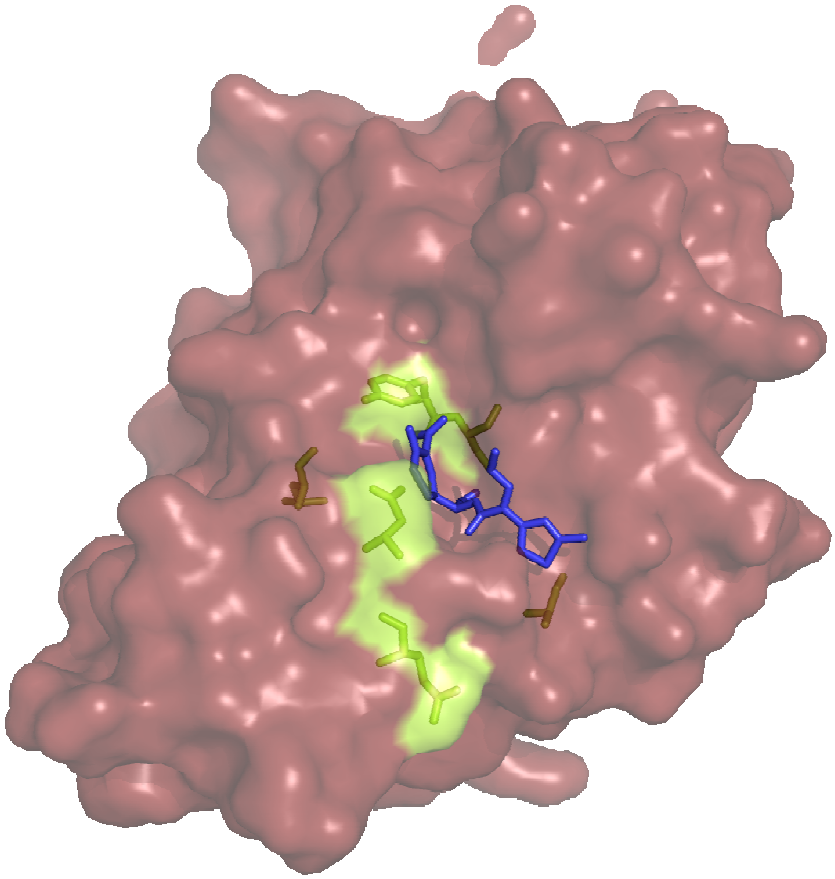
The evolutionary accessibility of a computationally designed TEM β-lactamase mutant



Joseline Houwman – 881022369110

Supervisors: Merijn Salverda and Arjan de Visser

Laboratory of Genetics

May 2010

Master Thesis Genetics – GEN-804039

Wageningen University, Wageningen

Cover illustration: The 3D structure of the Hayes TEM β-lactamase, with cefotaxime (blue) modeled in the active site and with an estimation of the possible conformations of the Hayes mutations (yellow).

Abstract

Fitness landscapes are visual representations of the fitness of different genes or genotypes in a certain environment. The highest fitness is represented by a peak in the landscape, while deleterious mutations are dips or valleys. The fitness landscape of the antibiotic degrading enzyme TEM-1 β-lactamase contains at least one peak when drawn for the antibiotic cefotaxime. This peak is caused by a natural occurring mutant, but in recent years much research has been done on other mutations that increase TEM-1's resistance to cefotaxime. One of these studies resulted in the Hayes mutant, where all but one of the six mutations were predicted by computational methods. This thesis deals with the probability of the natural occurrence of the Hayes mutant.

By testing the resistance to cefotaxime of the single and double mutants, it was found that while the Hayes mutant does represent a peak in the TEM-1 fitness landscape of cefotaxime, this peak is unconnected with the other large peak. Also the peak is very steep, as the single and double mutants of the Hayes mutant do not confer an increase in resistance to cefotaxime. Thus it seems to be the case that several neutral mutations have to become fixed, before any increase in fitness can be observed. Unfortunately, resolving the amount of neutral mutations necessary before ascending the top was beyond the scope of this thesis

As several double and one triple base pair changes are required for some of the mutations, it is unlikely that the Hayes mutant will ever be found in nature. Yet under the influence of genetic drift and fluctuating environments, there is still a small chance that this mutant might be isolated in the future.

Contents

Introduction

Fitness landscapes and sequence space

"Mountain ranges representing condensed areas of high fitness, small peaks in which a protein can become trapped, ridges connecting ranges, and the elusive Everest, representing the globally fittest protein"! Whilst reading this quotation from Voigt *et al*. (*1*), it may seem as if you have inadvertently recovered a paper on geology instead of evolution, yet actually it is a description of one of the oldest ideas in evolutionary biology: the fitness landscape (*2*). First proposed by Sewall Wright in 1932 (*3*) and nowadays widely used (*4*, *5*), these landscapes essentially provide a way to visualize the effect of mutations on the 'fitness' of a single gene or genotype. The underlying grid pattern (the x-y plane) of the fitness landscapes is the 'sequence space' of the protein, a concept proposed by Maynard Smith in 1970 (*6*). The sequence space is an arrangement of all possible amino-acid sequences in such a way that all neighbours can be converted into each other by one single substitution. By placing the ancestor gene and all observed mutants in this grid pattern, while the z-axis represents the fitness of the mutants with respect to the ancestor, one acquires a 'map' in which one can see at once what effect a certain mutation (or a combination thereof) would have on the fitness of the protein. This three dimensional approach is only applicable to the fitness landscape of a single gene, as in the genotype fitness landscape every single loci has its own axis, thus arriving at a multi-dimensional fitness landscape (*5*).

While most of the older work on fitness landscapes was theoretical (*5*), in recent years an increase can be observed in empirical studies to determine the shape of fitness landscapes for certain model organisms and proteins (*7*-*9*). These empirical fitness landscapes have been especially useful to show the effect of epistasis between mutations.

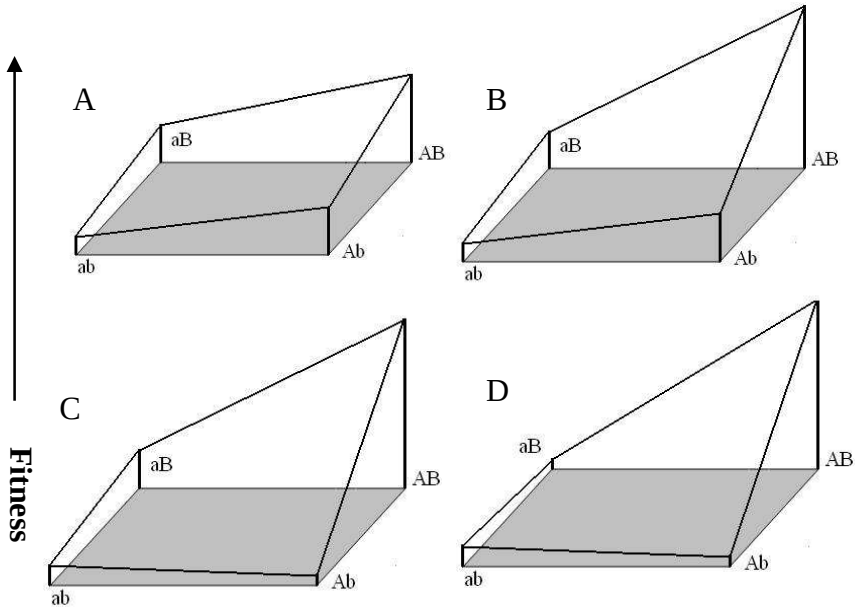


Figure 1 The effects of epistatic interactions on two loci with two alleles (A and B), starting from genotype ab (After (8)). The two upper pictures show the effect of no epistasis (A) and magnitude epistasis (B), while the lower two pictures show the effect of sign epistasis (C) and reciprocal sign epistasis (D).

In the case of no epistasis, the potential effect of a mutation on the fitness is independent of the genetic background in which it occurs. For epistasis however, the effect of the mutation of **a** to **A** and **b** to **B** is dependent on the occurrence of earlier mutations in the genetic background. This effect can be either positive or negative, see Figure 1. In the case of magnitude epistasis the change of **a** to **A** is positive in both cases, yet in the background of **B**, the effect is much larger than in the **b** background. For sign epistasis, the **a** to **A** mutation is positive in the background of **B**, but negative in the background of **b**. The last form of epistasis is reciprocal sign epistasis, where the mutation that causes the intermediate genotypes **Ab** and **aB** has a negative effect on fitness, while both second mutations increase the fitness again. This reciprocal sign epistasis is necessary for the occurrence of multiple peaks in a fitness landscape, though it does not guarantee it (*8*).

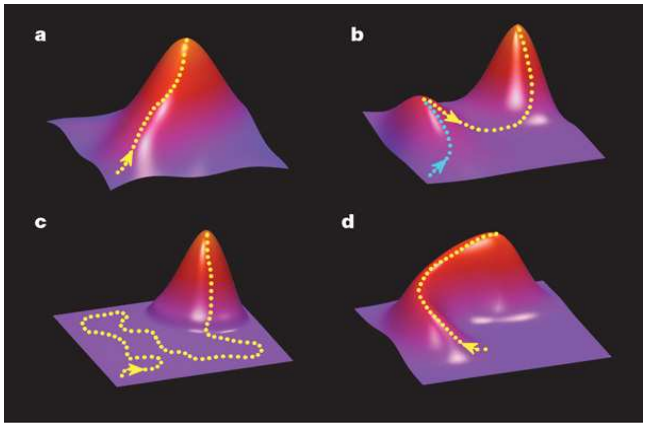


Figure 2 Different fitness landscape features: ***a****, a single smooth peak.* ***b****, a rugged landscape with two peaks, of which the smaller is an evolutionary trap.* ***c****, a neutral landscape.* ***d****, a reversion of original mutations leads to a detour landscape. From (8)*

This possible consequence of reciprocal sign epistasis, the emergence of multiple peaks on the fitness landscape, has great implications for the course of evolution. In the case of a single peak in the landscape, evolution would cause a steady increase in fitness in the population, until the top of the peak is reached. In a rugged landscape however, the 'danger' exists that a population becomes trapped on a local peak and can not escape unless it decreases its fitness, even when there is a higher peak in the neighbourhood (*2*) (Figure 2b). This trapping of populations at local fitness peaks can have a severe effect on the accessibility of evolutionary pathways, yet populations can escape these local peaks by reversing their original mutations (*10*) (Figure 2d). What complicates the research in sign epistasis is the fact that when not all fitness effects of all mutational neighbours in the fitness landscape are known, it is possible that a third mutation might relax or invert the original reciprocal sign epistasis, thus forming a 'ridge' between the two peaks (*5, 8*)(Figure 3).

While determining all mutational neighbours in a fitness landscape is one of the biggest challenges relating to the drawing of fitness landscapes, there are two other complicating factors: The high-dimensionality of genotype fitness landscapes and the specific environment for which most landscape are made.
As discussed in the first paragraph of this Introduction, when making a fitness landscape for an entire genotype, every loci of the organism has its own axis, thus making the fitness landscape multi-dimensional. As most biologists would rather not start thinking like theoretical physicists on string theory, instead of trying to understand the whole of the organism, single genes have been extensively used as models to try and understand the basics of evolutionary adaptation (*4, 8*). This is of course a rather simplistic approach, but with a better understanding of the mechanisms of evolution, research can start to focus on the evolutionary history of pathways and networks made out of multiple genes (*11*).

The other factor complicating the construction of fitness landscapes is the fact that it is actually not a 'rigid' mountain landscape, but more of a 'turbulent ocean' (*12*). This turbulence arises due the changing of the environment to which the organism or the gene has adapted. As fitness landscapes are usually designed for a specific environment, they are only applicable for those specific situations (*2*). This is very nicely illustrated by the fitness landscapes of the antibiotic resistance enzyme TEM-1 β-lactamase. When TEM-1 evolved the ability to degrade the 'new' antibiotic cefotaxime, it did this at the cost of the enzyme's ancestral ability to degrade the 'older' antibiotic ampicillin (*13*). Thus a peak in the enzyme's fitness landscape for cefotaxime, does not necessarily coincide with a peak in the ampicillin fitness landscape. In reality it is even the case that mutations which transfer a higher cefotaxime degrading ability usually decrease the capability of the β-lactamase to degrade ampicillin (*13*). Thus, what in one environment is a peak, might be a ridge in another environment or even a valley.

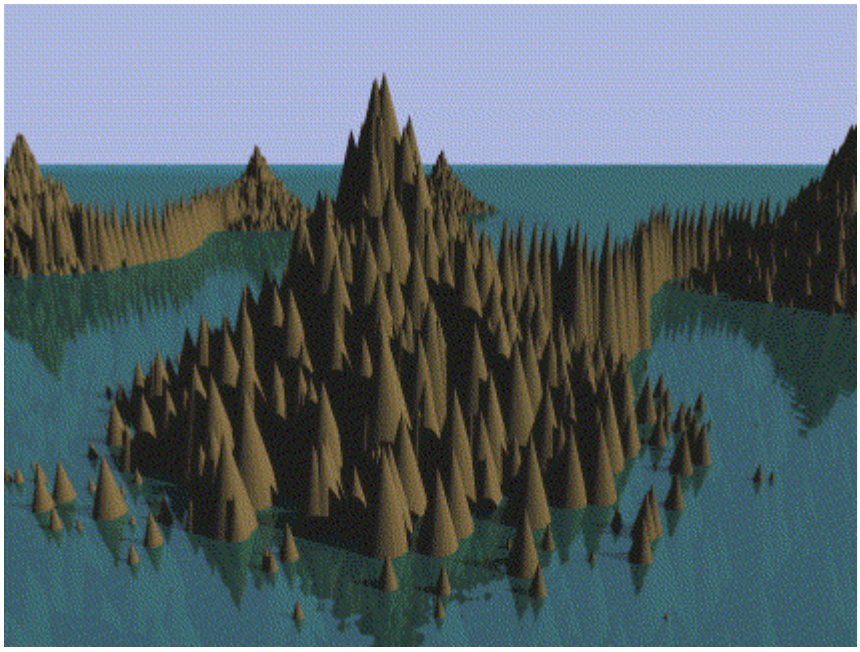


Figure 3 Seascape representation of fitness. Sea level represents minimum biological activity (i.e. fitness of the ancestor), while the land elevation is a measure for fitness increase due to mutations. Several 'islands' of fitness can be seen, joined by ridges. From (68)

Nevertheless, before these complications factors can be addressed, first basic research into the form of fitness landscapes and sequence spaces of proteins and the rules that govern these spaces has to be carried out. One of the evolutionary models that is used to explore the accessibility of sequence space, is the Strong Selection-Weak Mutation (SSWM) model by Gillespie (*14*). This model poses that overall mutation rates are low as well as that neutral and deleterious mutations have low probabilities of being fixed and can thus be ignored in natural selection. This then leads to the sequential fixation of beneficial mutations which sweep through the population and form a pathway from the lowest to the highest fitness. In recent years some debate has been going on, on whether the SSWM model is really the prevalent mode of selection, or whether the Weak Selection–Strong Mutation or the Strong Selection–Strong Mutation models are more accurate representations (*15*).

TEM-1 β-lactamase

A big problem of the modern medical world is the emergence of more and more antibiotic resistant bacteria. This problem is almost as old as the use of antibiotics itself (*9*) and is particularly pronounced in the case of TEM-1. Only two years after the introduction of the antibiotic ampicillin in 1961, the β-lactamase enzyme TEM-1, which confers resistance to

ampicillin, was isolated in Greece (*16*). β-lactam based antibiotics (see Figure 4) are one of the most widely used species of antibiotics, as they interfere with the biosynthesis of the bacterial cell wall (*17*). In response, β-lactamases have evolved which hydrolyze the β-lactam ring of the antibiotic, thus rendering it inactive.

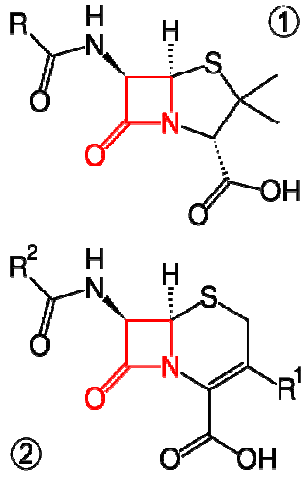


Figure 4 The core structure of penicillins (1) and cephalosporins (2). The β-lactam ring is shown in red. From the Wikipedia article Beta-lactam antibiotic.

Since bacteria steadily became more and more resistant to the 'older' antibiotics like penicillin and ampicillin, new β-lactam antibiotics were synthesized, the so-called cephalosporins. This was the start of the 'β-lactamase cycle' (*18*), in which pharmaceutical companies would produce new β-lactam antibiotics, while within a few years after their introduction, β-lactamases that could degrade these new antibiotics would emerge. These β-lactams are called extended spectrum β-lactamases (ESBLs), as they confer resistance to penicillins and cephalosporins. Even inhibitors of β-lactamases like clavulanic acid were not safe for this evolutionary arms race (*9, 19*). This increase in resistance to a particular antibiotic usually only requires a few, or sometimes as little as one single point mutation, to become active. Nowadays, TEM-1 occurs in almost 180 natural varieties[1] and many more mutants have been created during *in vitro* evolution experiments.

As this wide variety of mutants increases the knowledge of what possible effects the mutations would have on resistance and because of the fact that selecting for antibiotic resistance is relatively easy, TEM-1 has been used as a model system in biochemical engineering since 1976 (for a summarizing table see (*9*)). One emphasis of these studies lies on the ability of TEM-1 to degrade the 3rd-generation cephalosporin, cefotaxime. Originally as with most β-lactam antibiotics, TEM-1 was not able to catalyze the degradation of cefotaxime efficiently, resulting in a very low Minimal Inhibitory Concentration (MIC) against cefotaxime. MIC is defined as the lowest concentration of antibiotic that prevents visible growth of the bacteria and hence is used as a measure of resistance. In 1994 however, while using TEM-1 as a model for the new technique of DNA shuffling, Stemmer isolated a mutant with a ~30,000-fold increase in MIC against cefotaxime (*20*).

This mutant g4205c:A42G:G92S:E104K:M182T:G238S:R241H (numbering as in (*21*)), has since then been extensively studied and two important studies have been done regarding the potential evolutionary pathways to this mutant and its fitness landscape (*7, 22*). The first of these two studies was based on the assumption that only a succession of single beneficial mutations can lead to the top of a fitness peak (Strong Selection – Weak Mutation (*14*)). By building all single mutants, testing them for their MICs, selecting the highest one and building the remaining mutations in that background, etc. etc, Hall (*22*) showed that there was a pathway with steadily increasing MIC from TEM-1 and consequently presumably accessible by nature. Unexpectedly though, when eliminating mutation G92S and R241H, the mutant acquired a MIC against cefotaxime that was higher than reported by Stemmer (*20*). The remaining mutations have all been observed in clinical isolates and therefore a second study was done by Weinreich *et al.* (*7*) aiming to explore, whether there might be more pathways

[1] http://www.lahey.org/Studies/temtable.asp

than just the one involving the best intermediates. They showed that of the 120 possible pathways, only 18 were accessible by single mutation steps, with intermediates that had a steady increase of MIC, so no neutral or deleterious mutations. A follow-up study showed that if one allows reversions of original mutations, nine additional pathways can be followed (*10*). This obstruction of more than three quarters of all possible pathways to the fittest mutant is caused by sign epistasis, which in recent years has become a main focus in β-lactamase research, as has the study of fitness landscapes in general (*8, 9, 23*).

The search for new interesting techniques and mutants still continues to this day and in recent years the rise of faster computers has led to computational tools being used to 'predict' what effect mutations would have in advance. One of these tools is Protein Design Automation (PDA), a computational design tool which starts with the three-dimensional structure of the protein to be designed and then predicts the optimal sequence for that fold, allowing all or only a specified set of residues in the protein to change (*24*). The use of PDA in protein optimalization had a remarkable outcome in 2002. Combining PDA and experimental screening[1], Hayes *et al.* created a mutant of TEM-1 with some unusual characteristics (*25*). The mutant had a MIC which was only a eight-fold increment lower than the MIC of the Stemmer mutant (*20*) and none of its mutations have been observed in natural isolates before.

How did Hayes *et al*. make such a novel and interesting mutant? First, they focused their efforts on the active site area, selecting 19 positions that were allowed to change into all possible amino acids with the exception of cysteine and proline. This resulted in $7x10^{23}$ sequences, of which global energy of the folded protein was calculated by PDA and the 1000 sequences with the lowest global energy were screened for the occurrence of the different amino acids. These occurrences were counted and a cut-off of 10% was applied, that is if an amino acid occurred at a specific position in more than 10% of the 1000 sequences, it was included in the mutant library. This cut-off resulted in 9 mutated positions and in 172,800 unique mutant sequences, which were constructed using recursive PCR with overlapping oligonucleotides. This PCR was either error-proof, employing *Pfu* DNA polymerase, or error-prone, employing a mix of *Pfu* and *Taq* DNA polymerase. Libraries were then made and plated on increasing concentrations of cefotaxime to test the MIC of the clones. Clones from the error-proof library had a 640-fold increase in MIC against cefotaxime and clones from the error-prone library showed a 1280-fold increase. Three clones were then isolated, sequenced and several characteristics determined. Of these three mutants the subject of this thesis, PDA-2, contains six amino acid changes of which three require single point mutations, two necessitate double point mutations and one needs the entire codon to change with a triple point mutation.

Questions addressed in this thesis

Because of the fascinating characteristics of the Hayes mutant, the mutations which have never before been found in natural or laboratory isolates and the very high resistance to cefotaxime, an interesting evolutionary question can be asked:

Is the Hayes mutant evolutionary accessible?

[1] Patent number: US 6,403,312 B1

The easiest way to test this is to construct the fitness landscape for the Hayes mutant. This can be done by creating the intermediate steps of the mutant following the method described by Hall (*22*) and assuming that the SSWM model governs the emergence of new mutants. This leads to two possibilities: either there is a succession of fitness increasing intermediates or all intermediates are deleterious, neutral or only slightly beneficial. This then leads to the sub question:

Is the Hayes mutant accessible under SSWM conditions?

If there are intermediates which show a steady increase in MIC to cefotaxime, then it is only a matter of time before the Hayes mutant will be encountered in clinical isolates. However Hayes *et al.* stated that all single mutants did not seem to increase the fitness (*25*), so if the double mutants also show no increase in fitness then the SSWM model would not be able to explain the possible emergence of the Hayes mutant in nature. As the SSWM model is not the only evolutionary model, another sub question is:

Through what other evolutionary models might the Hayes mutant be accessible?

And finally, if all the mutations in the Hayes mutant are neutral or only slightly beneficial, it is very likely that adaptation will 'leap-frog' the moderately beneficial mutations, to arrive at strongly beneficial ones (*4*). Thus the single mutants or even the double mutants may revert to pathways that provide a larger fitness benefit in the beginning, such as pathways that include mutations G238S or R164S. Therefore my last research question is:

Will the (intermediates of the) Hayes mutant revert to more traditional pathways of resistance mutations?

Materials and methods

In the experiments described below, the gene TEM-1 is located on a plasmid and controlled with a promoter induced by isopropyl-β-D-thiogalactopyranoside (IPTG). A concentration of 50 ug/mL IPTG was added to solutions containing either ampicillin or cefotaxime, as this concentration creates expression levels of TEM which are comparable to expression levels in nature (*26*).

Bacterial strains and plasmids
Escherichia coli strain DH5αE (Invitrogen) was employed as host for all plasmids. For vector preparation, *E. coli* strain GM119 was used. Strains were grown at 37ºC and for liquid media at 220 rpm. Plasmid pACSE3 (*26*) was used as the host for cloning and expressing TEM-alleles. pACTEM is pACSE3 containing the TEM-1 allele from the plasmid pBR322 (*27*).

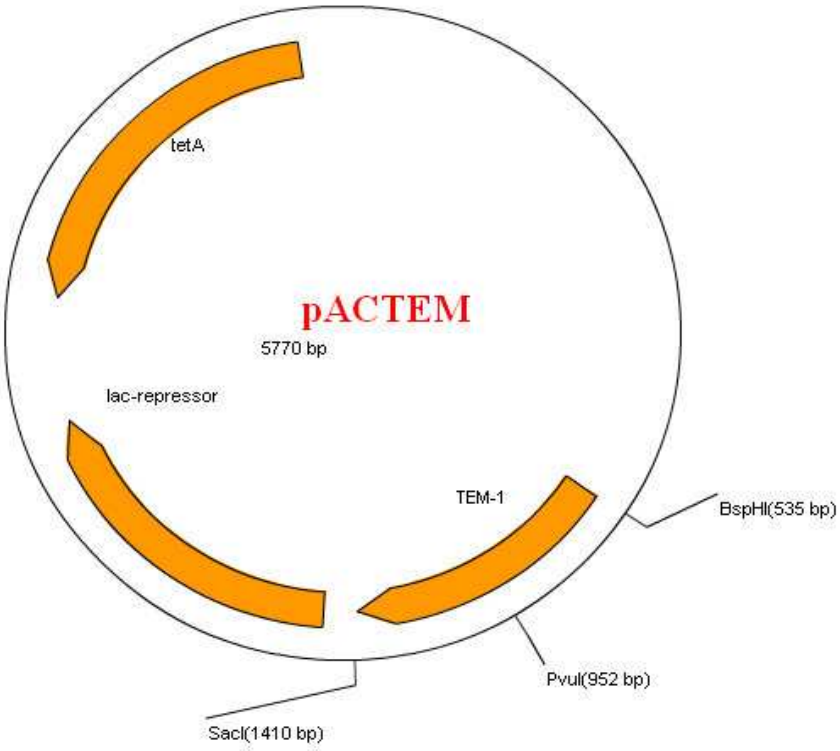


Figure 5 The pACTEM plasmid with important restriction sites. Made with EZ PLASMID MAP V1.9.

Media
Liquid broth (LB) was prepared by diluting 10 g trypticase peptone (tryptone), 5 g NaCl and 5 g yeast extract per liter demi-water. Mueller Hinton (Merck) and Mueller Hinton II (BD) broth were prepared according to manufacturers' instructions, in demi-water and MQ respectively. SOC-medium was made by adding 20 g tryptone and 5 g yeast-extract to 1 liter H_2O with 10 mM NaCl, 2.5 mM KCl, 10 mM $MgCl_2$ and 20 mM glucose. Solid media were made by adding 15 g of agar per liter. LB containing either tetracycline or ampicillin (LB+tet or LB+amp), was made by adding 15 μg/ml and 100 μg/ml respectively.

Antibiotics
Stock solutions of tetracycline, ampicillin and cefotaxime (all Sigma-Aldrich) were prepared in 70% ethanol, H_2O and 0.1 M $NaPO_4$, respectively. When either cefotaxime or ampicillin was included in the media, also IPTG was added to the solutions in a concentration of 50 μg/mL.

Vector preparation
pACSE3 was grown overnight (O/N) in LB+tet and purified with the GenElute™ Miniprep Kit (Sigma-Aldrich). Vector was prepared by digestion of pACSE3 with *BspHI*, *SacI* and dephosphorylated with Calf Intestinal Phosphatase (CIP)(all New England Biolabs).

Sequencing
The sequencing was done by Eurofins MWG Operon, using primers P3 and P4.

Site-directed mutagenesis
As we did not have access to the original sequence of Hayes *et al.*(*25*), the codon with the highest usage bias[1] was used for creating the double and triple point mutants (see Appendix I - Primers). Because of time limitations, the single point mutation intermediates of the double and triple mutations were not made.
Point mutations were introduced in the TEM-1 allele using the methodology described in Stratgene's QuikChange Site-Directed Mutagenesis Kit. PCRs were run using Pfu-polymerase (Fermentas) and the PCR program was adjusted to 16 cycles and an extension time of 6 min. For the construction of the single mutants the pACTEM plasmid was used as template in concentrations of 5 and 50 ng/μL. For the construction of TEM-alleles containing multiple mutations, PCRs were done on pACTEM already containing one or more of the desired mutations (see Appendix II). Primers can be found in Appendix I.
After transformation into DH5αE cells were grown for 24 h on LB-agar+tet. In order to eliminate possible mutations in the rest of the plasmid, several colonies were used as template in a PCR with the primers P3 and P4 (*26*), employing the following PCR-scheme: 1 cycle at 95º for 5 min, the 33 cycles of denaturation at 95º for 30 s, annealing at 61º for 30 s, extension at 72º for 75 s, followed by 1 cycle of 72º for 10 min. Successful incorporation of the desired mutations was verified by sequencing of the PCR-products. Subsequently, PCR-products containing the desired mutation were digested with *BspHI* and *SacI* (NEB), whose recognition sites flank the TEM-allele. After this digestion, the product was purified and religated O/N at 4ºC into naïve vector with T4 DNA ligase (Promega). The purified products were then transformed into DH5αE and plated on LB-agar+tet. The following day, colonies of each mutation were picked up and grown O/N for plasmid isolation and to obtain a glycerol stock of the pure cultures.
Purifications of PCR-, digestion and ligation products were done with the GenElute™ PCR Clean-up kit (Sigma-Aldrich), unless otherwise specified.

***In vitro* evolution**
To produce randomly mutated alleles of the mutant V103Q, error-prone PCR was done on pACTEM containing the V103Q substitution in five-fold replicates, according to the manufacturers' instructions of the GeneMorph® II Random Mutagenesis Kit (Stratagene). The conditions were set to introduce an average of 6.5 mutations per amplicon, with P3 and P4 as primers, an annealing temperature of 60º and an extension time of 75 s.
The resulting amplicons were purified with the GenElute™ PCR Clean-up kit (Sigma-Aldrich) and digested with *SacI*, *BspHI* and *DpnI* (NEB), and afterwards purified again. The samples were then ligated O/N at 4ºC into naïve vector with T4 DNA ligase (NEB). The next day, the ligase was heat inactivated and the ligation products were purified by isopropanol precipitation. The samples were then transformed into DH5αE and after a recovery period,

[1] The Codon Usage Database: http://www.kazusa.or.jp/codon/

they were transferred to 1 L bottles containing 500 mL LB+tet. Immediately after the transfer, 50 µL was plated on LB-agar+tet, to determine the library size. By growing the remainder of the culture O/N, the library was expanded. A glycerol stock of this enriched library was made and stored at -80ºC.

Selection of mutants

The following day, a series of bottles containing 50 mL of Mueller-Hinton and a series of cefotaxime (ctx) concentrations was used to select the mutated TEM-alleles with the highest resistance to ctx.

The range of ctx concentrations varied from 0.015 to 32 µg/mL (with a two-fold increase in concentration between bottles). These bottles were inoculated with a sample from the O/N enriched library, containing 10 times the determined library size and incubated for 48 h. After this period, the bottles with visible growth at the highest ctx concentration were selected. From each of these cultures a sample was taken, plated on LB-agar+tet and grown O/N. A single colony of each plate was picked and grown O/N in LB+tet for plasmid isolation and a pure culture glycerol stock.

MIC-assays

MIC is defined as the lowest concentration of antibiotic that prevents visible growth of the bacteria. To exclude phenotypic variation in the host, DH5αE clones containing naïve TEM-alleles were used in the MIC-assays. Roughly 1.5 $x10^4$ cells were grown in 300 µL Mueller-Hinton II containing IPTG and concentrations of cefotaxime ranging from 0.015 to 16 µg/mL (two-fold increments) for 24 h at 37ºC. The MIC values were then determined by visual inspection.

Results

All single and double mutants of the Hayes mutant were successfully obtained and tested for their MIC against cefotaxime. Because none of the mutants displayed a significant increase in MIC value and there was not enough time left to create all the triple mutants, we decided to make the 6-fold (complete) mutant and from there work back to the 5-fold mutants (see Appendix V – Constructing the 6- and 5-fold mutant). The necessary triple mutants V103Q:Y105N:I127L and L169A:S235Y:G236S were successfully created, but unfortunately, we could not get the protocol for the creation of the 6-fold Hayes mutant to work. Therefore no five-fold mutants were made either.

MIC-assays

The MIC-assays were done in three-fold (see Appendix I) and the numbers given in Table 1 and Table 2 are therefore averages.

Table 1 The MIC values for cefotaxime of the single mutants and the increase/decrease thereof respective to the wild-type TEM-1.

Mutant	**MIC (µg/mL)**	**-Fold Increase**
pACTEM	0.03	NA
V103Q	0.03	1
Y105N	0.03	1
I127L	0.015	-2
L169A	0.03	1
S235Y	0.03	1
G235S	0.015	-2

Table 2 The MIC values for cefotaxime of the double mutants and the increase/decrease thereof respective to the wild-type TEM-1.

Mutant	**MIC (µg/mL)**	**-Fold Increase**
pACTEM	0.03	NA
V103Q/Y105N	0.03	1
V103Q/I127L	0.03	1
V103Q/L169A	0.015	-2
V103Q/S235Y	0.03	1
V103Q/G236S	0.03	1
Y105N/I127L	0.03	1
Y105N/L169A	0.015	-2
Y105N/S235Y	0.06	2
Y105N/G236S	0.03	1

I127L/L169A	0.015	-2
I127L/S235Y	0.03	1
I127L/G236S	0.03	1
L169A/S235Y	0.015	-2
L169A/G236S	0.03	1
S235Y/G236S	0.03	1

In vitro evolution of mutation V103Q

Due to contaminations[1], only two lines of the *in vitro* evolution of V103Q contained the desired mutation. As mutation V103Q ultimately did not contribute to a higher resistance to cefotaxime, the experiment was abandoned. Therefore only the first round of evolution of two lines are shown below and no MIC measurements were done on either of them.

Table 3 The substitutions found after one round of in vitro evolution on the V103Q mutant. Shown in grey are the mutations built into the plasmid before the experiment.

DNA site	Mutation	Amino acid substitution				
		1	2	3	4	5
		V103Q		V103Q		
592	T → C			S130*		
660	A → G	H153R				
693	G → A			R164H		
719	A → G			I173V		
720	T → C	I173T				
747	T → C			M182T		
754	A → C	A184*				
765	C → T	T188I				
777	A → T			K192I		
824	A → C	I208L				
838	G → A	E212*				
921	G → C	R241P				
1038	A → G	E281G				

* Silent mutations

[1] For more information, please see Appendix VII – Other experiments

Possible functional effects of the individual mutations that make up the Hayes mutant

The six-fold mutant that is described by Hayes *et al.* is unusual because none of the mutations have been found before in nature and only three have been found in *in vitro* experiments. There are two Internet databases that keep track of all emerging TEM alleles, the Lahey[1] table, which gives an overview of known clinically isolated mutants and the LACtamase Engineering Database[2] (LACED), which automatically gathers new TEM alleles published in public databases (*28*). Neither of these two databases lists any of the mutations found in the Hayes mutant. Besides a list of known mutations the LACED database also contains a list of crystal structures of TEM alleles.

Nevertheless some conclusions can be drawn, like what effect the amino acid changes will have on the overall workings of the enzyme. In Figure 6, the positions of the amino acids which are mutated to form the Hayes mutant are shown with respect to cefotaxime. Though not all mutated amino acids in the Hayes mutant differ enormously from the original amino acids, most of them are significantly altered. This affects the folding of the enzyme, the accessibility of the active site to cefotaxime and leads to stability effects correlated to the change in the active site. When residues are described as 'interacting', this means that any of their atoms (with the exclusion of hydrogen atoms) are within a distance of 4.5 Å (*29*).

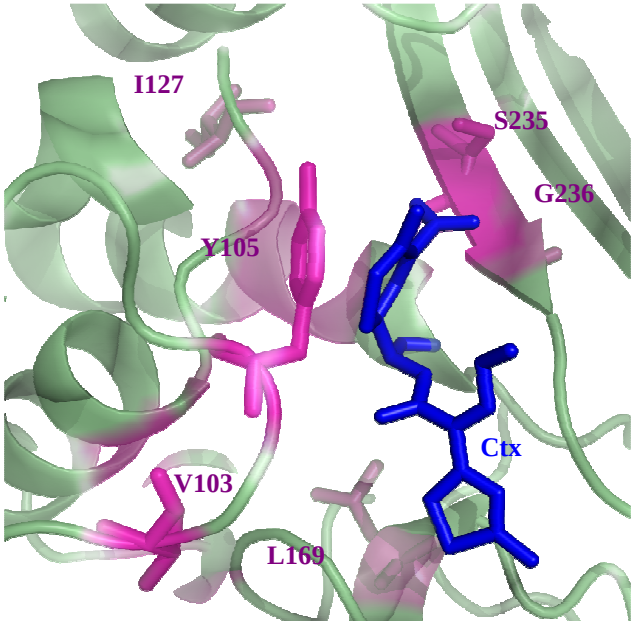


Figure 6 The positions of the Hayes mutations with respect to the cefotaxime molecule (the shown amino acids are still the original TEM--1 amino acids). The more pink the colour of the residue, the closer to the surface of the enzyme it is, residues I127 and L169 are thus completely buried in the enzyme core. The 3D-structure of TEM-1 β-lactamases (PDB-code: 1BTL) was modeled with cefotaxime (blue) in the binding pocket. The cefotaxime was obtained form the crystal structure of TOHO-1 ß-lactamase (PDB-code: 1IYO).

V103Q

The function of V103 is unknown with respect to catalysis of antibiotics. However, the amino acid at position 103 is involved in the binding of the β-lactamase inhibitor protein (BLIP), together with positions Y105, S235 and G236S (*30-32*). As no experiments with BLIP were done by Hayes *et al.*, this does not explain its role in the increased resistance to cefotaxime. Palzkill and Botstein have suggested that a hydrophobic side chain seems to be a predominant requirement at this position, as only 7 out of the 20 mutants they found had a polar side chain and none of these mutants increased the resistance to ampicillin (*33*).

[1] http://www.lahey.org/Studies/temtable.asp
[2] http://www.laced.uni-stuttgart.de/

Because the mutation from a Val to a Gln requires the entire codon to change instead of just one base pair, it is not surprising that this particular mutation has never before been found in nature.
Nevertheless, several precursors and the mutation itself have been found in mutagenesis experiments. Precursors were discovered in the random library studies by Palzkill and Botstein, namely Glu and Leu (*34*), and in the random libraries of Cantu *et al.*, which also contained Leu substitutions (*35*). Regrettably no DNA sequences are given in either study, so it is not possible to say if these mutations are indeed precursor steps to the final mutation V103Q. The mutation to Gln was found in a random mutagenesis experiment done by Huang *et al*, but no MIC measurements of the discovered mutant was done (*36*).

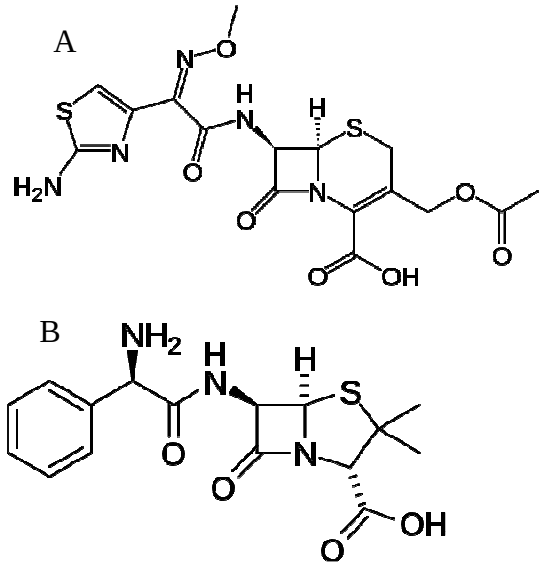


Figure 7 A, the structure of cefotaxime. B, the structure of ampicillin. From the Wikipedia articles Cefotaxime and Ampicillin.

The increased resistance towards cefotaxime can be explained by the change of the involved amino acid from a rather small, hydrophobic group to a bulky and uncharged polar group. Where first there was no possibility of forming hydrogen bonds with the side chain, now there are two acceptors (the N and the O) and two donors (the NH_2) for hydrogen bonds. This may increase the stability of the protein, or a hydrogen bond may be formed with the substrate, as cefotaxime has more H-bonding sited available than ampicillin (see Figure 7). Due to the larger size of cephalosporins compared to penicillins, when β-lactamases show increased cefotaxime resistance, the cavity of the active site has often been enlarged (*13*). This enlargement often results in a loss of stability, therefore all extra stability increases are a nice coincidence.
Usually the size change of the side chain may cause problems with respect to steric hindrance of side chains in the neighborhood. However this effect could be countered by the down-sizing of the side chain of position 105 and by the fact that both these positions are in a loop. This would probably make it easier for the mutant amino acids to find their optimal conformations, as the side chains are not as restricted in their mobility as they would have been in a β-sheet or an α-helix (*37*). This allowance of a bulky side chain is also confirmed by the fact that when the Val is replaced with a Phe, which is also a rather bulky amino acid, the resistance to cefotaxime increases 8-fold (*34*). The mutation to a Phe is in combination with a E104 mutation to Arg, which on itself has no effect on cefotaxime resistance (*38*), but may affect the MIC in combination with the V103 mutation.

Y105N

Palzkill and Botstein thought that position 105 was not essential to binding of substrate or catalysis (*33*), yet this view has changed in recent years, with Y105 playing a large role in stabilization of the substrate. Together with residue P107, the aromatic ring of the Tyr is responsible for the correct positioning of the substrate in the active cleft (*39*). Doucet *et al.* concluded that this positioning is caused by the stacking of the aromatic ring of the Tyr with the aromatic ring(s) of the antibiotic (*40*)(see Figure 6 and Figure 8). In addition to this

positioning role of position 105, Pelletier *et al.*, found that it also acts as a 'gatekeeper' with small residue replacements as Y105G and Y105A allowing discrimination between penicillin and cephalosporin substrates (*41*).
Besides the correct positioning and recognition of the substrate, the residue in position 105 is also responsible for binding of BLIP to the surface interface of TEM (*30-32*). To prevent collision with the Y50 of the BLIP, the side chain of Tyr is rearranges itself into a crowded space, surrounded by the three specificity determining regions of BLIP (*42*).

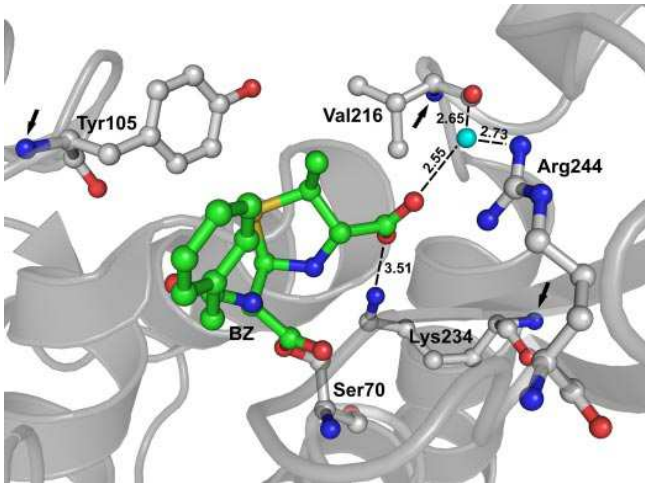


Figure 8 Positioning of the aromatic side chain of Y105 with the substrate benzylpenicillin. From (40)

The mutation of Tyr to Asn has not been found in TEM alleles before, even though it only requires one base pair substitution. Asn as the amino acid at position 105 is however found in the β-lactamases of *Streptomyces cacaoi* (*21*) and *Nocardia lactamdurans* (*38*).
Even though there is no occurrence of Y105N in natural TEM alleles, it is the most well studied of the Hayes mutations and therefore substitution Y105N has been found many times in experiments. The first of these occurrences is in the random libraries of Palzkill and Botstein, where it is found in 4 out of the 20 mutants (*34*). In these studies the mutation is found in combination with E104Y and E104R and does not increase the resistance to cefotaxime. In 2004, Doucet *et al.* did a site-saturation study on position 105 and inevitably encountered the mutation Y105N (*39*). They also did not see an increase in the MIC to cefotaxime due to this mutation. While studying the robustness of TEM-1 to simultaneous substitutions in the active-site region, De Wals *et al*. (*43*), encountered the following two mutants containing Y105N: E104V/Y105N/G238S/E240T and E104V/Y105N/G238S/E240L. This second mutant conferred a 16-fold increase in cefotaxime resistance and a four-fold increase in amp resistance. As the mutant with E240T grants only a four-fold increase in cefotaxime resistance, the mutation Y105N most likely does not contribute significantly to cefotaxime increase on its own. The increase observed probably comes mostly from the G238S mutation, which on its own infers a 16-fold increase in cefotaxime resistance (*7*). The LACED database lists a crystal structures of TEM with an artificial Y105N mutation (*31*).

As the aromatic ring at position 105 is essential to positioning of the substrate, it is not surprising that most substitutions at this location are also aromatic. Why then does Asn, which is polar and non-aromatic, provide an increase in antibiotic resistance? This is in part due to the amide group ($H_2NC=O$) at the end of the residue, which gives Asn the planar conformation necessary for the correct stacking and thus positioning of the substrate within the active cleft. Doucet *et al.* also found that the position of the Asn residue differs in bound and unbound state (*39*), thus suggesting that the side chain of Asn positions itself for optimal stacking. A second observation that Doucet *et al.* made about Asn is that it possesses only a few degrees of freedom as is the case with aromatic side chains (*44*).

There are several other possibilities why Asn confers a greater resistance. First, it is smaller than Tyr and its insertion into the enzyme probably increases the size of the active cleft, making it easier accessible to chephalosporins. Secondly Tyr can accept and donate one H-bond, whereas Asn can donate and accept two H-bonds.

I127L

Again, the role of this residue is not completely understood. It is thought that, while it is deeply buried in the enzyme (*36*) and thus not easily accessible to solvent (see Figure 6), it is located closely to K73, which is known to have an important role in β-lactam catalysis (*45*). Also in a simulation of Fisette *et al.* it is shown that I127 and S235 make a hydrogen bond with the active site residue K234, thus stabilizing the active site when no substrate is bound (*46*) (see Figure 9).

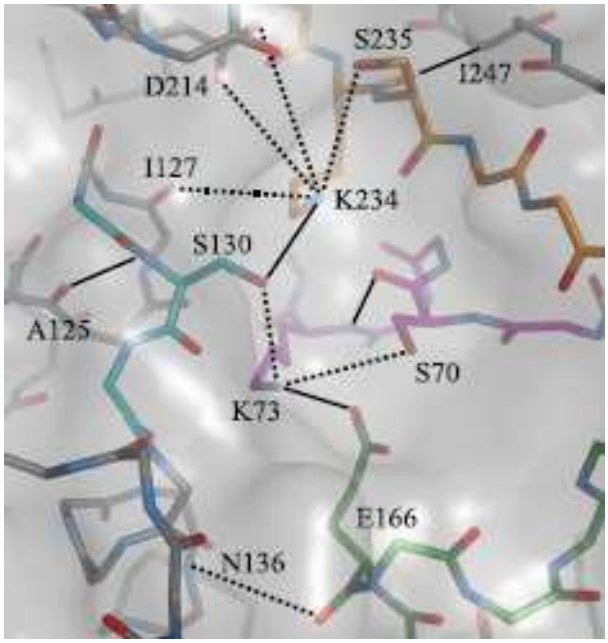


Figure 9 The hydrogen bond network that stabilizes the active site of TEM-1 when no substrate is bound. From the Supplementary data of (46)

The substitution of Ile to Leu in position 127 has not been identified before in natural TEM-1 isolates in a published paper. It has however been identified in other Class A β-lactamases of *Bacillus licheniformis*, *Klebsiella oxytoca* and *Streptomyces fradiae* (*21*). The consensus sequence is a Val at position 127 and this mutation occurs twice in the Lahey table, in TEM-80 and TEM-81 respectively. Both mutants also carry the mutation M69L and TEM-80 has the additional mutation N276D. The mutations M69L and N276D in TEM-80 affect inhibitor resistance and I127V was shown to have very little effect on its own (*47*).
In an *in vitro* evolution experiment on the first 140 amino acids of TEM-1, the mutation I127L was found, but the MIC of the individual mutation was not tested (*48*).

The mutant I127V has no effect on binding of inhibitor and also no effect on the MIC for cefotaxime. It was postulated that the lack of a methyl group on Val with respect to Ile, might change the conformation of the active site (*47*). As Val has the same conformation as Leu, except that it has one CH2 group less, Leu probably has no more effect than Val. Or if the methyl-group has an effect, than Leu, with one methyl-group more than Ile might have an even bigger effect in stabilizing the substrate. The effect of a change form Ile to Leu is probably not very significant, as all properties remain the same (both are small hydrophobic molecules), the only difference residing in the different conformation of the C-atoms in their respective side chains.

L169A

The mutation in position 169 is very interesting, as it occurs in a chain of six amino acids, positions 166-172, which form a large part of the substrate-binding pocket (*34*) and were no mutations have ever been found before (see LACED). Also this stretch of amino acids is conserved throughout the Class A β-lactamases (*21*). The residue L169 is buried in the hydrophobic core of the enzyme (*36*) (see Figure 6) and interacts with the residue F72 (*49*).

When no substrate is present in the active cleft, then L169 makes two hydrogen bonds with water molecules, thus stabilizing the active site (*37*) (see Figure 10).

As stated above, no natural occurring TEM alleles with a substitution in position 169 have been described.
In *in* vitro experiments however, the substitution of a Leu for an Ala is found once in combination with E168G and N170G (*34*). No significant increase in MIC-values for cefotaxime was observed, though the mutant does have an eight-fold drop in ampicillin resistance. Another occurrence of L169A can be found in a study by Cantu *et al.* (*35*).

Both Ala and Leu are small, hydrophobic amino acids and thus no big difference in properties is expected. However, Ala is somewhat smaller than Leu and the increased resistance to cefotaxime can then be explained by a better fit of the antibiotic in the active cleft.

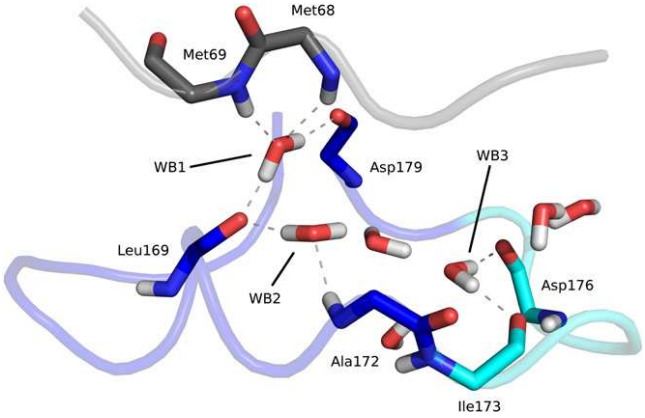


Figure 10 The formation of water bridges which stabilize the substrate-empty active cleft. From (37)

S235Y

The Ser of position 235 is not only important for the stabilization of the active site without substrate (see Figure 9), but is also very important for stabilizing the bound substrate, as the hydroxyl group makes a H-bond with the carboxylic acid group of benzylpenicillin (*50, 51*) (see Figure 11). It was postulated by Imtiaz *et al.* that the residue at position 235 contributes more to the catalysis of cephalosporins than of penicillins (*52*). It is also involved in binding of the BLIP (*32*). Huang *et al.*, found that the position 235 was intolerant to substitutions and thus invariant in TEM (*36*), though in fifty percent of the class A lactamases it is changed to a Thr and this is also the residue found in the consensus sequence (*21*).

This fifty-fifty division between Ser and Thr is also reflected in the fact that half of the mutants of Palzkill and Botstein contained Ser and the other half contained Thr (*33*). No example of the substitution of a Tyr for a Ser at position 235 has ever been mentioned in literature, neither for clinically isolated TEM alleles nor in *in vitro* experiments. The S235Y mutation was the only mutation that was not designed by PDA in the Hayes *et al.* experiment, but rather it was found after an *in vitro* evolution round. Hayes *et al.* hypothesize that the PDA did not calculate the mutation to Tyr, as it predicted steric clashes to occur. Protein backbone motions however was not considered in the calculations of the PDA and may therefore be the reason that the steric clashes do not occur (*25*).
The LACED database lists the crystal structure of TEM with an artificial S235A mutation (*53*).

The effect of the mutation from Ser to Tyr does not stem from a change in polarity, as both residues have a hydroxyl group with which to bind the substrate. This hydrogen bonding with the hydroxyl group is so important that only other substitute, Thr, also has a hydroxyl group to bind the substrate. There are however two major differences: First Tyr is quite a lot larger than Ser and secondly, it has the hydroxyl is bound to a phenol ring. The larger size of the

residue Tyr might make it easier to find another hydrogen bonding partner, as cefotaxime has more opportunities to form hydrogen bonds than ampicillin (see Figure 7). This might be especially useful together with the mutation G236S, as that mutation opens the possibility of a hydrogen bond between residue 236 and the substrate.

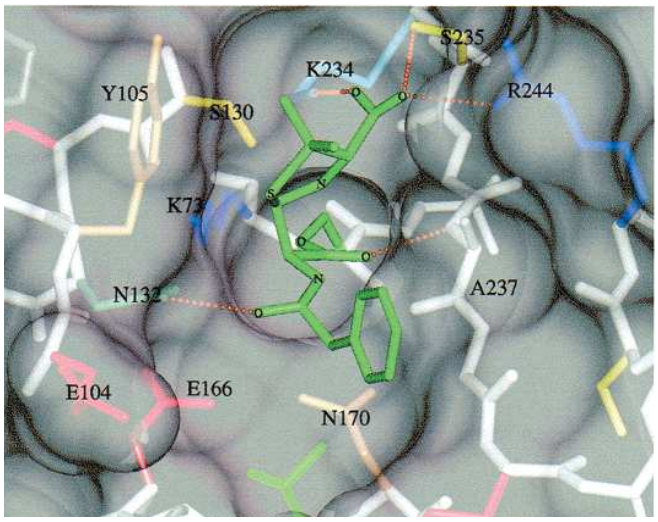


Figure 11 The hydrogen bond between the hydroxyl at position 235 and the substrate. From (51)

G236S

According to Huang *et al.* (*36*), the residue at position 236 is invariant in TEM and all other class A lactamases (*21*) and thus is intolerant to substitutions. Cantu *et al.* (*35*) postulated that substitutions other than Gly would cause a steric clash with S70, but also found that a Gly to Ala change was tolerated. While the role of the residue in antibiotic degradation is mysterious, it is known to be involved in the binding of BLIP (*32*).

While no natural occurring isolates with a mutation at position 236 exist, the substitution of Gly for a Ser is found in Palzkill and Botstein (*33*), but no MIC-values for cefotaxime are available.

The difference between the size of Gly and Ser is small; the greatest difference lies in their respective polarities. Due to the hydroxyl group of Ser, the residue at position 236 can make a hydrogen bond with the substrate.

In the 6-fold mutant

While it is impossible to say exactly what the effect of all mutations together would be without the crystal structure of the Hayes mutant, some speculations can be made. Cefotaxime has a carboxylic acid and a carbonyl in close proximity at the ‘head’ of the molecule (see Figure 7). These groups can interact with the two hydroxyl groups at positions 235 and 236. The amide of glutamine (103) can also interact with the hydrogen acceptors at the other end of cefotaxime. The substitution at 105 increases the size of the active site, without losing the stabilization of the substrate due to the planar conformation of the residue. The substitution at 169 probably also increases the active site and so does the one at 127 as well.

Discussion

The fitness landscape of TEM-1

As can be seen in Table 1 and Table 2 no significant increase in MIC values was observed for either the single or the double mutants, except for the double mutant Y105N:S235Y. However, because the increase in MIC of this mutant is minimal, Y105N:S235Y is regarded as a neutral mutant with respect to the resistance of TEM to cefotaxime.
That no dramatic increases in MIC values are observed is in accordance with the findings of Hayes *et al.* (*25*), who found no increases in MIC of the single mutants.

These wild-type like MIC values enforce the idea that to reach the (local) peak of the Hayes mutant in the fitness landscape of TEM-1, you would need to pass a valley of neutral or even slightly deleterious mutations. The cefotaxime fitness landscape for TEM thus shows two peaks, one for the Stemmer mutant as found by Hall (*22*) and one for the Hayes mutant. The Stemmer mutant is surrounded by other fitness peaks like that of mutation R164S, but might be connected to them by ridges (*23*).

With this data it is nonetheless impossible to say that the Hayes peak is truly isolated, as no intermediates were made for the amino acid substitutions which required more than one point mutation. The mutation V103Q requires two intermediate amino acids and mutations L169A and G236S both require one. Because these intermediates were not tested, it might very well be the case that one of these intermediates does provide the possibility of a ridge between the Stemmer and the Hayes mutant. If this is the case then these peaks are not distinct and though the landscape of TEM-1 is rugged, there is still only one peak on its landscape with numerous ‘bumps’ on its slope.

Another unknown characteristic of the possible fitness peak of the Hayes mutant is the steepness of its slope. As no triple mutants or any of the 4- and 5-fold mutants were tested, it is not clear whether the increase in fitness and thus the slope of the peak starts at one of the triple mutant or if even more mutations are required before the fitness rises above the wild-type levels. As a result, the fitness map of the Hayes mutant is still very incomplete and will remain so, until further research is conducted.

Accessibility of the Hayes mutant under the SSWM model

Simply put, with the knowledge obtained by this research, the Hayes mutant is not accessible to nature under the conditions of the SSWM model as posed by Gillespie (*14*) and used by Weinreich *et al.* (*7*) to study the accessible pathways to the Stemmer mutant. There are two reasons why the SSWM model is not able to explain the natural evolution of the Hayes mutant: Firstly, none of the single and double mutants increase the MIC for cefotaxime and secondly, the occurrence of amino acid changes in the Hayes mutant, which require double and triple point mutations.

As the SSWM model says that only single beneficial mutations are fixed and sweep through the population, one would need a pathway of steadily increasing fitness towards the global optimum. Although there might be multiple pathways, only those that show an increase in fitness at every step will be available to nature (*7*). This is definitely not the case for the single and double mutants of the Hayes mutant, as they seem to be neutral or even slightly deleterious. Though one might argue that the double mutant Y105N:S235Y (see Table 2) may be the starting point of an evolutionary pathway to the Hayes mutant, the probability of both mutations being fixed seperately when neither increases fitness on its own is neglible when adherring to the SSWM model. Gillespie calculated that the chance that a beneficial mutation is substituted is proportional to its selective advantage (*14*), therefore mutations that confer a big increase in resistance are more easily fixed in a population, while smaller beneficial mutations are 'jumped over' (*54*). Hence, mutations like G238S and R164S will be more easily fixed than the mutations that are needed for the Y105N:S235Y mutant (see also Reversion of the Hayes mutant to traditional pathways in evolution).

The second reason for inaccessibility of the Hayes mutant under SSWM conditions is the occurrence of multiple simultaneous mutations. This is in part an artifact of the method in which the mutants were constructed, as no intermediates were made for mutations V103Q, L169A and G236S. Therefore no information about possible MIC increases by these intermediates is available and thus it is not known whether they may indeed be the starting point for a succession of beneficial mutations leading to the Hayes mutant. Nevertheless, the occurrence of multiple point mutations in the same codon in the Hayes mutant would make it inaccessible to nature according to the SSWM model as used by Weinreich *et al*. (*7*).

In the SSWM model a beneficial mutation has to be fixed before the next beneficial mutation can arise. With rates of beneficial mutations occurring in individual genomes (Ub) thought to be between 10^{-9} to 10^{-8} mutations per base pair per generation, the chance that a double point mutation could occur seemed to be negligible (*4, 6*). Yet in recent years more evidence has accumulated which suggests that these Ub values have been grossly underestimated. A study by Perfeito *et al.* suggested that Ub was around $2x10^{-5}$ and that 1 in every 150 mutations was therefore beneficial (*55*). The occurrence of double point mutations is suddenly not as rare as previously thought and this is reflected in Appendix IV – Double point mutations in the Lahey, where all double point mutations found in clinical isolates are listed.
The occurrence of so-called 'hypermutation' (*56*) may also stimulate the rise of double point mutations in nature. Where normal *E. coli* will have a mutation rate of 10^{-8} to 10^{-7}, so-called strong mutators already exhibit a mutation rate of 10^{-6} to 10^{-5} (*57*), which is only a hundred times lower than Mutazyme II DNA polymerase at its lowest mutational conditions. As much as 1% of the natural isolates of *E. coli* are stable and strong mutators, meaning that its high mutation rates is inherited (*56*). These higher mutation rates support the Strong Selection – Strong Mutation model as described by Sniegowksi and Gerrisch (*15*), where chance co-occurrence and competition between beneficial mutations characterize evolution.

Consequently, though there seems to be no adaptive pathway to the Hayes mutant and as it contains simultaneous mutations, in theory the Hayes mutant may arise under SSWM conditions. Due to mutation rates that are higher than previously thought and an incomplete map of the fitness landscape, the option of the Hayes mutant occurring in nature is certainly there. However, as Roy found in his reevaluation of the experiments of Weinreich *et al.* (*7*) and DePristo *et al.* (*10*), even when allowing neutral and double mutations, the repeatability

of evolution is still very high (*58*). For the Hayes mutant, this means that even while its occurrence is possible, nature will probably follow a more conventional pathway like those involving the mutations G238S and R164S as initial steps. The chance of the Hayes mutant occurring under the conditions set by the SSWM model thus seems extremely small and this is confirmed by the fact that none of the mutations or their intermediates have ever been isolated from clinical samples.

Accessibility of Hayes mutant under other evolutionary models

As the probability of the Hayes mutant occurring under SSWM conditions is probably very low, that leaves the question whether there is another evolutionary model that might make the Hayes mutant accessible in nature?

In the 1960's and 70's, Kimura proposed the neutral theory of molecular evolution, which states that most mutations are selectively neutral and that therefore most evolutionary change is based on genetic drift (*4*). In large populations, the influence of genetic drift is smaller than that of selection, while in small populations and when bottlenecks occur, genetic drift predominates. This intense genetic drift can lead to the exploration of the neutral networks in the sequence space of a genotype or gene. A neutral network is a part of sequence space in which the genotypes are connected by neutral mutations and thus show no increase in fitness. In the corresponding fitness landscape, neutral networks can be depicted as a flat sections of the fitness landscape (*59*) (see Figure 12). Especially for viruses, neutral networks seems to be a common mode to escape immunity (*60*), but also RNA secondary structures show neutral networks (*61*). It was suggested by Bloom *et al.* that neutral drift of proteins that are not under selective pressure, may prepare the protein for functional evolution when new functions are favoured by selection in the future (*62*). Bershtein *et al.* found that intense neutral drift changed the ancestor TEM-1 into a protein with many evolutionary possibilities and a high resistance towards deleterious mutations (*63*).

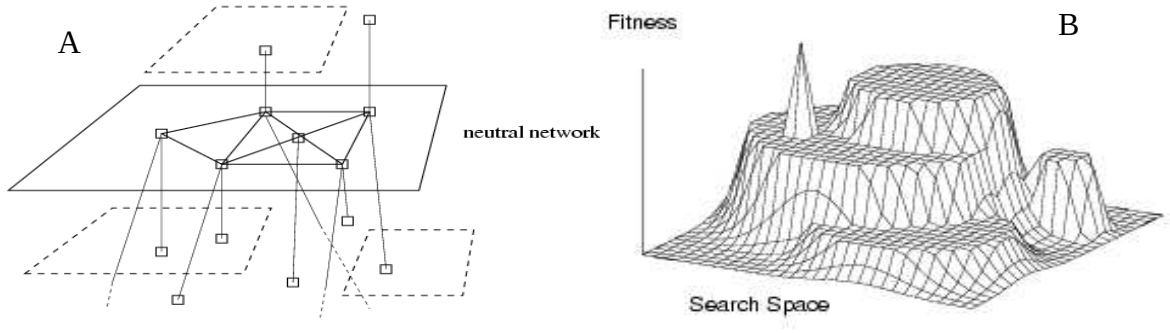


Figure 12 The common depiction of a neutral network (A) and a neutral fitness landscape (B). From (59)

Under intense neutral drift, the mutations of the Hayes mutant could possibly occur in nature and the neutral theory of molecular evolution could explain its emergence. Yet because the exact shape of the fitness landscape of the Hayes mutant is not known, it might very well be that no neutral mutations are required to make the Hayes mutant. If one of the intermediates of mutations V103Q, L169A or G236S does grant an increase in resistance towards

cefotaxime than it will be a starting point for an adaptive pathway in accordance with the SSWM model. However more research should be conducted on this possibility.

Another means by which the peak of the Hayes mutant can be reached is fluctuation of the environment. So far, the Hayes mutant has only been tested for the antibiotics cefotaxime and amipicillin (*25*), while the mutations might also provide resistance increases to other cephalosporins, or to protein inhibitors like BLIP (see also
Further suggestions for research). As stated before, fitness landscapes are mapped for one particular environment, yet in nature, environments change continuously. This can be seen in the continuing penicillinase activity of ESBLs, which severely constrains what mutations can be allowed in mutants that provide cephalosporin resistance. Because clinical environments still use penicillins for many purposes, the penicillinase activity of ESBLs can therefore not be sacrificed for higher resistance to cephalosporins. The β-lactamases must therefore provide resistance to two quite different antibiotics and this leads to two varying goals in evolution, as the β-lactamases need larger active sites to degrade cephalosporins, but when doing so lose penicillinase activity (*13*).

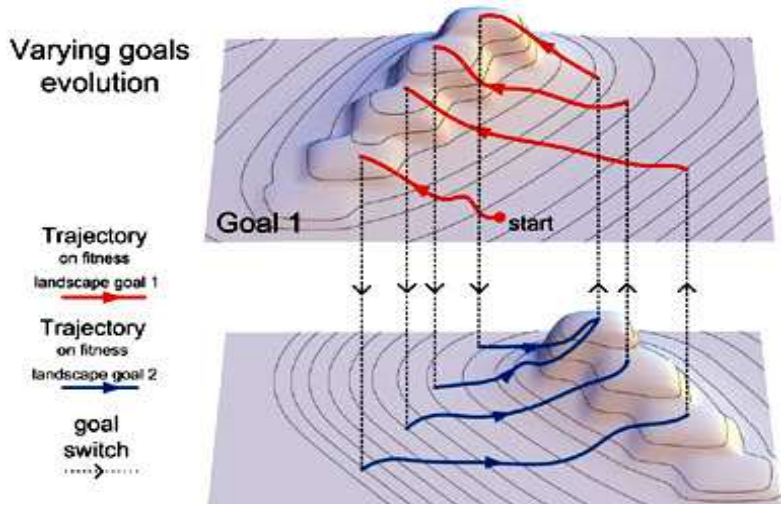


Figure 13 Due to changing environments, the goal of evolution changes. Therefore the protein meanders its way to a fitness point which has a global maximum for both goals. From (64)

This dilemma is nicely illustrated in Figure 13, where two different landscapes are shown on which the gene or genotype tries to find its way to the top, without losing the increase in fitness found on the other landscape.
These variations in environment and evolutionary goals seem to speed up the process of evolution, in fact the more complex the problem, the greater the acceleration (*64*). Combine this with the findings of Blazquez that antibiotics directly increase higher mutation rates (*56*) and it is not surprising that more and more resistant bacteria are found in clinical isolates. According to Blazquez, antibiotics increase mutation rates as they induce stressful conditions and so stimulate the SOS-system and error-prone DNA polymerases of the bacteria. One of these stressful conditions is the adding of β-lactams, as they prevent the forming of the cell wall of the bacterium (*56*). For some CTX-M β-lactamases mutants it has already been discovered that they can only occur when selected for at least two antibiotics, namely cefotaxime and ceftazidime (*65*). Antibiotics therefore not only select for efficient resistance mechanisms, but also act as diversifying and accelerating agents of evolution (*56, 64, 65*).

For the Hayes mutant this means that while some of its mutations that are neutral or deleterious in a cefotaxime environment may be beneficial in other environments and so can become fixed. The occurrence of several Hayes mutations may then only be a matter of time, yet more research should be done on this.

Reversion of the Hayes mutant to traditional pathways in evolution

Though we were not able to test how the Hayes mutant would react to *in vitro* evolution experiments, we did do a first round of *in vitro* evolution with the V103Q mutant. Unfortunately, due to contamination problems, only two out of five lines resulted in useful data, but still some conclusions can be drawn from them. The mutations that occurred after one round of directed evolution are listed in Table 3. Because of time constraints, only one round of evolution was done.

Of the mutations in line 1 only one has been described in the Lahey table, namely the H153R mutation. It has been suggested that this mutation may compensate for destabilization effects caused by other adaptive mutations (*9, 63*). The substitution I173T has been found in two laboratory studies (*23, 26*) and is a variation of the more common I173V substitution which is found in line 3. The effect of I173T is not known, yet the substitution I173V has been found to increase the resistance to the antibiotics cefepime (*26*) and cefotaxime (*23*). The mutation T188I is also found in Salverda *et al.* (*23*), but its effect on antibiotic resistance is unknown. The possible effects of the other mutations found in line 1 are as of yet still unidentified.

In line 3 the mutations I173V, M182T and R164H occur. These are three very common substitutions and all three are found in the Lahey table. As stated above, the I173V mutation increases the resistance to cefepime and cefotaxime. The substitution R164H is found in 18 clinical isolates and is thought to make a hydrogen bond with D179, thus increasing the degradation of the antibiotic ceftazidime (*66*). What is rather interesting about the substitution R164H, is that it is found in combination with substitution G238S, whereas the mutant R164S:G238S is shows reciprocal sign epistasis (*23*). This might be a sign that the mutation V103Q can be integrated into a G238S adaptive pathway. Furthermore the substitution M182T is a well-studied stabilizing mutation often found in the background of R164S or G238S (*13, 67*). While the mutations R164S and G238S increase the active site to make it accessible to cephalosporins, they also decrease the stability of the enzyme. The M182T substitution can compensate for this by making extra hydrogen bonds with residues nearby (*13*). The possible effect of the K192I substitution is not known.

While the extra mutations in line 1 seem to be stabilizing, those in line 3 are actually involved in increasing the resistance against cefotaxime. Especially the occurrence of R164H seems to suggest that mutation V103Q may become part of a G238S or R164S/H pathway. One reasons why V103Q may easily direct evolution towards a more common pathway is the fact that it is furthest away from the active site and may thus be more robust to other mutations(*68*). On the other hand the Y105N substitution may destabilize the active site and require stabilizing mutations that do not usually occur in the R164S or G238S pathway, thus opening ability to explore new pathways. More research into the pathway of the Hayes mutant and more *in vitro* evolution experiments should be done on the single and double mutants to truly explore the possibilities of these mutants.

It would be extremely unlikely for the 6-fold Hayes mutant to follow the R164S or G238S pathways. As a lot of substitutions have occurred in the active site, thus any additional substitutions there would probably severely disrupt the function of the enzyme. If it had been

possible to make the 6- and 5-fold mutant, this might have been observable in the cefotaxime MIC assays of the 5-fold mutants.

Conclusions

With the results presented in this thesis, the Hayes mutant can not be accessed under SSWM conditions. Even though mutation rates are higher than previously thought and there might be an adaptive pathway from one of the intermediates, evolution will probably favour pathways that provide a larger fitness increase earlier in the pathway. Examples of such pathways are those including mutations G238S and R164S. The individual mutations found in the Hayes mutant will therefore most likely be found in the background of G238S and R164S, if found at all.
The Hayes mutant would be accessible under the conditions of the neutral evolution theory, yet to obtain the intense genetic drift required, severe bottlenecks or small populations are called for. When combining these small populations with fluctuating environments, the likelihood of the Hayes mutant occurring in nature becomes more plausible. However, the situation in which the Hayes mutant may occur has to be researched further.

All in all more research has to be conducted before the emergence of the Hayes mutant in the future can conclusively be ruled out. There is however one thing that can be concluded and that is that where the mutant A42G:E104K:M182T:G238S can be seen as the Mount Everest of the TEM-1 fitness landscape for cefotaxime (7), the analogy for the Hayes mutant would be the K2. Mountaineers also call it the “Savage Mountain” due in part to its extremely complicated ascent and the fact that approximately twenty percent of all climbers have perished on their way to the top. In more scientific phrasing, it could be described as an ‘isolated island protein’ (*6*), in the sense that it will probably be a steep fitness peak surrounded by an impassable valley of neutral and deleterious mutations.

Further suggestions for research

As is probably clear by now, a lot of research still has to be done on the intermediates of the Hayes mutant. Not only do the intermediate point mutations of V103Q, L169A and G236S have be made, also their MICs have to be tested, to rule out the possibility of a pathway to the Hayes mutant under the strictures of the SSWM model. To get a detailed insight in the possible evolutionary pathway of the Hayes mutant, also the triple, quadruple, 5-fold and 6-fold mutant should be made and tested for their MICs.
Also, the adaptation of the single and double mutants to more traditional adaptation pathways involving G238S and R164S has to be studied in greater detail. Furthermore, the possibility of a neutral network forming around the Hayes mutant should also be explored in greater detail with *in vitro* experiments.

As suggested by Wang *et al.* (13) there is stability-activity trade-off between on one hand the increased resistance of TEM-1 to cephalosporins and on the other hand the loss of enzyme stability. It would be interesting to know whether the Hayes mutant also has such a trade-off, or if its stability is higher than that of the Stemmer mutant. If it is, than that might be another indication that it might be found in clinical isolates in the future.

Interestingly enough, after the PDA had computed the probability table for the various positions, the scientists chose Y105N over Y105Q, which had a higher probability. This is also done in the PDA-1 mutant, and there is no explanation given in the article. It might be that they thought that two tyrosines so close to each other would affect the tertiary folding/structure of the protein (*25*). The difference between the Hayes mutant with substitution Y105N and the one with Y105Q might give some interesting insights in the working of the mutant.

As four out of six Hayes mutations, V103Q, Y105N, S235Y and G236S, are involved in the binding of the protein inhibitor BLIP, experiments studying the resistance to BLIP might produce some surprising results. First of all, it is not known whether the Hayes mutant is resistant to BLIP, yet the possibility that it is, is certainly there. The mutations that occur at the binding sites of BLIP are quite extreme and it is therefore very unlikely that no difference in BLIP resistance is found. As Schroeder *et al.* found, the resistance to inhibition by BLIP is caused by a change in net charge on TEM (*69*). As many of the Hayes mutations change apolar to polar, this changes the charge of TEM and might therefore effect the binding of BLIP to TEM.
Secondly, the single Hayes mutations probably have a larger effect on the resistance to BLIP than the resistance to cefotaxime. Thus, one or several of the single mutations might have a increase in MIC and thus be the start of a adaptive pathway under the SSWM model. As natural environments are actually fluctuating, instead of the ‘static’ environment of a laboratory where mutant enzymes are selected against only one single antibiotic (*13*), this could lead to the emergence of the Hayes mutant in nature.

Acknowledgements

I would like to thank Merijn Salverda and Arjan de Visser for their efficient and enthusiastic supervision of this thesis. Thanks for all the nice discussions in the lab and during coffee breaks, for the enthusiasm when experiments went well and the support when they did not. Also a big thank you to Bertha Koopmanschap for showing me around in the lab and for practical tips. And last but certainly not least, I would like to thank the entire department of Genetics for the nice time!

References

1. Voigt, C. A., Kauffman, S., Wang, Z.-G., and Frances, H. A. (2001) Rational evolutionary design: The theory of in vitro protein evolution, in *Advances in Protein Chemistry*, pp 79-160, Academic Press.
2. Orr, H. A. (2009) Fitness and its role in evolutionary genetics, *Nat Rev Genet 10*, 531-539.
3. Wright, S. (1932) The roles of mutation, inbreeding, crossbreeding and selection in evolution, *Proceedings of the Sixth International Congress on Genetics*, 356-366.
4. Orr, H. A. (2005) The genetic theory of adaptation: a brief history, *Nat Rev Genet 6*, 119-127.
5. Weinreich, D. M., Watson, R. A., and Chao, L. (2005) Perspective: Sign epistasis and genetic constraint on evolutionary trajectories, *Evolution 59*, 1165-1174.
6. Smith, J. M. (1970) Natural selection and the concept of a protein space, *Nature 225*, 563-564.
7. Weinreich, D. M., Delaney, N. F., DePristo, M. A., and Hartl, D. L. (2006) Darwinian evolution can follow only very few mutational paths to fitter proteins, *Science 312*, 111-114.
8. Poelwijk, F. J., Kiviet, D. J., Weinreich, D. M., and Tans, S. J. (2007) Empirical fitness landscapes reveal accessible evolutionary paths, *Nature 445*, 383-386.
9. Salverda, M. (2008) On the natural and laboratory evolution of an antibiotic resistance gene, Wageningen University, Wageningen.
10. DePristo, M. A., Hartl, D. L., and Weinreich, D. M. (2007) Mutational reversions during adaptive protein evolution, *Mol Biol Evol 24*, 1608-1610.
11. Dean, A. M., and Thornton, J. W. (2007) Mechanistic approaches to the study of evolution: the functional synthesis, *Nat Rev Genet 8*, 675-688.
12. Whitlock, M. C., Phillips, P. C., Moore, F. B. G., and Tonsor, S. J. (1995) Multiple Fitness Peaks and Epistasis, *Annual Review of Ecology and Systematics 26*, 601-629.
13. Wang, X. J., Minasov, G., and Shoichet, B. K. (2002) Evolution of an antibiotic resistance enzyme constrained by stability and activity trade-offs, *Journal of Molecular Biology 320*, 85-95.
14. Gillespie, J. H. (1984) Molecular Evolution over the Mutational Landscape, *Evolution 38*, 1116-1129.
15. Sniegowski, P. D., and Gerrish, P. J. (2010) Beneficial mutations and the dynamics of adaptation in asexual populations, *Philos Trans R Soc Lond B Biol Sci 365*, 1255-1263.
16. Datta, N., and Kontomichalou, P. (1965) Penicillinase synthesis controlled by infectious R factors in Enterobacteriaceae, *Nature 208*, 239-241.
17. Frere, J. M., and Joris, B. (1985) Penicillin-sensitive enzymes in peptidoglycan biosynthesis, *Crit Rev Microbiol 11*, 299-396.
18. Matagne, A., Dubus, A., Galleni, M., and Frere, J. M. (1999) The beta-lactamase cycle: a tale of selective pressure and bacterial ingenuity, *Nat Prod Rep 16*, 1-19.
19. Knox, J. R. (1995) Extended-spectrum and inhibitor-resistant TEM-type beta-lactamases: mutations, specificity, and three-dimensional structure, *Antimicrob Agents Chemother 39*, 2593-2601.
20. Stemmer, W. P. (1994) Rapid evolution of a protein in vitro by DNA shuffling, *Nature 370*, 389-391.
21. Ambler, R. P., Coulson, A. F. W., Frere, J. M., Ghuysen, J. M., Joris, B., Forsman, M., Levesque, R. C., Tiraby, G., and Waley, S. G. (1991) A Standard Numbering Scheme for the Class-a Beta-Lactamases, *Biochemical Journal 276*, 269-270.
22. Hall, B. G. (2002) Predicting evolution by in vitro evolution requires determining evolutionary pathways, *Antimicrobial Agents and Chemotherapy 46*, 3035-3038.

23. Salverda, M., Debets, A.J.M., van der Oost, J., Hoekstra, R.F., de Visser, J.A.G.M. (2010) Initial mutations direct alternative pathways of protein evolution (unpublished), Wageningen University.
24. Dahiyat, B. I., and Mayo, S. L. (1996) Protein design automation, *Protein Sci 5*, 895-903.
25. Hayes, R. J., Bentzien, J., Ary, M. L., Hwang, M. Y., Jacinto, J. M., Vielmetter, J., Kundu, A., and Dahiyat, B. I. (2002) Combining computational and experimental screening for rapid optimization of protein properties, *Proceedings of the National Academy of Sciences of the United States of America* 99, 15926-15931.
26. Barlow, M., and Hall, B. G. (2002) Predicting evolutionary potential: In vitro evolution accurately reproduces natural evolution of the TEM beta-lactamase, *Genetics 160*, 823-832.
27. Bolivar, F., Rodriguez, R. L., Betlach, M. C., and Boyer, H. W. (1977) Construction and characterization of new cloning vehicles. I. Ampicillin-resistant derivatives of the plasmid pMB9, *Gene 2*, 75-93.
28. Thai, Q. K., Bos, F., and Pleiss, J. (2009) The Lactamase Engineering Database: a critical survey of TEM sequences in public databases, *BMC Genomics 10*, 390.
29. Voigt, C. A., Martinez, C., Wang, Z. G., Mayo, S. L., and Arnold, F. H. (2002) Protein building blocks preserved by recombination, *Nat Struct Biol 9*, 553-558.
30. Reichmann, D., Rahat, O., Albeck, S., Meged, R., Dym, O., and Schreiber, G. (2005) The modular architecture of protein-protein binding interfaces, *Proc Natl Acad Sci U S A 102*, 57-62.
31. Reichmann, D., Cohen, M., Abramovich, R., Dym, O., Lim, D., Strynadka, N. C. J., and Schreiber, G. (2007) Binding hot spots in the TEM1-BLIP interface in light of its modular architecture, *Journal of Molecular Biology 365*, 663-679.
32. Rudgers, G. W., and Palzkill, T. (1999) Identification of residues in beta-lactamase critical for binding beta-lactamase inhibitory protein, *J Biol Chem 274*, 6963-6971.
33. Palzkill, T., and Botstein, D. (1992) Probing beta-lactamase structure and function using random replacement mutagenesis, *Proteins 14*, 29-44.
34. Palzkill, T., and Botstein, D. (1992) Identification of amino acid substitutions that alter the substrate specificity of TEM-1 beta-lactamase, *J Bacteriol 174*, 5237-5243.
35. Cantu, C., Huang, W. Z., and Palzkill, T. (1997) Cephalosporin substrate specificity determinants of TEM-1 beta-lactamase, *Journal of Biological Chemistry 272*, 29144-29150.
36. Huang, W., Petrosino, J., Hirsch, M., Shenkin, P. S., and Palzkill, T. (1996) Amino acid sequence determinants of beta-lactamase structure and activity, *J Mol Biol 258*, 688-703.
37. Bos, F., and Pleiss, J. (2009) Multiple Molecular Dynamics Simulations of TEM beta-Lactamase: Dynamics and Water Binding of the Omega-Loop, *Biophysical Journal 97*, 2550-2558.
38. Petit, A., Maveyraud, L., Lenfant, F., Samama, J. P., Labia, R., and Masson, J. M. (1995) Multiple substitutions at position 104 of beta-lactamase TEM-1: assessing the role of this residue in substrate specificity, *Biochem J 305 (Pt 1)*, 33-40.
39. Doucet, N., De Wals, P. Y., and Pelletier, J. N. (2004) Site-saturation mutagenesis of Tyr-105 reveals its importance in substrate stabilization and discrimination in TEM-1 beta-lactamase, *J Biol Chem 279*, 46295-46303.
40. Doucet, N., Savard, P. Y., Pelletier, J. N., and Gagne, S. M. (2007) NMR investigation of Tyr105 mutants in TEM-1 beta-lactamase: dynamics are correlated with function, *J Biol Chem 282*, 21448-21459.
41. Pelletier, J. N., DeWals, P-Y., Doucet, N., Volpato, J. (2006) Modifying enzym specifity by combinatorial active site mutations, in *Bozen 2006*, Bozen, Italy.
42. Zhang, Z., and Palzkill, T. (2004) Dissecting the protein-protein interface between beta-lactamase inhibitory protein and class A beta-lactamases, *J Biol Chem 279*, 42860-42866.

43. De Wals, P. Y., Doucet, N., and Pelletier, J. N. (2009) High tolerance to simultaneous active-site mutations in TEM-1 beta-lactamase: Distinct mutational paths provide more generalized beta-lactam recognition, *Protein Sci 18*, 147-160.
44. Doucet, N., and Pelletier, J. N. (2007) Simulated annealing exploration of an active-site tyrosine in TEM-1 beta-lactamase suggests the existence of alternate conformations, *Proteins 69*, 340-348.
45. Matagne, A., and Frere, J. M. (1995) Contribution of Mutant Analysis to the Understanding of Enzyme Catalysis - the Case of Class-a Beta-Lactamases, *Biochimica Et Biophysica Acta-Protein Structure and Molecular Enzymology 1246*, 109-127.
46. Fisette, O., Morin, S., Savard, P. Y., Lague, P., and Gagne, S. M. (2010) TEM-1 backbone dynamics-insights from combined molecular dynamics and nuclear magnetic resonance, *Biophys J 98*, 637-645.
47. Arpin, C., Labia, R., Dubois, V., Noury, P., Souquet, M., and Quentin, C. (2002) TEM-80, a novel inhibitor-resistant beta-lactamase in a clinical isolate of Enterobacter cloacae, *Antimicrob Agents Chemother 46*, 1183-1189.
48. Gorter, F. (2009) The effect of genetic constraints on the *in vitro* evolution of TEM-1 β-lactamase, in *Laboratory of Genetics*, Wageningen University, Wageningen.
49. Shimamura, T., Ibuka, A., Fushinobu, S., Wakagi, T., Ishiguro, M., Ishii, Y., and Matsuzawa, H. (2002) Acyl-intermediate structures of the extended-spectrum class A beta-lactamase, Toho-1, in complex with cefotaxime, cephalothin, and benzylpenicillin, *J Biol Chem 277*, 46601-46608.
50. Diaz, N., Sordo, T. L., Merz Jr, K. M., and Suarez, D. (2003) Insights into the acylation mechanism of class A beta-lactamases from molecular dynamics simulations of the TEM-1 enzyme complexed with benzylpenicillin, *J Am Chem Soc 125*, 672-684.
51. Ness, S., Martin, R., Kindler, A. M., Paetzel, M., Gold, M., Jensen, S. E., Jones, J. B., and Strynadka, N. C. (2000) Structure-based design guides the improved efficacy of deacylation transition state analogue inhibitors of TEM-1 beta-Lactamase(,), *Biochemistry 39*, 5312-5321.
52. Imtiaz, U., Manavathu, E. K., Lerner, S. A., and Mobashery, S. (1993) Critical hydrogen bonding by serine 235 for cephalosporinase activity of TEM-1 beta-lactamase, *Antimicrob Agents Chemother 37*, 2438-2442.
53. Fonze, E., Charlier, P., To'th, Y., Vermeire, M., Raquet, X., Dubus, A., and Frere, J. M. (1995) TEM1 beta-lactamase structure solved by molecular replacement and refined structure of the S235A mutant, *Acta Crystallogr D Biol Crystallogr 51*, 682-694.
54. Orr, H. A. (2002) The population genetics of adaptation: The adaptation of DNA sequences, *Evolution 56*, 1317-1330.
55. Perfeito, L., Fernandes, L., Mota, C., and Gordo, I. (2007) Adaptive mutations in bacteria: high rate and small effects, *Science 317*, 813-815.
56. Blazquez, J. (2003) Hypermutation as a factor contributing to the acquisition of antimicrobial resistance, *Clin Infect Dis 37*, 1201-1209.
57. Matic, I., Radman, M., Taddei, F., Picard, B., Doit, C., Bingen, E., Denamur, E., and Elion, J. (1997) Highly variable mutation rates in commensal and pathogenic Escherichia coli, *Science 277*, 1833-1834.
58. Roy, S. W. (2009) Probing evolutionary repeatability: neutral and double changes and the predictability of evolutionary adaptation, *PLoS One 4*, e4500.
59. Verel, S. (2005) Étude et Exploitation des Réseaux de Neutralité dans les Paysages Adaptatifs pour l'Optimisation Difficile, Université de Nice-Sophia Antipolis, Nice.
60. van Nimwegen, E. (2006) Epidemiology. Influenza escapes immunity along neutral networks, *Science 314*, 1884-1886.
61. Huynen, M. A. (1996) Exploring phenotype space through neutral evolution, *J Mol Evol 43*, 165-169.
62. Bloom, J. D., Romero, P. A., Lu, Z., and Arnold, F. H. (2007) Neutral genetic drift can alter promiscuous protein functions, potentially aiding functional evolution, *Biol Direct 2*, 17.

63. Bershtein, S., Goldin, K., and Tawfik, D. S. (2008) Intense neutral drifts yield robust and evolvable consensus proteins, *J Mol Biol 379*, 1029-1044.
64. Kashtan, N., Noor, E., and Alon, U. (2007) Varying environments can speed up evolution, *Proc Natl Acad Sci U S A 104*, 13711-13716.
65. Novais, A., Comas, I., Baquero, F., Canton, R., Coque, T. M., Moya, A., Gonzalez-Candelas, F., and Galan, J. C. (2010) Evolutionary trajectories of beta-lactamase CTX-M-1 cluster enzymes: predicting antibiotic resistance, *PLoS Pathog 6*, e1000735.
66. Perilli, M., Celenza, G., Fiore, M., Segatore, B., Pellegrini, C., Luzzaro, F., Rossolini, G. M., and Amicosante, G. (2007) Biochemical analysis of TEM-134, a new TEM-type extended-spectrum beta-lactamase variant produced in a Citrobacter koseri clinical isolate from an Italian hospital, *J Antimicrob Chemother 60*, 877-880.
67. Sideraki, V., Huang, W., Palzkill, T., and Gilbert, H. F. (2001) A secondary drug resistance mutation of TEM-1 beta-lactamase that suppresses misfolding and aggregation, *Proc Natl Acad Sci U S A 98*, 283-288.
68. Axe, D. D. (2000) Extreme functional sensitivity to conservative amino acid changes on enzyme exteriors, *J Mol Biol 301*, 585-595.
69. Schroeder, W. A., Locke, T. R., and Jensen, S. E. (2002) Resistance to beta-lactamase inhibitor protein does not parallel resistance to clavulanic acid in TEM beta-lactamase mutants, *Antimicrob Agents Chemother 46*, 3568-3573.

Appendix I - Primers

The base pairs shown in **bold** are the base pairs that are changed. Base pairs shown in *italic*, are the original changes in the primers for the double mutations and for the primers of the 5-fold mutations.

P3
For: 5′ tcatccggctcgtataatgtgtgga 3′

P4
Rev: 5′ actctcttccgggcgctatcat 3′

Single Mutations

V103Q:
For: 5′ gcatacactattctcagaatgacttg**CAG**gagtactcaccagtcacagaaaagc 3′
Rev: 5′ gcttttctgtgactggtgagtactc**CTG**caagtcattctgagaatagtgtatgc 3′

Y105N
For: 5′ ctcagaatgacttggttgag**A**actcaccagtcacagaaaagc 3′
Rev: 5′ gcttttctgtgactggtgagt**T**ctcaaccaagtcattctgag 3′

I127L
For: 5′ gagaattatgcagtgctgcc**T**taaccatgagtgataacactgc 3′
Rev: 5′ gcagtgttatcactcatggtta**A**ggcagcactgcataattctc 3′

L169A
For: 5′ cgttgggaaccggag**GC**gaatgaagccatacc 3′
Rev: 5′ ggtatggcttcattc**GC**ctccggttcccaacg 3′

S235Y
For: 5′ ctggtttattgctgataaat**A**tggagccggtgagcgtgggtctcg 3′
Rev: 5′ cgagacccacgctcaccggctcca**T**atttatcagcaataaaccag 3′

G236S
For: 5′ ctggtttattgctgataaatct**A**g**C**gccggtgagcgtgggtctcg 3′
Rev: 5′ cgagacccacgctcaccggc**G**c**T**agatttatcagcaataaaccag 3′

Double mutations

V103Q:Y105N
For: 5′ ctcagaatgacttg*CAG*gag**A**actcaccagtcacagaaaagc 3′
Rev: 5′ gcttttctgtgactggtgagt**T**ctc*CTG*caagtcattctgag 3′

G236S:S235Y
For: 5′ ctggtttattgctgataaat**A**t*Ag**C***gccggtgagcgtgggtctcg 3′
Rev: 5′ gcagacccacgctcaccggc*Gc**T***a**T**atttatcagcaataaaccag 3′

6-fold mutant

PvuI P3 supplement

Rev: 5′ caaggcgagttacatgatcc 3′

PvuI P4 supplement

For: 5′ gcagtgctgccttaaccatg 3′

5-fold mutant

V103:Y105N

For: 5′ cactattctcagaatgacttg**GTT**gag*A*actcaccagtcacag 3′
Rev: 5′ ctgtgactggtgagt*T*ctc**AAC**caagtcattctgagaatagtg 3′

Y105:V103Q

For: 5′ ctcagaatgacttg*CAG*gag**T**actcaccagtcacagaaaagc 3′
Rev: 5′ gcttttctgtgactggtgagt**A**ctc*CTG*caagtcattctgag 3′

I127

For: 5′ gagaattatgcagtgctgcc**ATA**accatgagtgataacactgc 3′
Rev: 5′ gcagtgttatcactcatggt**TAT**ggcagcactgcataattctc 3′

L169

For: 5′ cgttgggaaccggag**CTG**aatgaagccatacc 3′
Rev: 5′ ggtatggcttcatt**CAG**ctccggttcccaacg 3′

S235:G236S

For: 5′ ctggtttattgctgataaa**TCT***A*g*C*gccggtgagcgtgggtctcg 3′
Rev: 5′ cgagacccacgctcaccggc*G*c*T***AGA**tttatcagcaataaaccag 3′

G236:S235Y

For: 5′ ctggtttattgctgataaaTAT**G**g**A**gccggtgagcgtgggtctcg 3′
Rev: 5′ cgagacccacgctcaccggc**T**c**C***ATA*tttatcagcaataaaccag 3′

Appendix II – Making the double mutants

Mutation to be built	**Plasmid (template)**	**Primer**
V103Q:Y105N	V103Q	V103Q:Y105N
V103Q:I127L	I127L	V103Q
V103Q:L169A	L169A	V103Q
V103Q:S235Y	V103Q	S235Y
V103Q:G236S	V103Q	G236S
Y105N:I127L	I127L	Y105N
Y105N:L169A	L169A	Y105N
Y105N:S235Y	Y105N	S235Y
Y105N:G236S	Y105N	G236S
I127L:L169A	L169A	I127L
I127L:S235Y	I127L	S235Y
I127L:G236S	I127L	G236S
L169A:S235Y	L169A	S235Y
L169A:G236S	L169A	G236S
S235Y:G236S	G236S	G236S:S235Y

Appendix III – MIC values

Mutant	Replicate	MIC value
TEM-1	1	0.03
	2	0.06
	3	0.03
V103Q	1	0.015
	2	0.03
	3	0.015
Y105N	1	0.03
	2	0.06
	3	0.03
I127L	1	0.03
	2	0.015
	3	0.015
L169A	1	0.03
	2	0.03
	3	0.06
S235Y	1	0.03
	2	0.03
	3	0.03
G236S	1	0.015
	2	0.015
	3	0.015

Mutant	Replicate	MIC value
	1	0.03
V103Q:Y105N	2	0.03
	3	0.03
	1	0.03
V103Q:I127L	2	0.03
	3	0.03
	1	0.015
V103Q:L169A	2	0.015
	3	0.015
	1	0.03
V103Q:S235Y	2	0.03
	3	0.03
	1	0.03
V103Q:G236S	2	0.03
	3	0.03
	1	0.03
Y105N:I127L	2	0.03
	3	0.03
	1	0.015
Y105N:L169A	2	0.015
	3	0.015
	1	0.06
Y105N:S235Y	2	0.06
	3	0.06

Mutant	Replicate	MIC value
	1	0.03
Y105N:G236S	2	0.03
	3	0.03
	1	0.015
I127L:L169A	2	0.015
	3	0.015
	1	0.03
I127L:S235Y	2	0.03
	3	0.03
	1	0.03
I127L:G236S	2	0.03
	3	0.03
	1	0.015
L169A:S235Y	2	0.015
	3	0.015
	1	0.015
L169A:G236S	2	0.03
	3	0.03
	1	0.03
S235Y:G236S	2	0.03
	3	0.03

Appendix IV – Double point mutations in the Lahey table

Table 4 From http://www.lahey.org/Studies/temtable.asp

Isolate	Mutation	Original codon	Mutant codon	Accessible by
TEM-60	L51P	ctc	cct	
TEM-101	A280V	gct	gtc	
TEM-125	M69L	atg	cta	M69I
TEM-130	D35P	gat	cct	
TEM-137	E240R	gag	agg	E240K
TEM-142	G238N	ggt	aat	G238S
TEM-178	R43T	cga	aca	
TEM-178	V44S	gtg	tcg	
TEM-178	R178A	cgt	gct	

Appendix V – Constructing the 6- and 5-fold mutant

Because neither the single mutants nor the double mutant showed any significant increase in MIC and there was not enough time to construct all 20 triple mutants, we decided to start at the complete mutant and work our way back from there to the 5-fold mutants.

6-fold (complete) mutant
All used materials are the same as in Materials and methods, unless otherwise stated.
The mutant as found by Hayes *et al* (*25*) was made by first constructing the triple mutants V103Q/Y105N/I127L (mutant 1) and L169A/S235Y/G236S (mutant 2). The enzyme *PvuI* has a restriction site in between the two subsets of three mutations of the two triple mutants (see Figure 5). Subsequently, two different approaches were taken:
1. Vector and insert were prepared from both triple mutants by digestion of their respective isolated plasmids. For the vector of mutant 1 (vector 1), the plasmid was digested with CIP, *PvuI* and *SacI* (all NEB). The insert of mutant 1 (insert 1) was prepared by digesting the plasmid with *PvuI* and *PagI* (both Fermentas). The vector of mutant 2 (vector 2) was made by a digestion with CIP, *PvuI* (both NEB) and *PagI* (Fermentas). The insert of mutant 2 (insert 2) was made by digesting the plasmid with *PvuI* and *SacI* (both NEB). The digested samples were then separated on a 1%-agarose gel and purified from the gel with use of the Agarose Gel DNA Extraction Kit (Roche). Vector 1 and insert 2 were then ligated together O/N at 4ºC with T4 DNA ligase (Promega), as were vector 2 and insert 1.
2. A PCR was done on the triple mutants with primers P3 and PvuI P3 supplement for mutant 1 and primers P4 and PvuI P4 supplement for mutant 2. The PCR-program was the same as for a normal P3P4 PCR. The PCR-products were purified and digested with *PagI*, *PvuI* (both Fermentas) and *DpnI* (NEB) for mutant 1 and *SacI*, *PvuI* (both Fermentas) and *DpnI* (NEB) for mutant 2. After purifying the mutants were religated together and into naïve vector O/N at 4ºC with T4 DNA ligase (Promega).
After purification of the ligations products of method 1 and 2, the products were transformed into DH5αE and plated on LB-agar+tet+amp (with an ampicillin concentration of 50 μg/mL). The next day colonies were used as template in a P3P4 PCR and sent to the sequencer to verify the successful incorporation of the mutations.

5-fold mutants
To make the 5-fold mutants, follow the protocol as described in Materials and methods – Site-directed mutagenesis. As template use the 6-fold mutant and primers can be found in Appendix I - Primers.

Unfortunately no colonies grew on the plates and so the 6-fold mutant was not created and therefore neither were the 5-fold mutants.

Appendix VII – Other experiments

Growth problems on agar-plates

While doing MIC assays of the single and double mutants, I found that while all mutants grow fine in liquid media, most of the mutants did not do so well on agar-plates.
In the original MIC assay protocol, after diluting the O/N culture of the mutant to 10^5 cells/mL, another dilution is made of $2.5x10^3$ cells/mL. when 50 μL of this is plated out on LB+agar+tet, the next approximately 125 colonies should be visible. Just as with the MIC assay, the plating on agar-plates was done in threefold.
In all the MIC assays that were done, only the mutants S235Y, Y105N:S235Y and S235Y:G236S had colony counts that were comparable to the expected number. All other mutants either showed no colonies, or only up to ten. Even pACTEM did not have normal colony numbers, even though the strain was the same used by Merijn and he never experienced any problems with it.

Two possible explanations were thought off and corresponding experiments were designed:
1. The promoter of the β-lactamase is 'leaky', meaning that even when not induced with IPTG, a small amount of enzyme will be produced. If this enzyme misfolds due to the mutations, this could lead to aggregation of the misfolded enzyme (*67*) and would then prevent growth of colonies on plate. This misfolding can easily be detected with a MIC assay of ampicillin. The natural resistance of TEM-1 to ampicillin is 4096 μg/mL (*22*), so if a dramatic decrease is seen in a mutant, this can be sign of misfolding of the enzyme (*67*). Thus an ampicillin MIC assay was done with pACTEM and the mutant V103Q with ampicillin concentrations from 3 to 7062 μg/mL. However, both pACTEM and the mutant grew to a concentration of 1766 μg/mL, indicating that no misfolding of the mutant protein has occurred.
2. Another idea was that the bacteria had a decreased fitness for growing on agar plates. Therefore bacteria that did form colonies would have an increased fitness and give this on to their offspring, who would then also form colonies on plate. So we grew mutants O/N in liquid LB+tet, diluted the culture to $2.5x10^3$ cells/mL and plated that on LB+agar+tet. The next day colonies that had grown were picked up and grown O/N in LB+tet. These were then again diluted to $2.5x10^3$ cells/mL and plated on LB+agar+tet.
The expectation was that more colonies would then form on these plates, this was however not the case.

Contaminations with earlier experiments

When the first round of *in vitro* evolution of V103Q was sequenced, three of the lines contained combinations of mutations from earlier experiments done by Merijn and Bertha (see Table 5). These mutations could stem from several sources, but the most likely would be the pipettes. They had recently been autoclaved, bus this did not seem to have destroyed the DNA and/or plasmid of the mutants.
I scraped the inside of Bertha and Merijn's 200 μL and 1000 μL pipettes with a sterile toothpick and dissolved the scraped off stuff in MQ. The subsequent mixture was then electroporated into DH5αE and plated on LB+agar+tet. The next day four colonies for each pipette were picked up, PCRed with P3P4 and sent to the sequencer. The mutations that were found in these samples can be seen in Table 6 with the experiments they came from. To avoid further contaminations, filter tips were used in *in vitro* experiments in stead of normal tips.

Table 5 The contaminations found in three lines of the in vitro *evolution experiment of V103Q*

Amino acid substitutions				
1	2	3	4	5
	E104K		E104K	E104K
	P107*			
	P145*			
	M182T		M182T	M182T
	G238S		G238S	G238S
				E281D
	H289Q			

Table 6 Contaminations found in the 200 and 1000 µL pipettes

Amino acid substitutions						
200 µL Merijn 1[1]	200 µL Merijn 2[2]	200 µL Merijn 3	200 µL Bertha 1[3]	200 µL Bertha 2[4]	200 µL Bertha 3[5]	1000 µL Merijn
						K34K
L49M						
E104K	E104K		E104K	E104K		
		Y105N				
		I127L				
R164S	R164S				R164S	
						L169A
					A172G	
			M182T	M182T		
				D214G		
V216I	V216I					
						S235Y
						G236S
			G238S	G238S		
					E240K	
G267R	G267R					
S268N	S268N					
Q278H	Q278H					
			E281D			

[1] L2R2 E104K:R164S
[2] L2R2 E104K:R164S
[3] L2R1 E104K:G238S
[4] L1R1 E104K:G238S
[5] Has not been identified before