2, between the different amino acid matrices and the matrix of mean effect on MIC score of mutations in TEM-1.

Building a predictive linear model of MIC in the TEM-1 library

Using R, we could build a linear predictive model of the data. We sampled randomly 100 to 400 mutants from the library and build a linear model with all three factors (BLOSUM 62 penalty matrix, residue accessibility and ΔΔG prediction by PopMusic) and their interactions with the "lm" function of R. This model was used to predict the MIC of the other mutants and the correlation between predicted and observed MIC was computed. For each size of the learning sample (the ones used to build the model), a 1000 different combination of mutants were used. With learning sample size exceeding 200, an average correlation of 67% was recovered.

Model based correlation between MIC scores and ΔΔG predictions. While we can easily compute the correlation between the predicted ΔΔG and the MIC score of mutations, the impact of ΔΔG on MIC should ideally be computed through equation 1 from the text. However for that purpose we need the ΔG of TEM-1, $\Delta G_{TEM\text{-}1}$, to compute the ΔG of mutants as $\Delta G_{mutant}=\Delta G_{TEM\text{-}1} +\Delta\Delta G_{mutant}$. Yet using the values of $\Delta G_{TEM\text{-}1}$ of -8 kcal/mole as measured *in vitro* (3), a very poor correlation was found (at most 0.15 (0.11) using PopMusic (FoldX) predictions excluding the active site). As many biological parameters may affect $\Delta G_{TEM\text{-}1}$, making for instance some non-native forms more likely, we decided to use $\Delta G_{TEM\text{-}1}$ as a free parameter used to maximize the correlation between predicted ΔΔG and MIC scores. The maximum correlations between the predicted fraction of folded protein and observed MIC scores was 0.52 (0.45) for PopMusic (FoldX) predictions excluding the active site. The value of $\Delta G_{TEM\text{-}1}$ maximizing the correlation was slightly different for the two predictions (-1.57 kcal/mole for PopMusic and -1.96 kcal/mol for FoldX), we therefore used the value maximizing both that was -1.73 kcal/mole (Fig. S6).

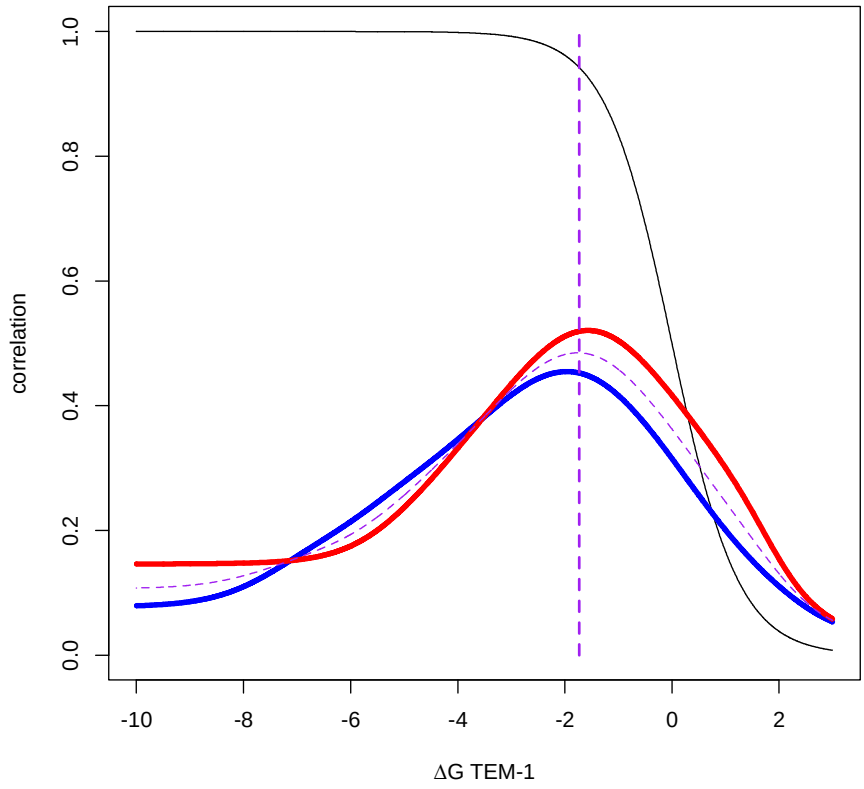


Fig. S6: Correlation between observed MIC and the predicted MIC derived with equation 1 from main text for different values of ΔG_{TEM-1} on the x-axis. The active site is excluded. In red, the predictions from PopMusic are used, in blue the ones from FoldX and in purple, the average between both correlations is plotted. The vertical dotted line show the value that maximizes both correlations, -1.73 kcal/mole, that is used further in the text. Black line shows the fraction of folded protein as a function of ΔG.

Correlation between growth rate and molecular determinants. We performed growth curves for 757 mutants with a single amino acid change. These experiments have been performed in triplicate at two concentrations of amoxicillin, 6 and 100 mg/L, and we determined for both concentrations the maximum growth rate for each mutant as a direct fitness linked determinant. In the manuscript, we used the MIC as a proxy of the fitness, and dealt with the variance of the mutants' MIC explained by the different determinants alone or in combination. Here, we have been interested in the correlation of MGR with the same parameters (Table S3).

With the MIC method, all mutants with small effects have the same MIC: therefore, we cannot conclude from the MIC statistical analysis if the factors we identified as important over the whole range of MIC remain relevant for small effect mutations. Yet, we found that

the growth rates at high concentration of mutants having the wild-type MIC of 500 mg/L were still correlated significantly with these factors (Table S4).

	Variance explained	
	MGR at 6 mg/L	MGR at 100 mg/L
BLOSUM62	0.09	0.10
Accessibility	0.10	0.20
ΔΔG Popmusic	0.14	0.15
ΔΔG foldX	0.09	0.09
BLOSUM62 + Accessibility	0.20 (0.25)	0.31 (0.32)
BLOSUM62 + ΔΔG Popmusic	0.18 (0.19)	0.21 (0.21)
BLOSUM62 + ΔΔG foldX	0.13 (0.16)	0.15 (0.15)
Accessibility + ΔΔG Popmusic	0.16 (0.18)	0.23 (0.23)
Accessibility + ΔΔG foldX	0.16 (0.20)	0.24 (0.24)
BLOSUM62 + Accessibility + ΔΔG Popmusic	0.22 (0.28)	0.32 (0.34)
BLOSUM62 + Accessibility + ΔΔG foldX	0.22 (0.30)	0.32 (0.34)

Table S3. Fraction of variance of the mutant's MGR explained by the different factors alone or in combination. MGR were determined for growth curves performed at 6 and 100 mg/L of amoxicillin. The adjusted R square is given for the combination of factors without or with (in parenthesis) interactions among factors.

	Variance	p-value
BLOSUM62	0.008	7.10^{-2}
Accessibility	0.07	10^{-6}
ΔΔG Popmusic	0.07	$4.8.10^{-6}$
ΔΔG foldX	0.03	$3.9.10^{-3}$

Table S4. Fraction of variance of the mutant's MGR explained by the different factors. This analysis is restricted to mutants with MIC=500 mg/L for MGR measured at 100 mg/L.

Correlation between initial velocity and molecular determinants. To further assay the robustness of our results based on MIC, we also estimated the initial velocity of cell extracts on the substrate nitrocefin, a chromogenic beta-lactam. Though the measurement of V_0 in 96 well plates was noisy, all the factors identified previously correlated with V_0 measurements whether they were corrected by the culture optical density or not (Table S5).

	Variance explained	
	No OD correction	OD Correction
BLOSUM62	0.09	0.06
Accessibility	0.11	0.07
ΔΔG Popmusic	0.08	0.06
ΔΔG foldX	0.09	0.05
BLOSUM62 + Accessibility	0.19 (0.19)	0.13 (0.13)
BLOSUM62 + ΔΔG Popmusic	0.13 (0.14)	0.09 (0.10)
BLOSUM62 + ΔΔG foldX	0.14 (0.14)	0.08 (0.09)
Accessibility + ΔΔG Popmusic	0.12 (0.17)	0.10 (0.10)
Accessibility + ΔΔG foldX	0.17 (0.17)	0.09 (0.11)
BLOSUM62 + Accessibility + ΔΔG Popmusic	0.19 (0.22)	0.13 (0.14)
BLOSUM62 + Accessibility + ΔΔG foldX	0.22 (0.23)	0.14 (0.15)

Table S5. Fraction of variance of the mutant's log-transformed initial velocity ($\log(V_0)$) explained by the different factors. Results are displayed with or without optical density correction.

Fitting the multiple distributions with the biophysical model. The model-based approach can be used to go further and be used not only TEM-1 mutants MIC, but also to fit other distributions such as the ones obtained in M182T background. Compared to the first three parameters model described previously, two additional parameters are required: the stability of TEM-1, $\Delta G_{TEM\text{-}1}$ and the stabilizing effect of M182T, $\Delta\Delta G_{M182T}$. Either we can estimate those parameters on the whole dataset, or we can use the value from the literature for the stabilizing effect of M182T and the value of $\Delta G_{TEM\text{-}1}$ that we have estimated using the correlation approach and the PopMusic and FoldX predictions (next section). As previously we used the "optim" method to find an optimal combination of parameters, yet as the parameters have some biological meaning, we also used a Monte Carlo Markov Chain with Metropolis Hasting (MCMCMS) method to evaluate the posterior distribution of each of these parameters.

Model with 5 free parameters. First, 5 parameters were estimated, using the distribution of MIC measurements on the wild types in TEM-1 (1571) and M182T (n=855) as well as on the distribution of MIC of mutants in TEM-1 (n=990) and M182T (n=198) backgrounds. (For the sake of convergence of the optimization methods, the lower tail of the distribution of MIC of

the wild types, which does not match with the Gaussian distribution assumed for measurement noise, was omitted). The parameters were the following:

-M, the "maximum" parameter, which associates a maximum MIC to a genotype with 100% of folded proteins.

-$\Delta G_{TEM\text{-}1}$, the ΔG of wild type TEM-1

-$\Delta\Delta G_{M182T}$, the ΔΔG of M182T, in other words the impact of M182T on TEM-1 free energy.

- $\mu = \overline{\Delta\Delta G mutants}$, the mean change in free energy, ΔΔG, of mutants, in other words the mean impact of mutants on free energy.

- σ, the standard deviation of ΔΔG of mutants.

While the first parameter is in units of MIC score, the others are in kcal/mole. The model used to fit the data was the following:

$$MIC\ score\ (Wild\ \ Type\ TEM1) = M + log2\left(\frac{1}{1 + e^{\frac{\Delta GTEM1}{kT}}}\right) + error$$

$$MIC\ score\ (Wild\ \ Type\ M182T) = M + log2\left(\frac{1}{1 + e^{\frac{\Delta GTEM1 + \Delta\Delta GM182T}{kT}}}\right) + error$$

$$MIC\ score\ (TEM1\ mutants) = M + log2\left(\frac{1}{1 + e^{\frac{\Delta GTEM1 + N(\mu,\sigma)}{kT}}}\right) + error$$

$$MIC\ score\ (M182T\ mutants) = M + log2\left(\frac{1}{1 + e^{\frac{\Delta GTEM1 + \Delta\Delta GM182T + N(\mu,\sigma)}{kT}}}\right) + error$$

The posterior distribution was evaluated from 5 independent chains, which ran for 100,000 cycles, using Metropolis Hasting rejection method and a burning of 10,000 cycles. One every 100 points in the chain was sampled for statistical analysis. At each cycle, each parameter had 50% chance of being shifted by a normal distribution of mean 0, and standard deviation 0.1.

Parameter	Median	95% Confidence Interval	unit
M	-0.37	[-0.40, -0.34]	MIC score
M	387	[380, 396]	mg/L
$\Delta G_{TEM\text{-}1}$	-1.38	[-1.56, -1.22]	kcal/mole
$\Delta\Delta G_{M182T}$	-2.04	[-2.69, -1.42]	kcal/mole
$\mu = \overline{\Delta\Delta Gmutants}$	0.66	[0.38, 0.92]	kcal/mole
σ	2.69	[2.43, 2.98]	kcal/mole

Table S6: Posterior distribution of the parameters of the 5 parameter biophysical model of protein stability (including active site)

Parameter	Median	95% Confidence Interval	unit
M	-0.37	[-0.40 -0.34]	MIC score
M	387	[379, 395]	mg/L
$\Delta G_{TEM\text{-}1}$	-1.38	[-1.57, -1.23]	kcal/mole
$\Delta\Delta G_{M182T}$	-2.52	[-3.22, -1.83]	kcal/mole
$\mu = \overline{\Delta\Delta Gmutants}$	0.52	[0.24, 0.78]	kcal/mole
σ	2.51	[2.26, 2.80]	kcal/mole

Table S7: Posterior distribution of the parameters of the 5 parameter biophysical model of protein stability (excluding active site)

All parameters, except M, have been used here as fitting parameters while they also have some estimates in the literature. We can therefore make some comparison, focusing mostly on the values obtained without the active site (Table S6 and Table S7). The discrepancies come from $\Delta G_{TEM\text{-}1,}$ which is several units higher than expected (-1.38 compared to -8 kcal/mole (3)), from σ that is higher than what has been described (around 2.5 compared to 1.7 kcal/mole (3)) and to a lesser extent from μ that is less than expected (0.5 compared to 1 kcal/mole (3)). Here are some suggestions that might explain these differences. First μ and σ were extrapolated from pools of mutants with biophysical informations extracted from the literature (3), and the assumption that the distribution of mutation effect on $\Delta\Delta G$ may be conserved across all proteins may be wrong. Second, some biases may exist in the estimation

of these parameters from collections of mutants. An underestimation of σ is possible if only mutants that conserved some activity were analyzed. Alternatively, biophysical measures in the laboratory may differ from the ones *in vivo*. The high density of proteins in the cell may destabilize the protein more than predicted *in vitro*, moreover mutations may affect the folding dynamics of the proteins. Finally, the destabilizing effect of $\Delta\Delta G_{M182T}$ is perfectly in the range predicted *in vitro* (6).

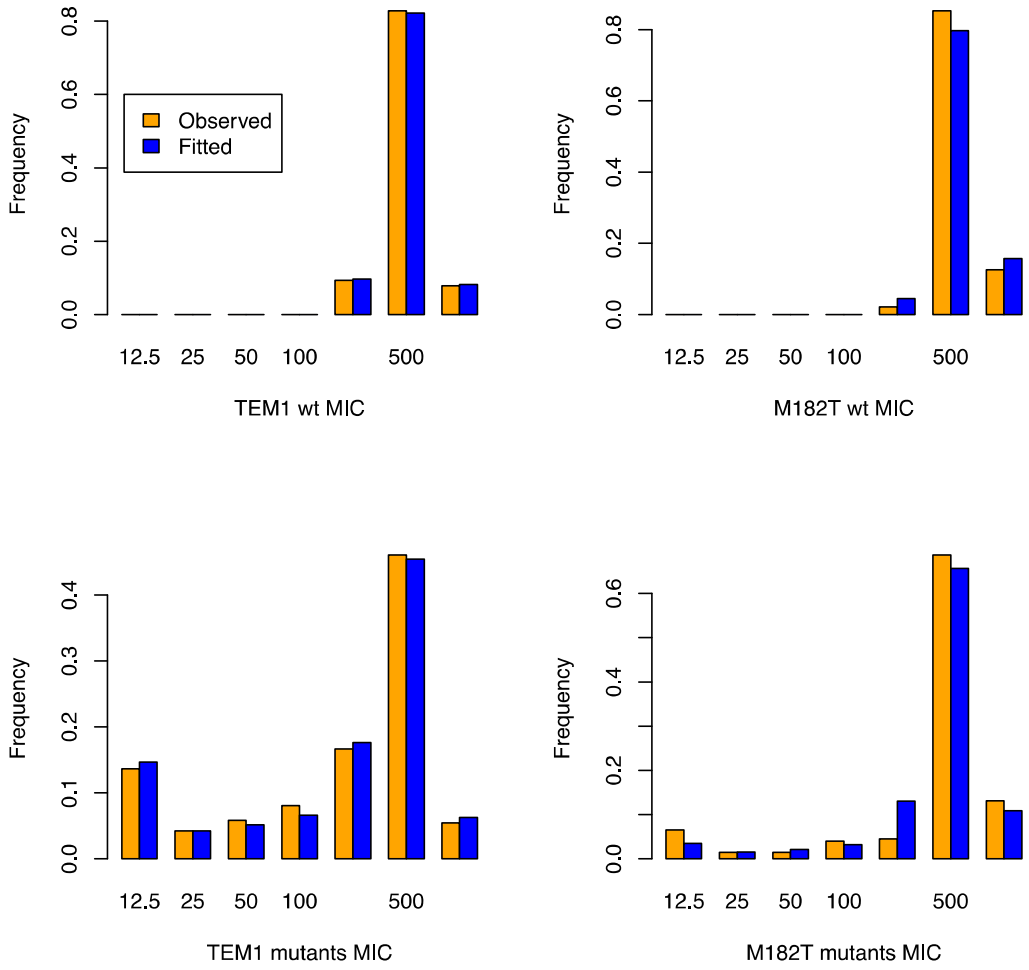


Fig. S7: Visual representation of the quality of the fit provided by the biophysical model of stability with 5 parameters. Blue bars represent the fitted data and orange bars the observed data (including active site).

***Fit on three parameters*:** We also performed the fit using a model in which the values of $\Delta\Delta G_{M182T}$, and $\Delta G_{TEM\text{-}1}$ were fixed (Table S8). First, the log-likelihood of the chain (median -2874.95) was very similar to the one of the 5 parameter model of previous section (median -2870.23), the difference between the two being almost completely explained by the 2 units

penalty classically associated to each additional parameter. Hence this model is almost as good as the previous one with only three parameters. Second, with that configuration, the value of μ used in the literature is in the 95% confidence interval. Finally, using the median values of each parameter, we tested if the 4 distributions of predicted MICs differed from observed ones and found no significant difference. (KS tests, p=0.88 for TEM-1, p=0.30 for M182T, p=0.97 for TEM-1 mutants, and p=0.99 for M182T mutants). Figure S8 presents the result of this fit with a continuous distribution to reveal that the differences among distributions are not just a factor of the discrete decomposition of MIC.

Of note, the slight increase in MIC associated with mutation M182T suggested by the model was validated experimentally using many replicates in liquid. It was also suggested in other works (7)

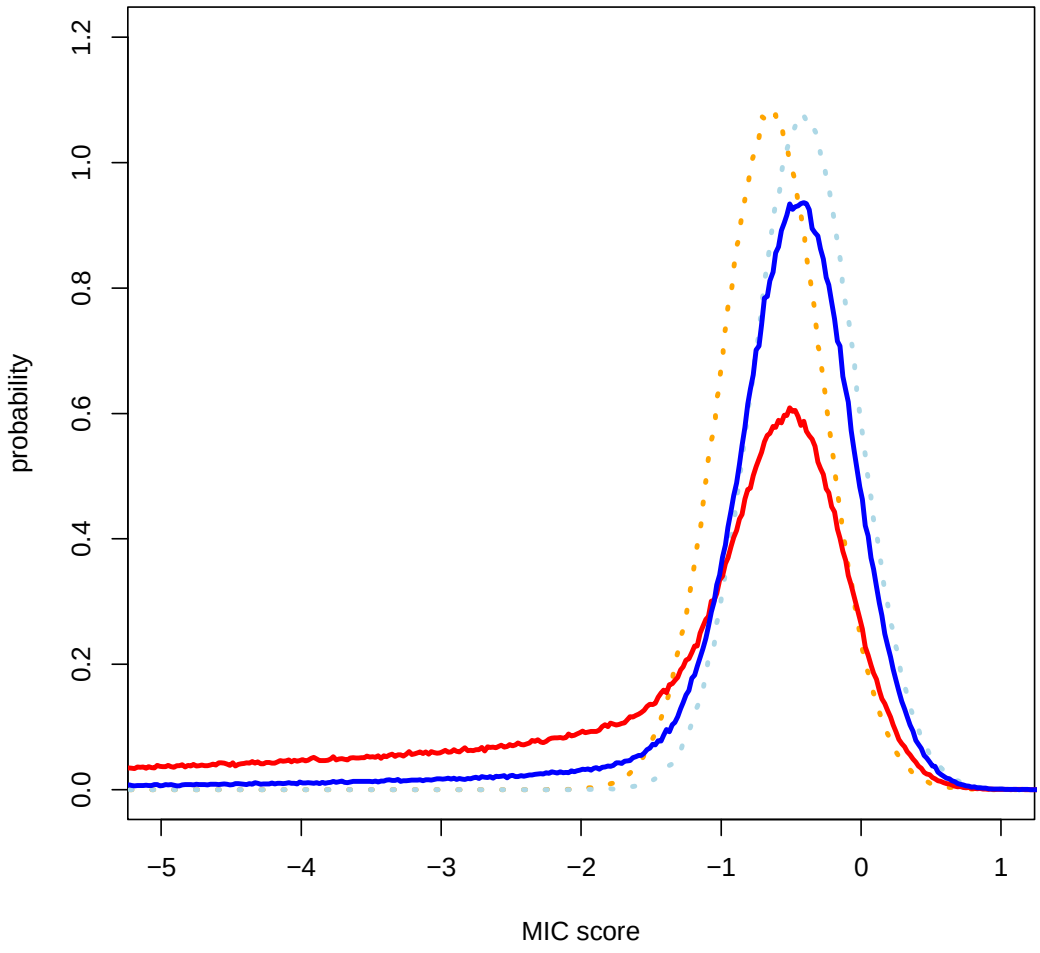


Fig. S8: Probability distribution of the fitted models, in dotted orange the wild type TEM-1, in red the mutants of TEM-1, in dotted light blue, the wild type M182T, and in blue the mutants of M182T.

Parameter	Median	95% Confidence Interval	unit
M	-0.40	[-0.42 -0.39]	MIC score
M	377	[372, 382]	mg/L
ΔG_{TEM-1}	-1.73	fixed	kcal/mole
$\Delta\Delta G_{M182T}$	-2.7	fixed	kcal/mole
$\mu = \overline{\Delta\Delta Gmutants}$	0.76	[0.47, 1.01]	kcal/mole
σ	2.62	[2.36, 2.90]	kcal/mole

Table S8: Posterior distribution of the parameters of the 3 parameter biophysical model of protein stability (excluding active site).

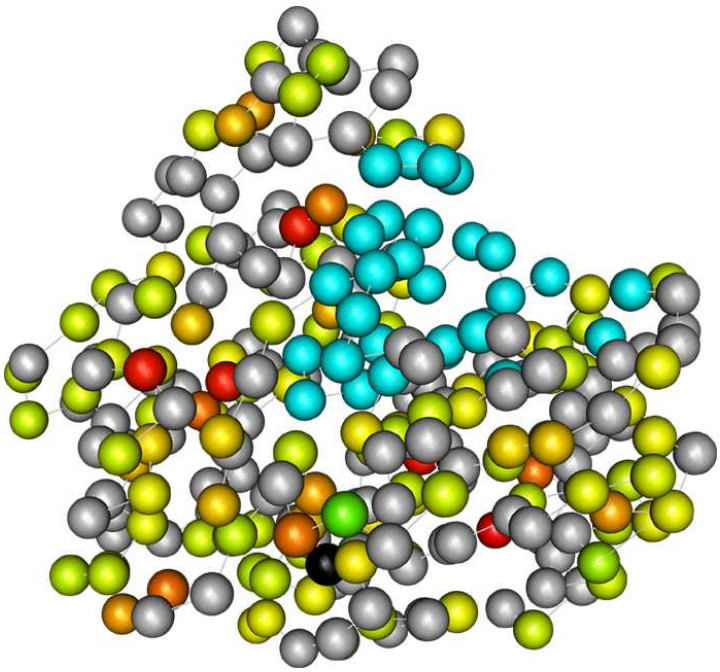


Fig S9: Representation of the epistatic interactions on the three dimensional structure of the protein. The difference between MIC scores in M182T and MIC scores in TEM-1 is presented by a gradient from red to yellow (no interaction) to green. Red represents mutations inactivating TEM-1 and with no detectable effect in M182T background. Amino acids without information are in grey, the active site is in blue, and residue M182 is in black. The mutations showing strong interactions (red) with M182T background are not close to residue 182.

Single mutants informations. Primary data of single mutants of TEM-1 and M182T libraries are available in Dataset S1 and Dataset S2 respectively, downloadable on the PNAS web site.

REFERENCES

1. Drummond DA, Iverson BL, Georgiou G, Arnold FH (2005) Why high-error-rate random mutagenesis libraries are enriched in functional and improved proteins. *J Mol Biol* 350, 806-16.
2. Doucet N, Savard PY, Pelletier JN, Gagne SM (2007) NMR investigation of Tyr105 mutants in TEM-1 beta-lactamase: dynamics are correlated with function. *J Biol Chem* 282, 21448-59.
3. Wylie CS, Shakhnovich EI (2011) A biophysical protein folding model accounts for most mutational fitness effects in viruses. *Proc Natl Acad Sci U S A* 108, 9916-21.
4. Kawashima *et al.* (2008) AAindex: amino acid index database, progress report 2008. *Nucleic Acids Res* 36, D202-5.
5. Koshi JM, Goldstein R A (1995) Context-dependent optimal substitution matrices. *Protein Eng* 8, 641-5.
6. Wang X, Minasov G, Shoichet BK (2002) Evolution of an antibiotic resistance enzyme constrained by stability and activity trade-offs. *J Mol Biol* 320, 85-95.
7. Brown NG, Pennington JM, Huang W, Ayvaz T, Palzkill T (2010) Multiple Global Suppressors of Protein Stability Defects Facilitate the Evolution of Extended-Spectrum TEM ß-Lactamases *J Mol Biol* 404, 832-46.