Universiteit Leiden
The Netherlands

Essentiality of conserved amino acid residues in β-lactamase

Chikunova, A.

Citation
Chikunova, A. (2022, May 31). *Essentiality of conserved amino acid residues in β-lactamase*. Retrieved from https://hdl.handle.net/1887/3304732

Version:	Publisher's Version
License:	Licence agreement concerning inclusion of doctoral thesis in the Institutional Repository of the University of Leiden
Downloaded from:	https://hdl.handle.net/1887/3304732

Note: To cite this publication please use the final published version (if applicable).

Chapter 4

Conserved residues Glu37 and Trp229 play an essential role in protein folding of β-lactamase

Based on the research article:

Aleksandra Chikunova, Max P. Manley, Misbha Ud Din Ahmad, Tuğçe Bilman, Anastassis Perrakis, Marcellus Ubbink (2021). Conserved residues Glu37 and Trp229 play an essential role in protein folding of β-lactamase, *FEBS J.* 288(19): 5708-5722

Abstract

Evolutionary robustness requires that the number of highly conserved amino acid residues in proteins is minimized. In enzymes, such conservation is observed for catalytic residues but also for some residues in the second shell or even further from the active site. β-lactamases evolve in response to changing antibiotic selection pressures and are thus expected to be evolutionarily robust, with a limited number of highly conserved amino acid residues. As part of the effort to understand the roles of conserved residues in class A β-lactamases, we investigate the reasons leading to the conservation of two amino acid residues in the β-lactamase BlaC, Glu37 and Trp229. Using site-directed mutagenesis, we have generated point mutations of these residues and observed a drastic decrease in the levels of soluble protein produced in *Escherichia coli*, thus abolishing completely the resistance of bacteria against β-lactam antibiotics. However, the purified proteins are structurally and kinetically very similar to the wild type enzyme, only differing by exhibiting a slightly lower melting temperature. We conclude that conservation of Glu37 and Trp229 is solely caused by an essential role in the folding process, and we propose that during folding Glu37 primes the formation of the central β-sheet and Trp229 contributes to the hydrophobic collapse into a molten globule.

Introduction

Most proteins are under evolutionary pressure to minimize the number of essential residues, yielding evolutionary robust proteins. Proteins for which the function depends on the exact nature of many amino acid residues are prone to be rendered inactive by random mutations. Nevertheless, in many proteins the conservation of some residues is high, and it is not always obvious why that is the case. In enzymes, the few residues involved in ligand binding and catalysis are generally highly conserved for obvious reasons. However, also other residues, surrounding the active site or even at more distant locations, are conserved. Conservation is taken as a proxy for functional relevance and is extensively used in research for the prediction of functionally important regions[34–36,108]; yet it is often not evident what the function of such residues is. Many factors contribute to the proper functioning of proteins, such as rapid and correct folding, a stable three-dimensional structure, precise positioning of the catalytic residues in the active site, correct dynamic properties, targeting to the correct location in the cell and post-translational modifications. We have started an effort to determine the functions of all conserved residues in β-lactamases and establish whether high conservation equates functional necessity. Specifying the roles of all conserved residues in an enzyme can contribute to a better understanding of the evolutionary robustness of proteins. B-lactamases catalyze the reaction of hydrolysis of β-lactam antibiotics, the largest group of antibiotics, providing bacteria with resistance against these compounds[197]. The current study focuses on the relevance of two of the highly conserved residues in the β-lactamase from Mycobacterium tuberculosis, BlaC. This enzyme is a broad-spectrum class A β-lactamase encoded by a chromosomal gene[101]. The presence of BlaC in this pathogen is the reason that β-lactam antibiotics are not normally used for tuberculosis treatment[73].

Residue Glu37 (Ambler numbering[89]) is present in 99% of class A β-lactamases[109,140,173,198] (across 494 sequences) and has not been reported to be involved in protein function in this class. In class B (metallo-β-lactamases) this residue is also conserved and mutations lead to lower activity[199,200]. The Cα of this residue is located 20 Å away from the Cα of the active site Ser70. In BlaC Glu37 interacts with Ala42, Leu44, Tyr60 and Arg61 (Figure 4.1). Alanine, aspartate, glutamine or serine were also found at this position in 1% of sequences.

Residue Trp229 is 98% conserved in class A β-lactamases. In TEM-1, another class A β-lactamase, mutations at this position significantly decrease enzyme activity[131]. Trp229 is in close proximity to an allosteric site in TEM-1, where two non-competitive inhibitors have been shown to bind[201]. Trp229 is situated 24 Å away from the active site serine, and in BlaC Trp229 interacts with Phe225, Tyr259, Pro254 and Pro226 (Figure 4.1). Alanine, phenylalanine, tyrosine, serine or cysteine are found at position 229 in 2% of sequences.

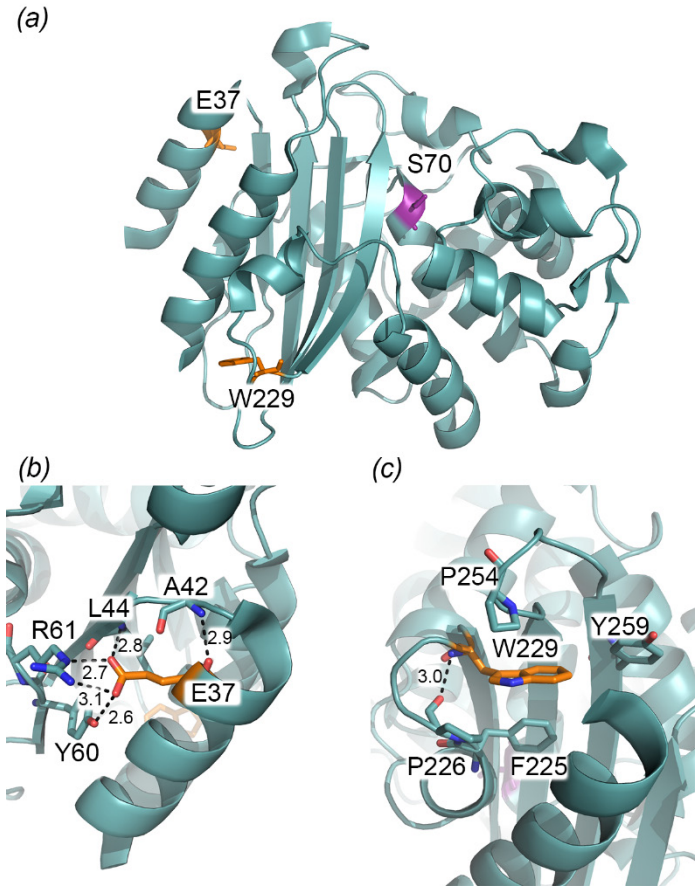

Figure 4.1. Location of Glu 37 and Trp229 in BlaC. (A) Structure of BlaC (PDB entry 2GDN[101]) with residues Glu37 and Trp229 shown in orange sticks and the catalytic residue Ser70 in violet; (B) Detail of the structure surrounding Glu37. Interactions involving Glu37 are drawn in black dashed lines; (C) Detail of the structure surrounding Trp229. Interaction involving backbone of Trp229 is drawn in black dashed line.

To reveal the reason for conservation of these residues in class A β-lactamases, site-directed mutagenesis was performed, replacing Glu37 and Trp229 with a range of amino acids with different properties. We report that the production levels are strongly reduced, indicating that both residues are important for correct folding. Apart from a moderate decrease in thermal stability, revealed by thermal shift assays, the mutations have surprisingly little effect on the structure and activity of the enzyme, as probed with kinetic measurements, CD and NMR spectroscopy and X-ray crystallography. The possible roles of Glu37 and Trp229 in the folding process are discussed.

Results

To investigate the role of Glu37 in BlaC, it was mutated to Ala, Asp, Leu and Gln. The mutation to Ala was chosen to ensure that all interactions involving the side chain (Figure 4.1) are no longer possible. The Asp mutant maintains the functional group but reduces the size, probing interactions in which exact positioning is critical. The mutation to Leu introduces a side chain of the exact same surface area[202] as the wild type residue but with non-polar nature. Gln is a mimic that eliminates the charge but maintains size, polarity and ability to form hydrogen bonds, while allowing to also donate a hydrogen in such interactions.

The aromatic nature of Trp229 was mimicked with mutations to Phe and Tyr. The latter also resembles the amphipathic nature of Trp with its hydroxyl group. The other mutants were Leu, Gln and Ala. Properties of the side chains of these residues are in Table 4.1.

E37 and W229 are essential for function

The survival of BlaC wild type and variants in *Escherichia coli* was tested by applying cells cultures of different dilution on agar plates containing various concentrations of ampicillin and carbenicillin. A catalytically inactive mutant with active site serine 70 replaced by alanine was used as a negative control. Cells producing wild type protein are able to grow with 100 μg mL^{-1} of ampicillin and 500 μg mL^{-1} of carbenicillin, whereas mutant producing cells have very low survival in the analyzed antibiotic concentration range (Figure 4.2a, Figure S4.1).

Cultures that produce BlaC Glu37 mutants do not perform better than the negative control with minimal inhibitory concentrations (MIC) of 3 μg mL^{-1} and 20 μg mL^{-1} for ampicillin and carbenicillin, respectively. Only W229F and W229Y mutants do slightly better and are not able to grow at 5 μg mL^{-1} and 50 μg mL^{-1} of ampicillin and carbenicillin, respectively. Interestingly, cultures producing W229A, W229L and W229Q were outperformed by the negative control at 2 μg mL^{-1} of ampicillin (Figure 4.2b). The observed effect could be explained by increased production of toxic protein, induced by the presence of antibiotic in mutant producing cultures, an effect previously described for some proteins[203]. In cultures incubated at 25°C Glu37 mutants still show no activity against antibiotics. However, for the mutants W229F and W229Y improved growth was observed at 25 °C both on plate and in liquid cultures (Figure 4.2c, Figures S4.2-S4.3), suggesting that the mutant proteins can be formed in reasonable amounts but have reduced thermostability.

Table 4.1. Amino acid residues properties

	Residue	Side-chain ASA (Å²)	Side chain polarity	Charge	Aromaticity
Glu37	**Glu**	138	+	Negative	
	Ala	67	–	Neutral	
	Asp	106	+	Negative	
	Leu	137	–	Neutral	
	Gln	144	+	Neutral	
Trp229	**Trp**	217	–		+
	Ala	67	–		–
	Phe	175	–		+
	Leu	137	–		–
	Gln	144	+		–
	Tyr	187	+		+

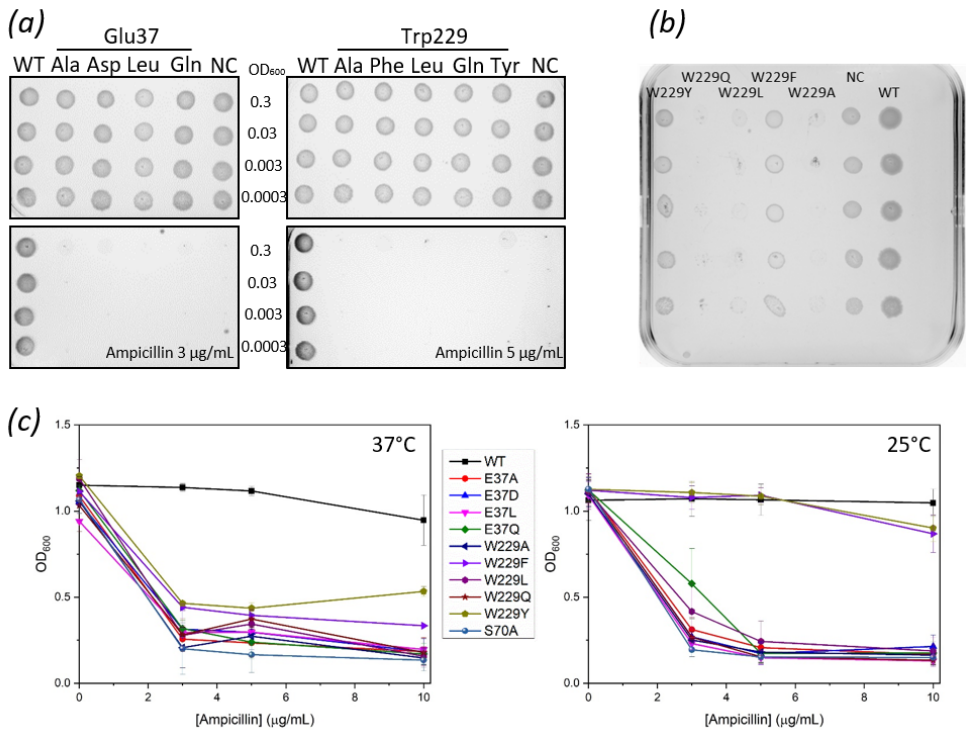

Figure 4.2. Glu37 and Trp 229 mutants are mostly inactive *in vivo*. (A) Plates showing growth at 37˚C of *E. coli* cells expressing wild type (WT) or mutant BlaC (negative control, NC is the S70A mutant) with no antibiotics (top) or a low concentration of ampicillin (bottom, MIC for wild type BlaC > 100 mg mL^{-1}); (B) Plate with 2 µg mL^{-1} ampicillin. The growth of W229A, W229L and W229Q mutant producing cultures is inhibited, whereas the negative control (S70A mutant) still grows; (C) End point OD_{600} values (15 hours at 37 ˚C and 25 hours at 25 ˚C) of *E. coli* cultures expressing wild type and mutant blaC genes as a function of ampicillin concentration. Error bars represent standard deviation of three replicates.

Mutants of E37 and W229 yield less soluble enzyme

For *in vitro* characterization, the genes for the mature wild type and variants BlaC were over-expressed without signal sequence to obtain cytosolic soluble protein. Production of soluble protein was checked using SDS-PAGE. The results (Figure 4.3a) show a large decrease in the amount of soluble protein for Glu37 mutants. The total amount of protein, however, remains similar.

For protein quantification the signal intensities were compared to each other and to samples with known protein concentrations (Figure S4.4) using ImageLab software (BioRad). For all variants, the BlaC total yield varied from 138 µg mL^{-1} to 206 µg mL^{-1} of cell culture, while the soluble fraction contained 70-86% of protein for wild type and only up to 12% for Glu37 mutants. For BlaC E37L, W229A, W229L and W229Q no soluble protein was detected. For

BlaC W229F and W229Y 60-69% of soluble protein was present, however, after purification most of it turned out to be unfolded.

To check the effect of additional chaperones in the cell, expression was also carried out in a strain, in which the production of chaperones could be induced (JM109(DE3) pGKEJ-8). Two chaperone combinations were used, namely the DnaK-DnaJ-GrpE combination and GroES-GroEL. In the presence of chaperones, slightly less insoluble protein was found for the mutants and wild type. However, the amount of BlaC in the soluble fractions did not increase, as judged by SDS-PAGE and an activity test based on nitrocefin conversion in the culture (Figure 4.3b-4.3c).

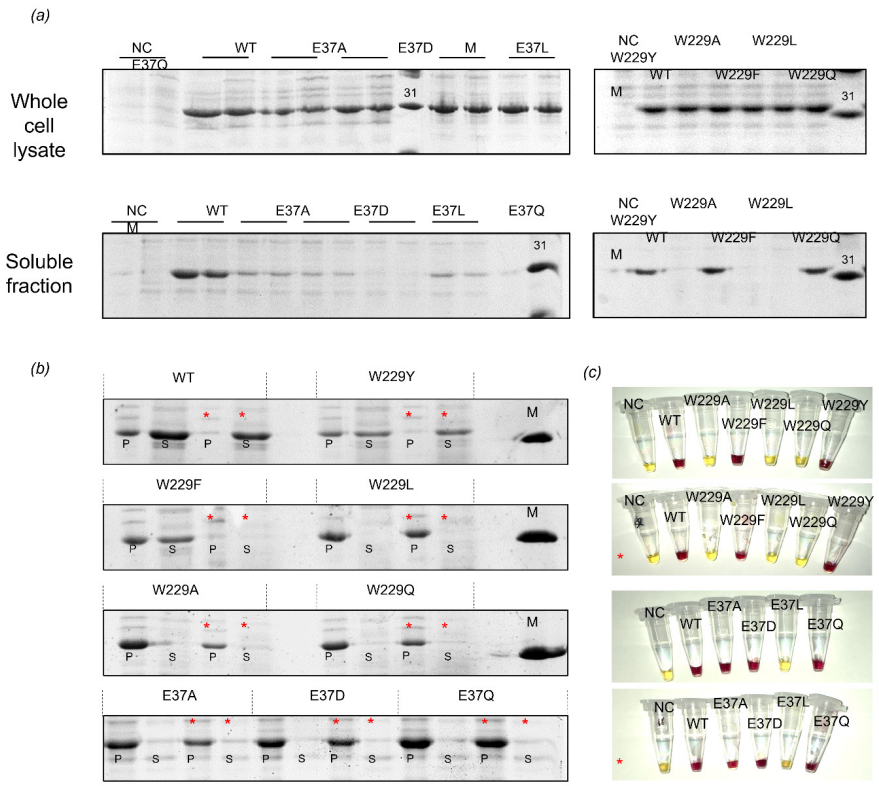

Figure 4.3. (A) Production levels of total BlaC and soluble fraction for wild type (WT) and variants. SDS PAGE showing whole cell lysate and soluble fraction after induction of gene expression. Negative control (NC) indicates non-induced cell culture for the wild type BlaC. M, markers with the 31 kDa band indicated; (B) SDS PA-gels showing pellet (P) and supernatant (S) content of lysed cultures grown with and without chaperones expression. Samples from cultures with chaperones are marked with asterisk. Note that the induction of W229F culture with chaperones failed. M corresponds to the marker band of 31 kDa. The amount of protein in the pellet is reduced in the presence of chaperones, the amount of protein in supernatant, however, is not affected; (C) Nitrocefin activity test for BlaC presence in the supernatant of lysed cultures. Cultures producing chaperones are marked with an asterisk.

E37 and W229 mutants are active in vitro but have reduced thermostability and refolding ability

Insofar mutants yielded soluble BlaC, the enzymes were purified for *in vitro* characterization. The far-UV circular dichroism (CD) spectra showed that folded fractions of all mutant proteins display a secondary structure similar to wild type BlaC, with two peaks at around 210 nm and 222 nm, characteristic of a predominantly helical protein (Figure 4.4a).

Melting temperatures were determined using a thermal shift assay, following the fluorescence signal of SYPRO Orange dye, which increases upon binding to hydrophobic areas of the protein that become exposed during thermal unfolding (Figure 4.4b; Table 4.2). These mutations decrease the thermal stability of BlaC by 2-4 ˚C. Specificity constants of nitrocefin conversion for BlaC variants were shown to be only marginally smaller than that of wild type BlaC (Table 4.2, Figure 4.4c), suggesting that the inability of *E. coli* cells expressing mutants to grow in the presence of β-lactams was due to a lack of sufficient folded, soluble protein rather than inactivity of the mutant enzymes.

A refolding experiment was carried out to compare the extent of refolding of WT BlaC and mutants after denaturation by the chaotropic agent guanidinium and dilution to various final concentrations of this agent. The fraction of refolded, active enzyme was probed by the rate of nitrocefin conversion. BlaC E37A is much more sensitive to the presence of the chaotropic agent than wt BlaC. The sensitivity of BlaC W229F is similar to wt BlaC but refolding is slower (Figure 4.4d).

Table 4.2. Melting temperatures and specificity constants of wild type and mutant BlaC.

	T_{melt} [a] (˚C)	ΔT_{melt} [b] ($T_{melt[mutant]}$-$T_{melt[WT]}$, ˚C)	k_{cat}/K_M (x10^5 $M^{-1}s^{-1}$)	K_M (μM)	k_{cat} (s^{-1})
C-terminal His-tag					
WT	52	-	4.4 ± 0.5	212 ± 17	107 ± 12
E37A	48	-4	3.1 ± 0.9	146 ± 24	44 ± 8
E37D	50	-2	2.8 ± 0.4	217 ± 36	60 ± 9
E37Q	49	-3	4.5 ± 0.9	176 ± 25	67 ± 13
No His-tag					
WT	52	-	4.7 ± 0.5	281 ± 18	133 ± 7
W229F	49	-3	3.2 ± 0.1	403 ± 51	129 ± 10
W229Y	49	-3	3.4 ± 0.6	284 ± 48	98 ± 2

[a] Error 0.5 ˚C; [b] Error 0.7 ˚C

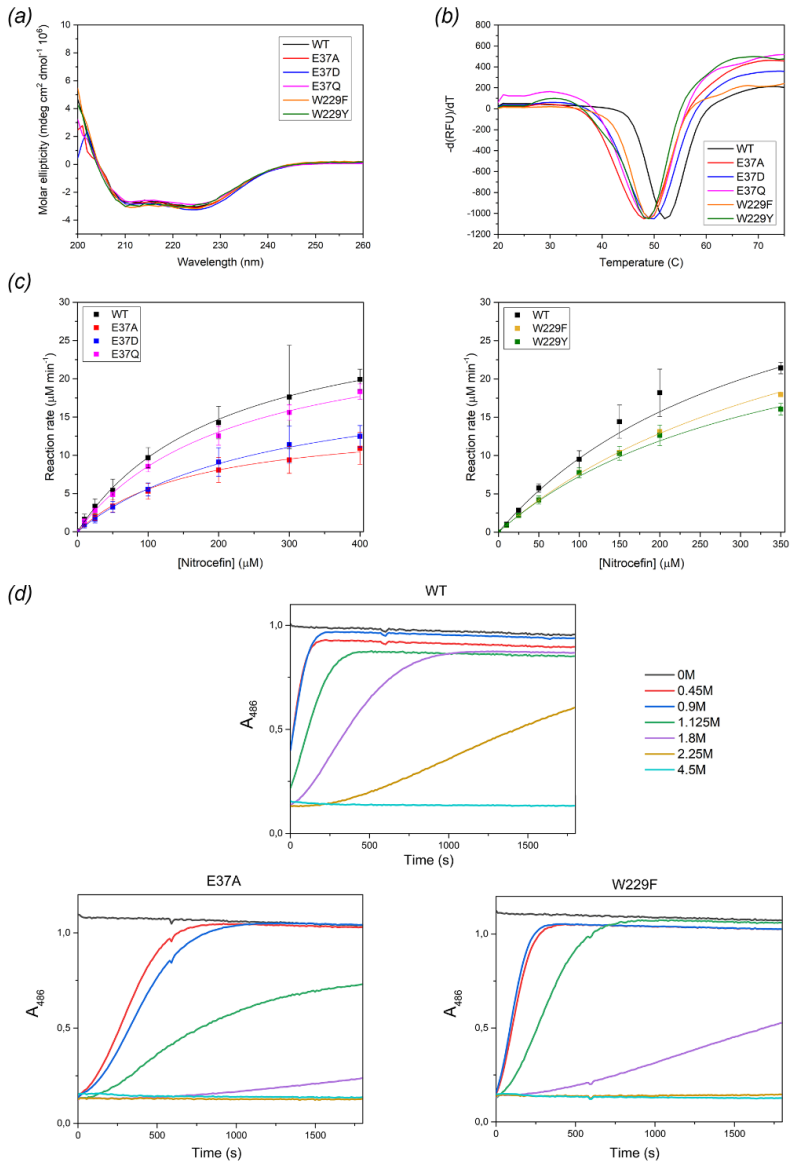

Figure 4.4. Secondary structure and thermostability of WT and mutant BlaC. (A) CD spectra of wild type and soluble BlaC mutants; (B) Normalized negative derivative of thermal shift assay (SYPRO orange binding). The minimum in the curve represents the melting temperature; (C) Kinetics curves of BlaC wild type and mutants for nitrocefin. Experiments were performed at 25 ˚C in 100 mM sodium phosphate buffer, pH 6.4. Error bars represent the standard deviation of five measurements for E37D and E37Q mutants and three measurements for all other samples. A His-tag is present at the C-terminus for E37 variants and wild type at the left panel. No His-tag is present for W229 variants and wild type at the right panel. Experimental data were fitted to the Michaelis-Menten equation (equation 2); (D) Refolding assay for BlaC wt and E37A and W229F. Samples were denatured in 4.5 M Gdn for 20 min on ice and then diluted into the Gdn concentration indicated in the graph with a buffer that also contained nitrocefin (100 mM final concentration). The final enzyme concentration was 0.9 µM. The conversion of the substrate was followed over time at 486 nm and 25 ˚C. The enzyme used for the curves at 0 M Gdn was not unfolded. The curves are averages of three replicate experiments.

E37 and W229 mutations only cause local changes in structure

Chemical shifts perturbations (CSPs) are an indication of the extent of structural changes caused by point mutations. Figure 4.5 shows the backbone amide CSP maps obtained from TROSY-HSQC NMR[204] spectra of the ^{15}N labelled BlaC variants plotted on the crystal structure of wild type BlaC (Figure 4.5). Two features stand out. Firstly, significant CSPs are only localized to regions nearby the mutations, as most clearly shown by W229F and W229Y BlaC. This observation is in line with the CD results that indicate that the secondary structure is unaffected. In combination with the fact that the mutants are capable of converting nitrocefin, it is likely that the three-dimensional structures of the mutants are very similar to that of the wild type BlaC. Indeed, the crystal structure of the E37A mutant we determined at 1.5 Å resolution, does not show large differences compared to the wild type (Table 4.3, Figure 4.6).

The loss of hydrogen bond ability of residue 37 seems to be compensated by a well-ordered water molecule that is not present in the structure of the wild type enzyme (Figure 4.6a). The structures of wild type and E37A BlaC superimpose with an average root-mean-square deviation (RMSD) between Cα atoms of 0.29 Å and the position of catalytic residues match with the great precision (Figure 4.6b). We were unable to crystallise any W229 mutants.

The CSP maps for the three E37 mutants (E37A, E37D and E37Q) are similar, as are the ones for W229F and W229Y. This finding suggests that the mutations disrupt specific interactions that are not or not completely rescued by the mutant side chain. For example, the similarity of the maps of E37A and E37Q indicates that the alanine and glutamine side chains cause the same changes in the chemical environments of the involved amides, relative to a glutamate side chain. Also an aspartate, which is expected to maintain the negative charge present on the glutamate side chain, causes similar CSPs. Thus, both length and charge of the glutamate appear to be critical in wild type BlaC. Similarly, but perhaps less surprisingly, neither tyrosine nor phenylalanine can completely replace the indol of W229.

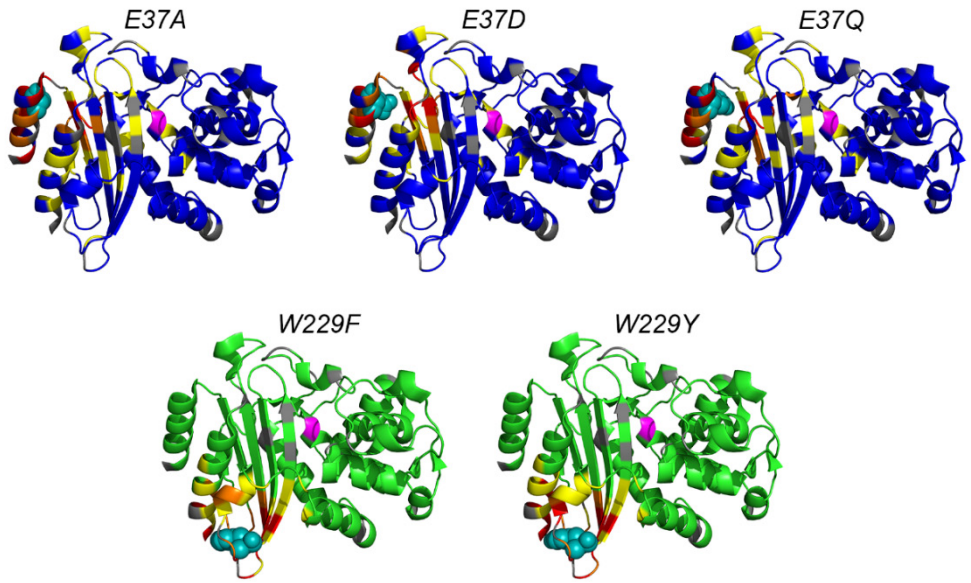

Figure 4.5. Structural effects of the single point mutations. The backbone amide CSP, as defined in the Materials and Methods section, observed for E37A, E37D, E37Q, W229F and W229Y relative to wild type BlaC are plotted on the structure (PDB 2GDN[101]). Blue/green, $\Delta\delta < 0.025$ ppm; yellow $0.025 < \Delta\delta < 0.050$ ppm; orange $0.050 < \Delta\delta < 0.100$ ppm; red, $\Delta\delta > 0.1$ ppm; grey, no data. The mutated residue is modelled and shown as a teal space fill. The location of the catalytic residue S70 is shown in magenta.

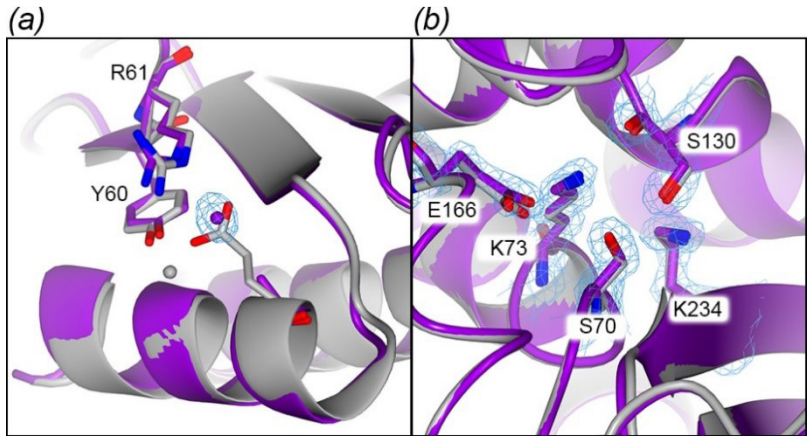

Figure 4.6. Overlay of E37A mutant crystal structure in violet (7A5U) and wild type BlaC in grey (PDB 2GDN6). (A) Mutation site close-up. Residues 37, Tyr60 and Arg61 are shown in sticks. Water found on the surface in wild type structure (grey sphere) is found to be buried in the mutant structure (violet sphere). The 2mFo-DFc electron density corresponding to this water is shown in blue chicken wire; (B) Active site close-up. Catalytic residues are shown in sticks. Blue chicken wire shows the 2mFo-DFc electron density of catalytic residues in mutant structure.

Table 4.3. Data collection and refinement statistics for the structures of BlaC E37A, PDB entry 7A5U.

Data Collection	7A5U
Wavelength (Å)	0.966
Resolution (Å)	54.30-1.50 (1.53-1.50)
Space group	P 1 21 1
Unit cell a, b, c (Å)	38.56, 54.74, 54.51
α, β, γ	90.00, 94.99, 90.00
$CC_{1/2}$	99.9 (76.8)
R_{pim} (%)	2.3 (35.8)
$\lvert I/\sigma I \rvert$	9.1 (1.5)
Completeness (%)	91.5 (97.0)
Multiplicity	2.5
Unique reflections	33194
Refinement	
Atoms protein/ligands/water	2021/9/148
B-factors protein/ligands/water (Å²)	16/20/31
R_{work}/R_{free} (%)	12.5/17.6
Bond lengths RMSZ/RMSD (Å)	0.998/0.014
Bond angles RMSZ/RMSD (°)	1.065/1.772
Ramachandran plot preferred/outliers	247/2
Ramachandran plot Z-score	-0.51
Clash score	2.23
MolProbity score	1

Discussion

The results presented in this work show that BlaC variants of both Glu37 and Trp229, once folded, behave like the wild type protein. NMR spectra do not show major changes compared to the wild type BlaC spectrum, and most CSPs are found in the vicinity of the mutation site. Also, the kinetic parameters of nitrocefin hydrolysis are comparable to those of the wild type enzyme. However, cells producing mutated protein were not able to grow even on the lowest concentrations of antibiotics. In each case, a large fraction or even all of the mutant protein was insoluble and found in the pellet.

Residue Glu37 is 20 Å away from the active site, yet it is 99% conserved[109140]. In BlaC Glu37 makes a hydrogen bond with the side chain of Tyr60 and two hydrogen bonds with Arg61, fixing the turn in between two β-strands. Furthermore, it makes a side chain-to-backbone interaction with Leu44 and backbone-to-backbone interaction with Ala42, which "staples" the first α-helix to the first β-strand. Whereas the backbone-to-backbone interaction can be achieved with any other residue, the remaining hydrogen bonds require a polar side chain. Multiple-sequence alignment shows that residues Tyr60 and Arg61 are present in only 56% and 85% of the sequences, respectively. However, the nature of these residues remains conserved. In position 60 mostly aromatic residues are found and position 61 is occupied exclusively by amino acids that contain an amide, amine or guanidinium group in the side chain. Furthermore, in multiple structures the guanidinium group of arginine is located 3.4-4.2 Å from the aromatic ring of residue 60, which makes it possible to form a cation-π interaction. This suggests that for Arg, His or Lys at position 61 such cation-π interaction is also plausible, and a N-H-π interaction could be present for Asn or Gln residues (Figure 4.7)[205]. The hydrogen bond between Tyr60 and Glu37 is not conserved because position 60 is not always a Tyr residue [84,101,190,191,206,207]. On the other hand, a hydrogen bond between Glu37 and residue 61 is conserved and the 60-61 N-H-π or cation-π interaction also seems to be present in all β-lactamases of class A, and we propose that the loss of these hydrogen bonds is the cause of the 3 °C decrease in melting temperature of BlaC upon mutation of Glu37.

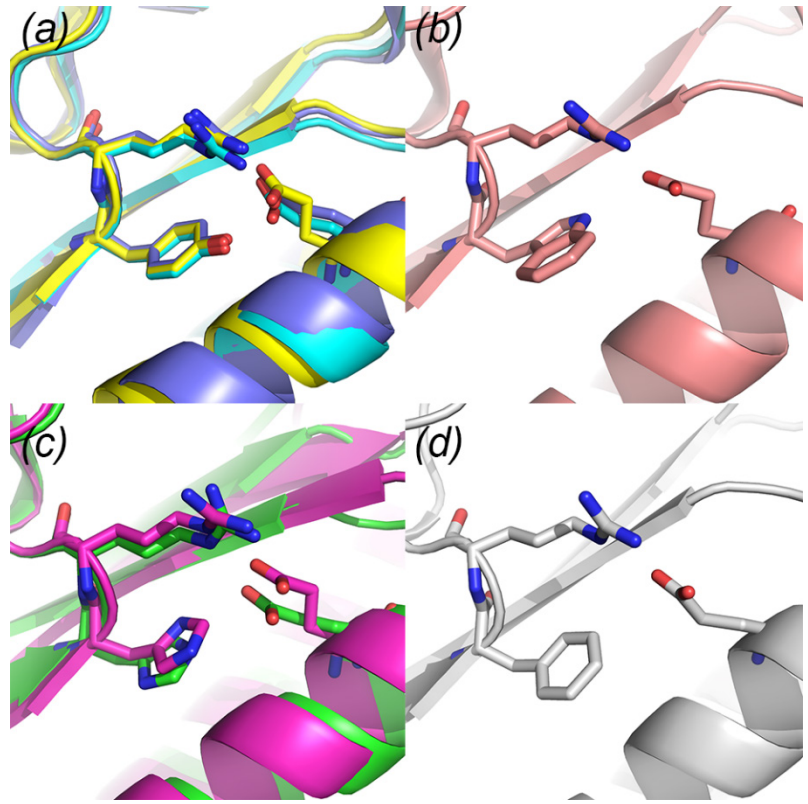

Figure 4.7. Structural overlay of some of the β-lactamase crystal structures, residues 37, 60 and 61 are represented with sticks. (A) Tyr60 containing structures: BlaC - purple (2GDN[101]), β-lactamase from *Streptomyces albus* - yellow (1BSG), BS3 from *Bacillus licheniformis* – cyan (1I2S[206]); (B) Trp60: SHV1 from *Klebsiella pneumoniae* – salmon (2ZD8[190]); (C) His60: GES1 from *Klebsiella pneumoniae* – green (2QPN[175]), β-lactamase from *Mycobacterium abscessus* – magenta (4YFM[84]); (D) Phe60: TEM1 from *Escherichia coli* – white (1BTL[191]).

Our results show that both charge and exact location of the carboxyl group are important because neither Asp nor Gln performs better than Ala in replacing the Glu. Thus, the Glu at position 37 contributes to stability but at the same time is not critical for enzymatic activity, as mutants exhibit activity at a level similar to wild type BlaC. *In vivo*, however, Glu37 mutants confer no resistance against β-lactam antibiotics at all and the amount of soluble protein is very low. Clearly, Glu37 also has a critical role in protein folding. We propose that Glu37 is involved in the early stages of the folding process, through formation of an intermediate that facilitates formation of the central β-sheet. The sheet consists of two β-strands from stretches of amino acids located close the N-terminus and three that are close to the C-terminus (Figure 4.8). The multiple interactions of Glu37 with residues 42, 44, 60 and 61 could prime the formation of the β-sheet by stimulating the interactions between the first and second β-strand (in blue in Figure 4.8). *In vitro* the E37A variant is clearly hampered in refolding in the presence of low concentrations of Gdn. This result is an indication that Glu37 is relevant

for folding, but it is noted that cellular folding differs from *in vitro* refolding in the presence of a chaotropic agent. In prokaryotic cells, folding already starts on the ribosome[208,209], so formation of this interaction may happen very early on, even before the remaining part of the protein has emerged from the ribosome channel.

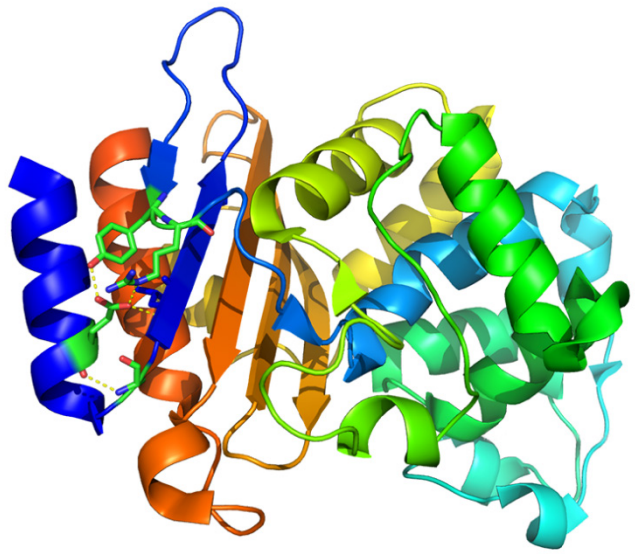

Figure 4.8. The potential role of Glu37 interactions for folding. The structure of BlaC (PDB 2GDN)[101] is shown in cartoon representation with rainbow colours, from blue (N-terminus) to orange (C-terminus). Glu37 and the residues to which it forms hydrogen bonds (dashed lines) are shown in sticks, with carbons, oxygens and nitrogens in green, red and blue, respectively. It is proposed that the interactions of Glu37 prime the formation of the two blue β-strands, required for formation of the β-sheet with the remaining three β-strands that are distant in the primary structure.

Residue Trp229 is 98% conserved in class A β-lactamases. It is located in a turn between α-helix 12 and β-sheet 3 (Figure S2.9) and, just like Glu37, is found more than 20 Å away from the active site Ser70. In BlaC Trp229 makes a backbone-to-backbone interaction with Pro226 and a number of stacking interactions, including a perpendicular π-stacking interaction with Phe225, shifted parallel π-stacking interaction with Tyr259, and a stack-like interaction with Pro254. Pro226 most likely interacts with Trp229 via an L-shaped CH-π interaction (Figure 4.9)[210–212]. These residues are not highly conserved in β-lactamases, with the exception of Pro226. Interestingly, even though β-lactamases are conserved structurally, the sequence as well as the precise structure of the loops differ considerably in this region. A consensus is that other aromatic residues play the same roles as Phe225 and Tyr259 in BlaC (Figure 4.9). Thus, in the folded protein, Trp229 facilitates staple interactions and occupies a large volume inside the hydrophobic pocket. Tryptophan has been shown to perform this role in multiple proteins[213,214].

For the class A β-lactamase TEM-1, it was reported that mutation of Trp229 to Ala, Phe or Tyr drastically decreases the yield but also reduces the activity of the enzyme[131]. Another aromatic residue at the position 229 can still form stacking interactions inside the hydrophobic pocket, although those with Pro226 and Pro254/252 are likely to be weaker, because the indol π-cloud is more electron rich[211]. Due to the smaller size of Phe and Tyr, the van der Waals interactions will be sub-optimal. Together, these effects can explain the slightly lower melting temperatures obtained for W229Y and W229F BlaC. The insolubility of three mutants and low yields of two Trp229 mutants coupled with their poor in vivo performance indicate that Trp229 is essential for mediating correct and efficient folding of BlaC. Aromatic mutants are able to confer resistance to β-lactam antibiotics at reduced growth temperatures (25 °C), indicating that the balance of misfolding and productive folding can easily be shifted in this case. Analogous to the Glu37 mutants, once folded, the mutant proteins show an overall fold and activity similar to those of the wild type BlaC. It is less obvious why Trp229 is critical for the folding process, because many of the surrounding residues are not conserved to a high degree and, unlike Glu37, Trp229 is not involved in stapling distant parts of the primary structure together. Trp229 is part of the large hydrophobic core of the protein formed by the three C-terminal β-strands that stabilizes the entire aβ domain of the β-lactamases and thus could play a role in ensuring a rapid hydrophobic collapse and formation of the molten globule state[215–217], which then evolves to the folded state.

The yield of soluble, folded protein did not improve by increasing the concentration of chaperones for either Glu37 or Trp229 mutants. Despite seeing a decrease in the amount of protein within the pellet, the yield of soluble protein did not increase. It is plausible that chaperones help misfolded BlaC to undergo proteolysis, a process which has been well documented in eukaryotic systems[218–220] and *E. coli*[221,222]. Apparently, the chaperones do not enhance correct folding.

In summary, we find that residues Glu37 and Trp229 are not only conserved but also essential residues. Mutations dramatically reduces the yield of soluble, folded protein. In the case of Glu37, this reduction is proposed to be attributable to a role in the early folding process by priming the formation of the central β-sheet, which consists of parts of the protein widely separated in the primary structure. Given it large size and hydrophobic and aromatic interactions, Trp229 could be involved in the hydrophobic collapse and the formation of the molten globular state. Interestingly, though in the folded state both residues are engaged in interactions that affect the thermostability to some degree, they are not critical for the overall structure and activity.

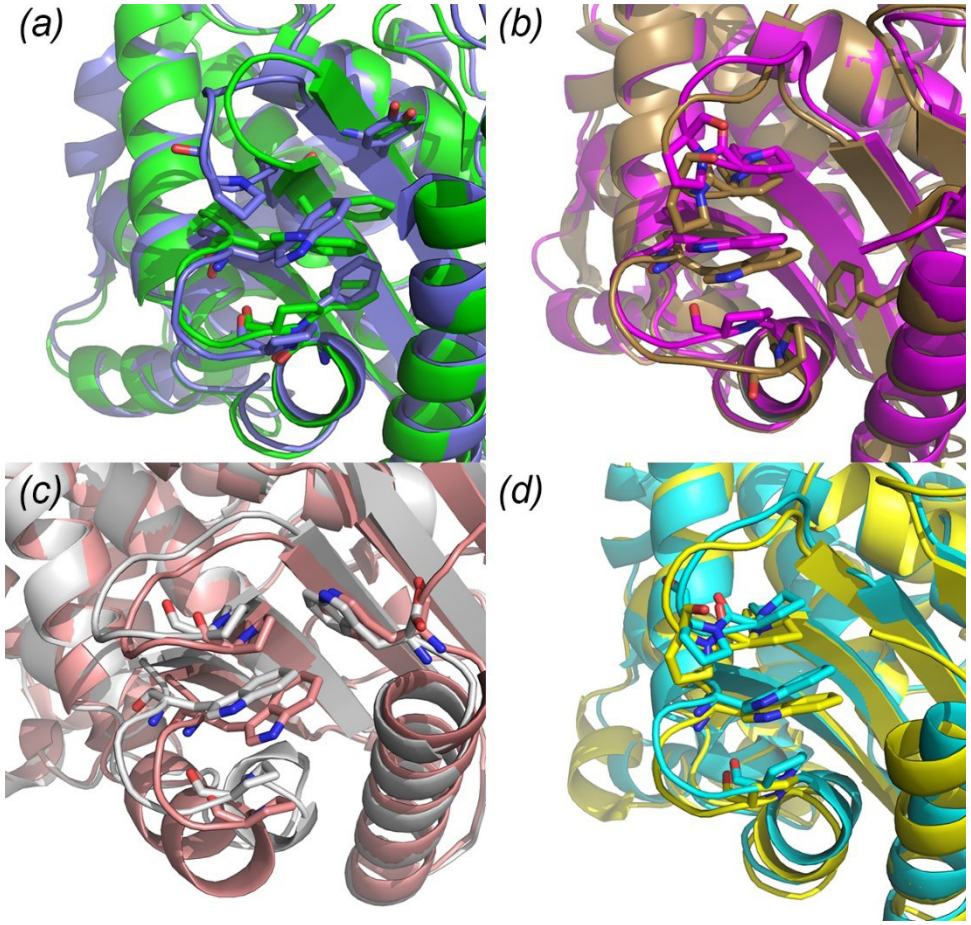

Figure 4.9. Structural overlay of some of the β-lactamase crystal structures, residues 229, 226, 254 and aromatic residues from hydrophobic pocket are represented with sticks. (A) Phe225-Pro226-Trp229-(Phe249)-(Pro254)-Tyr259 chain: BlaC - purple (2GDN[101]), GES1 from *Klebsiella pneumoniae* – green (2QPN[175]); (B) Pro226-Trp229-Pro251-Pro252-(Phe287) chain: β-lactamase from Burkholderia phymatum – gold (5VPQ), β-lactamase from *Mycobacterium abscessus* – magenta (4YFM[84]); (C) Pro226-Trp229-Pro252-Trp290 chain: TEM1 from *Escherichia coli* – white (1BTL[191]), SHV1 from *Klebsiella pneumoniae* – salmon (2ZD8[190]); (D) Pro226-Trp229-Pro251-Pro252 chain: β-lactamase from *Streptomyces albus* - yellow (1BSG), BS3 from *Bacillus licheniformis* – cyan (1I2S[206]).

Materials and methods

Production of BlaC variants

Site-directed mutagenesis was performed using whole plasmid site-directed mutagenesis. Plasmid pUK21 carrying the *blaC* gene, encoding a signal peptide for transmembrane transport or the plasmid pET28a+ carrying the *blaC* gene with the code for a N-terminal or C-terminal His(6)-tag was used as template. Sequences of the primers were designed to introduce single amino acid substitutions (Table S2.3). Mutated plasmids were transformed into competent *E. coli* KA797 cells. Colonies were selected on LB agar plates containing 50 μg mL^{-1} kanamycin. The presence of the mutations was confirmed by sequencing, performed by Baseclear BV.

For recombinant production of mutant proteins, *E. coli* strain BL21pLysS (DE3) was used in combination with the pET28 based plasmids. Cells were cultured in LB medium at 37˚C until the optical density at 600 nm reached 0.6, at which point protein production was induced with 1 mM isopropyl β-D-1-thiogalactopyranoside (IPTG), followed by incubation of the cultures at 18 ˚C for 16 hours. For the production of isotope labeled proteins for NMR experiments, M9 medium was used, instead of LB, with ^{15}N ammonium chloride as the sole nitrogen source. Components of the M9 medium can be found in Table S2.4. Proteins carried a N-terminal (Trp229 mutants) or C-terminal (Glu37 mutants) His-tag. The C-terminal His-tag was chosen for Glu37 mutants to avoid possible influence of the tag on folding due to the close proximity of mutation site to the N-terminus. Wild type protein was produced in both tag variants and there was no observable difference between the yields. Cells were harvested by centrifugation and lyzed with French Press in 20 mM Tris-HCl buffer, pH 7.5 with 500 mM NaCl. The first purification was conducted with 5 mL HisTrap Nickel column (GE Healthcare) in 20 mM Tris-HCl buffer, pH 7.5 with 500 mM NaCl, protein was eluted with the same buffer with 125 mM imidazole. N-terminal His-tags were removed by overnight incubation with 0.2 mg mL^{-1} His-tagged Tobacco Etch Virus (TEV) protease at 4˚C in 25 mM Tris-HCl buffer, pH 8.0 with 100 mM NaCl, 1 mM EDTA and 5 mM DTT followed by an additional HisTrap Nickel column purification to separate BlaC without His-tag from uncleaved BlaC and TEV protease. SDS-PAGE analysis revealed single bands of similar sizes for all variant enzymes. Protein concentration was determined by absorption at 280 nm, using the theoretical extinction coefficient 29 910 $M^{-1}cm^{-1}$ [223].

Protein quantification in cultures

Overnight bacterial cultures induced with 1 mM IPTG were normalized to the same optical density at 600 nm and treated either with SDS PAGE sample buffer or with lysis buffer (25 mM Tris-HCl, 100 mM NaCl, pH 8.0 with 1 mM EDTA and 1 mg mL^{-1} lysozyme) to separate soluble and insoluble fractions. Whole lysate, soluble and insoluble fraction samples were

analyzed using SDS PAGE. Gels were stained with Coomassie Brilliant Blue. The signal intensities were compared to each other and to samples with known protein concentrations using ImageLab software (BioRad).

In vivo activity studies

The survival of the *Escherichia coli* cells carrying pUK-based plasmids with wild type or mutant BlaC genes was tested on LB-agar plates with various concentrations of antibiotics. The sequence for the mature protein was coupled to a signal sequence for the Tat-system, because under physiological conditions BlaC is produced in the cytoplasm and translocated through the cell membrane[165]. All plates contained 50 μg mL^{-1} kanamycin and 1 mM IPTG. Cells were applied on the plates as 10 μL drops with OD_{600} values of 0.3, 0.03, 0.003 or 0.0003.

Circular dichroism

CD profiles were recorded using Jasco J-815 spectropolarimeter with a Peltier temperature controller (Jasco, MD). Measurements were performed in triplicate at 25˚C with 10-15 μM protein in 100 mM sodium phosphate buffer, pH 6.4. Spectra were acquired in 1 mm quartz cuvette at a scan rate of 50 nm/min and later normalized for concentration.

Thermal stability

Thermal stability of the proteins was analyzed by thermal shift assay (TSA) with SYPRO Orange dye (Invitrogen). The measurements were performed in triplicate in two independent experiments using the CFX 96 Touch Real-Time PCR Detection System from Bio Rad with 2x dye and 10 μM protein in 100 mM sodium phosphate buffer, pH 6.4 with the temperature range 20-80 ˚C. Melting temperatures were determined as an average of 6 measurements with standard deviation.

Kinetics

Determination of the Michaelis-Menten kinetic constants was done by measuring the absorption change at 486 nm for nitrocefin ($\Delta\varepsilon = 11300\ M^{-1}cm^{-1}$) in a PerkinElmer Lambda 800 UV-Vis spectrometer at 25 ˚C in 100 mM sodium phosphate buffer, pH 6.4; all measurements were performed in triplicate. The Δε was calculated by measuring a dilution series, which yielded a calibration line of absorption against concentration, with the concentration of the stock determined by quantitative NMR against 2.0 mM trimethylsilylpropanoic acid (TSP). The reactions were carried out at various concentrations of substrates and 2 nM or 5 nM concentration of BlaC, and initial rates of the hydrolysis were plotted against concentration of substrate and fitted separately for each experimental set to the Michaelis-Menten equation 4.1 using OriginPro 9.1, where v_0 is the initial reaction rate, $[S]_0$ the initial substrate concentration, V_{max} the maximum reaction rate and K_m the Michaelis-Menten constant. v_0 and $[S]_0$ are the dependent and independent variables, respectively, and K_m and V_{max} are the fitted parameters. $V_{max} = k_{cat}[E]$, k_{cat} is the rate-limiting conversion rate, and $[E]$ is the concentration of BlaC.

$$v_0 = \frac{V_{max}[S]_0}{[S]_0 + K_m} \qquad (4.1)$$

Refolding experiments

For refolding experiments initial unfolding of BlaC proteins was done using thermal denaturation and chemical denaturation. Thermal denaturation was found to be irreversible in analyzed temperature range (data not shown). For chemical denaturation 45 μM of BlaC WT and mutants were incubated for 20 minutes on ice with 4.5 M guanidinium chloride in 100 mM phosphate buffer (pH 6.4). Refolding was than initiated by dilution of samples with the same buffer containing 2.25 M, 1.8 M, 1.125 M, 0.9 M or 0.45 M guanidinium chloride with 100 μM nitrocefin. Untreated BlaC in 100 mM phosphate buffer was used as a positive control. The hydrolysis of nitrocefin was measured at 486 nm on a TECAN Infinite® M1000PRO plate reader. In each measurement the final BlaC concentration was 0.9 μM. Experiments were carried out in triplicate.

NMR Spectroscopy experiments

TROSY-HSQC spectra were recorded on a Bruker AVIII HD 850 MHz spectrometer at 25˚C in 100 mM phosphate buffer (pH 6.4) with 6% D_2O. Data were processed in Topspin 3.2 (Bruker). Spectra were analyzed with CCPNmr Analysis software. Peaks of the mutant spectra were assigned by comparison to peaks in the wild type BlaC spectrum and average chemical shift perturbations (CSP), $\Delta\delta$, of the ^1H ($\Delta\omega_1$) and ^{15}N ($\Delta\omega_2$) resonances of backbone amides were calculated using equation 4.2. Peaks that could not be assigned with certainty were assigned based on the smallest possible CSP (Figure S4.5).

$$\Delta\delta = \sqrt{\frac{1}{2}\left(\Delta\omega_1^2 + \left(\frac{\Delta\omega_2}{5}\right)^2\right)} \qquad (4.2)$$

Crystallization

Crystallization conditions for BlaC at a concentration of 10-15 mg mL^{-1} were screened for by the sitting-drop method using the JCSG+, PACT premier (Molecular Dimensions) and Index HT (Hampton Research) screens at 20 ˚C with 200 nL drops with 1:1 protein to screening condition ratio[224]. Crystals for BlaC E37A grew within four days in 0.1 M sodium acetate buffer pH 5, 0.2 M magnesium chloride and 20% w/v PEG 6K as precipitant). The crystals were mounted on cryoloops in mother liquor and 30% glycerol as cryo-protectant and vitrified by plunging in liquid nitrogen.

X-ray data collection, processing and structure solving

Diffraction data to 1.5 Å resolution of BlaC E37A crystals were collected at the European Synchrotron Radiation Facility on the MASSIF beamline[225]. The data were integrated using XDS[226] and scaled using Aimless[227] and data to 1.5 Å resolution were kept for structure solution and refinement. Overall completeness of this dataset is somewhat low (91.5%) due to the presence of ice rings. The structures were solved by molecular replacement using MOLREP[228] from the CCP4 suite[228] using PDB entry 2GDN[101] as a search model. Subsequently, building and refinement were performed using Coot and REFMAC[228]. Waters were added in REFMAC during refinement; we have modeled double conformation for residues N110, R128, D240 and V263. The final model falls on the 100th percentile of MolProbity[229] and showed that 98.85% of all residues were within the Ramachandran plot favoured regions with two outliers. The model was further optimized using the PDB-REDO webserver[230,231]. Data collection and refinement statistics can be found in Table 4.3. The data have been deposited in the Protein Data Bank, entry 7A5U.

Supporting Information

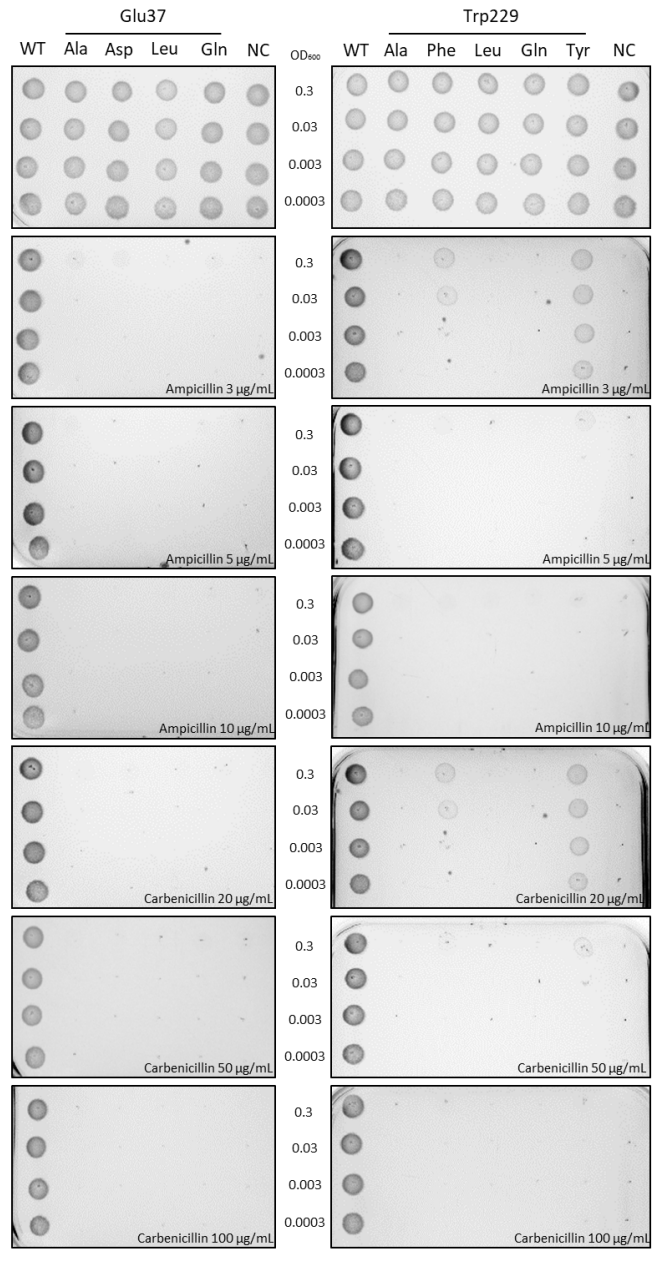

Figure S4.1. Plates showing growth of *E. coli* cells expressing wild type (WT) or mutant proteins (negative control, NC is the S70A mutant) with no antibiotics or various concentrations of antibiotics. Cultures were applied as the drops with OD shown in between the plates.

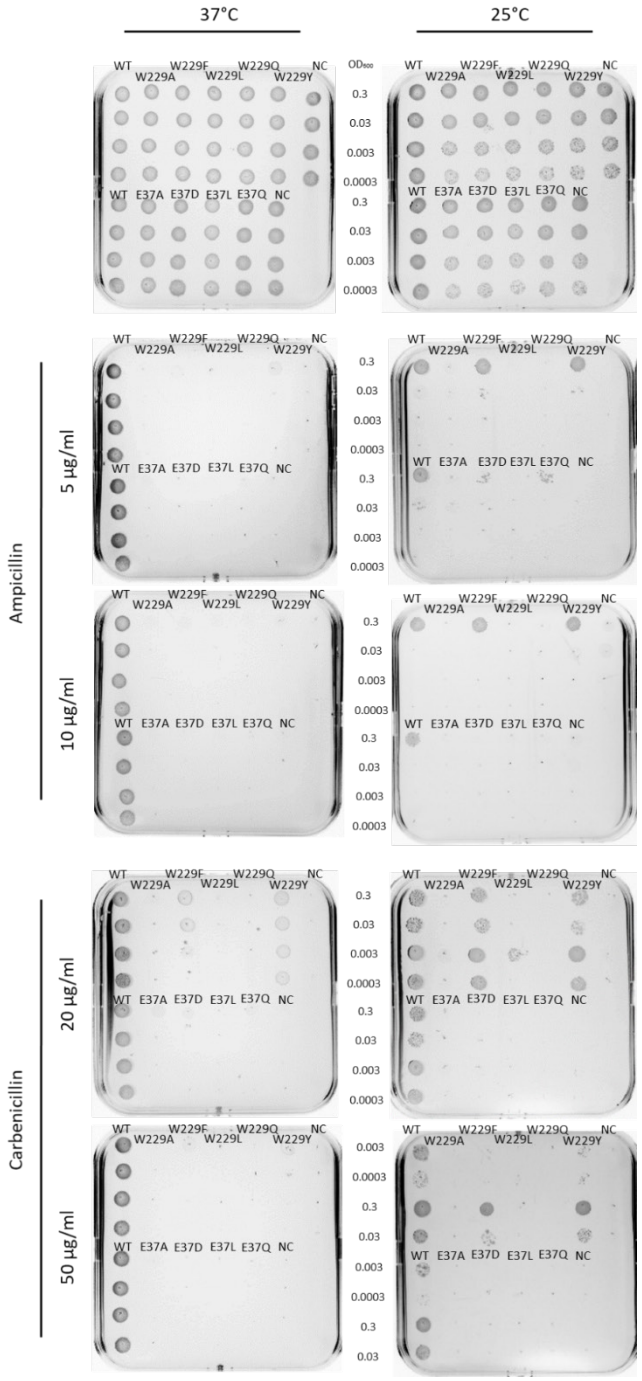

Figure S4.2. Plates with ampicillin or carbenicillin incubated at 37˚C or 25˚C. Cultures producing wild type (WT) or mutant BlaC (negative control, NC is the S70A mutant) were applied as the drops with OD shown in between the plates. For W229F and W229Y BlaC growth is better at 25 ˚C.

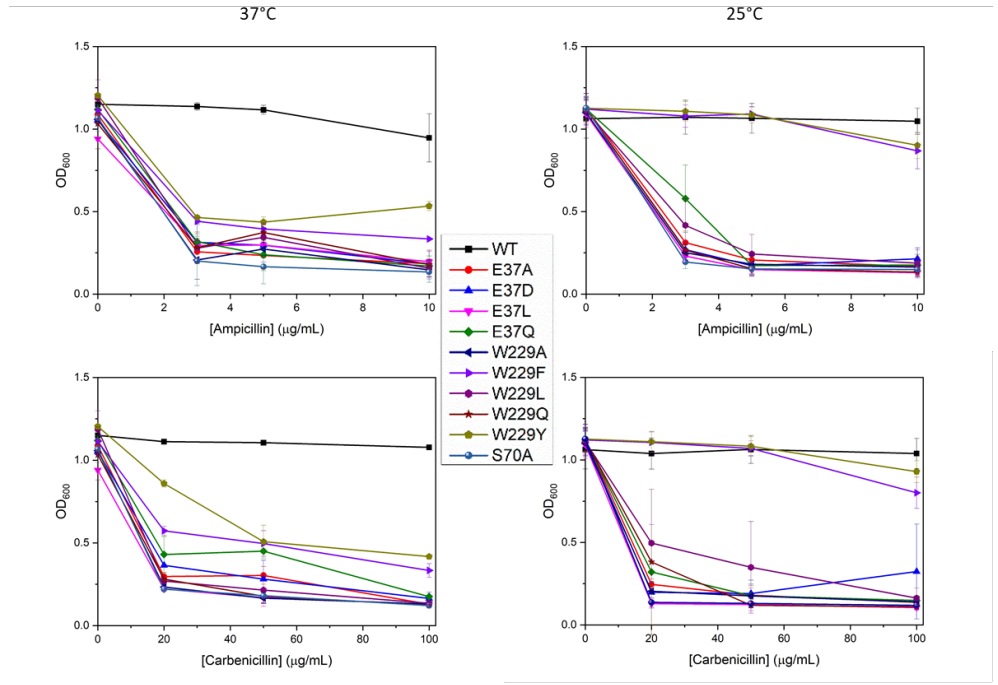

Figure S4.3. The end point (15 hours at 37 ˚C or 25 hours at 25 ˚C) of OD600 of wild type and mutant cultures growth with various concentrations of antibiotics. For W229F and W229Y BlaC growth is better at 25˚C. Error bars represent the standard deviation of three measurements.

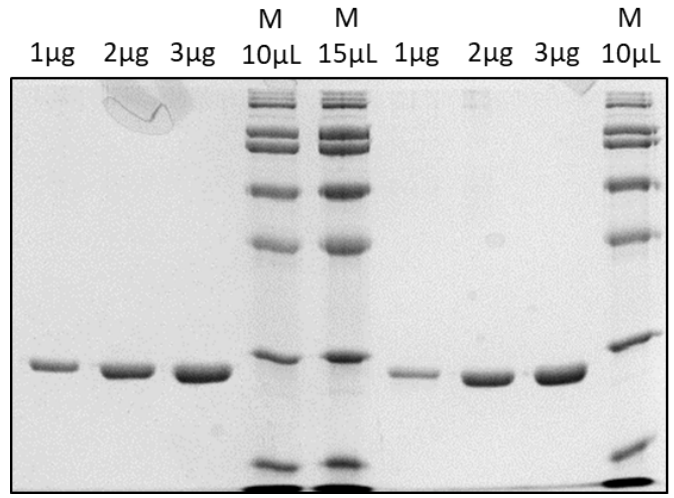

Figure S4.4. SDS PA-gel for determination of protein amounts in Marker (BioRad). Purified BlaC with known concentration was used in quantities shown on the figure. Analysis of the band intensities was performed with ImageLab software (BioRad).

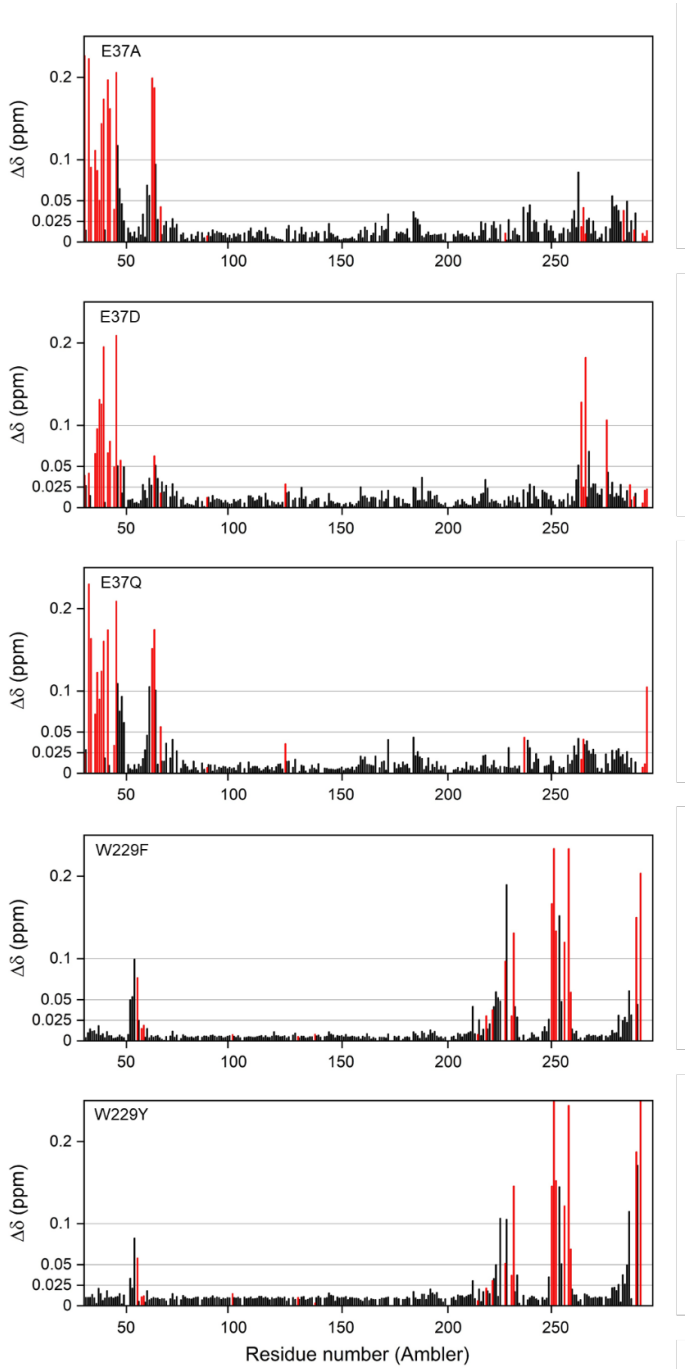

Figure S4.5. Average CSP for the backbone amides of the BlaC mutants. Residue numbers refer to Ambler numbering [89]. Black bars represent CSP for residues assigned with certainty. Red bars show CSP calculated as the minimal possible CSP.