On the natural and laboratory evolution of an antibiotic resistance gene

Promotor:
Prof. dr. R.F. Hoekstra
Hoogleraar in de Genetica
Wageningen Universiteit

Co-promotoren:
Dr. J.A.G.M. de Visser
Universitair Hoofddocent, Laboratorium voor Erfelijkheidsleer
Wageningen Universiteit

Prof. Dr. J. van der Oost
Persoonlijk Hoogleraar, Leerstoelgroep Microbiologie
Wageningen Universiteit

Promotiecommissie:
Prof. Dr. A. van Belkum, Erasmus Medisch Centrum, Rotterdam
Prof. Dr. J.A.M. Leunissen, Wageningen Universiteit
Prof. Dr. S. Tans, FOM Instituut voor Atoom- en Molecuulfysica, Amsterdam
Prof. Dr. P.H. van Tienderen, Universiteit van Amsterdam

Dit onderzoek is uitgevoerd binnen de onderzoeksschool Production Ecology and Resource Conservation

Merijn Salverda

On the natural and laboratory evolution of an antibiotic resistance gene

Proefschrift

ter verkrijging van de graad van doctor
op gezag van de rector magnificus
van Wageningen Universiteit,
Prof. Dr. M.J. Kropff,
in het openbaar te verdedigen
op woensdag 8 oktober 2008
des namiddags te vier uur in de Aula.

Salverda, Merijn Louis Marten (2008)
'On the natural and laboratory evolution of an antibiotic resistance gene'
Thesis Wageningen University – with references and summary in Dutch

ISBN: 978-90-8504-999-9

Thesis abstract

TEM-1 ß-lactamase is one of the most notorious antibiotic resistance enzymes around. It exists at high frequencies in antibiotic-resistant bacteria around the world and confers resistance to ß-lactam antibiotics, including penicillins (e.g. ampicillin) and cephalosporins. The enzyme displays a remarkable phenotypic plasticity in response to the introduction of new drugs; within a few years after the clinical debut of most new ß-lactam antibiotics resistance conferring variants of TEM-1 are isolated. Such a shift in resistance phenotype is typically caused by just a few amino acid substitutions. Until today, more than 150 variants of TEM-1 with a unique amino acid sequence have been identified.

Because of the clear link between genotype and phenotype (i.e. level of resistance or fitness) and because of the ease of selecting for increased antibiotic resistance, TEM-1 has been used as a model in studies that seek new methods to optimize proteins. These studies combine the power of *in vitro* mutagenesis and *in vivo* selection and have resulted in a wealth of information about which mutations can increase resistance when the enzyme is exposed to an antibiotic that it initially hydrolyzes inefficiently. At a later stage, these techniques were adopted and used to repeat and predict the natural evolution of TEM-1 under various selective conditions. Recently, TEM-1 is increasingly being used as an experimental model for the study of fundamental evolutionary questions, particularly those that benefit from the direct relationship between genotype and phenotype.

In this thesis, both the natural and laboratory evolution of TEM-1 are studied. The aim of the laboratory work is to increase our understanding of the way in which adaptive mutations interact. For this purpose, TEM-1 is mutagenized using error-prone PCR, which creates variation in the resulting copies of the TEM-1 gene. Mutated gene-copies are placed in bacteria which are subsequently selected for increased resistance to cefotaxime (an antibiotic that TEM-1 hydrolyzes poorly). By repeating this process multiple times in independent experiments, the mutations and mutational trajectories involved in the increase of cefotaxime resistance are studied. At a fundamental level, this has lead to a better understanding of the nature of mutation interaction and its consequences for evolutionary contingency and constraint. Evidence indicating that certain ‘silent’ mutations (i.e. mutations that alter the codon sequence but not the amino acid that the respective codon encodes) can also play a role

in increased resistance was found in these data as well.

A phylogenetic study of the sequences of the ~150 different TEM-alleles that have been isolated in hospitals and clinics so far indicates that recombination has played a significant role in the evolution of TEM-alleles, contrary to what is often assumed. Furthermore, amino acid substitutions present in these clinical isolates are compared to those found in laboratory evolution studies of TEM-1, in order to investigate to what extent laboratory evolution can be used as a predictive tool for the natural evolution of antibiotic resistance genes. This overview indicates that laboratory evolution very accurately repeats the natural evolution of TEM-1. Based on these findings, predictions are made about substitutions that may appear in future clinical TEM-isolates, and directions are given how laboratory evolution can be exploited as a predictive tool most efficiently.

Contents

-Chapter 1-

General introduction

Merijn Salverda

A brief history of antibiotics and the problem of antibiotic resistance

The use of antibiotics dates back to long before Alexander Fleming's famous discovery. Ancient cultures have been reported to use antibiotic containing materials like moulds, honey and plants for the treatment of skin infections (MOELLERING 1995) and the secretions of maggots, ancient helpers in wound-care (SHERMAN *et al.* 2000), have been shown to contain antimicrobial compounds (KERRIDGE *et al.* 2005). In 1928, Fleming discovered a halo of inhibition of bacterial growth around a *Penicillium notatum* mould that had infected one of his plate cultures of *Staphylococcus aureus*. Although several researchers had reported antimicrobial characteristics of moulds before him, Fleming was the first to demonstrate that a substance that was secreted by the mould, which he named 'penicillin', was the cause of this antimicrobial effect. Because of problems with the purification and mass production of penicillin it took until the end of the Second World War before the drug proved its success in clinical use. Ironically and perhaps illustrative of the problems that would soon suppress the initial optimism about this 'wonder drug' (Figure 1.1), the first enzyme able to destroy penicillin was identified before penicillin made its clinical debut in 1941 (ABRAHAM and CHAIN 1940). Experiments testing the ability of bacteria to evolve resistance to penicillin at the same time forecasted similar problems (ABRAHAM *et al.* 1941):

'In order to find an answer to the questions whether bacteria will acclimatise themselves to inhibitory concentrations of penicillin (...) a strain of Staph. aureus was cultivated for some months in broth in the presence of increasing quantities of (...) crude penicillin. Even after a few daily subcultures there was evidence of increased resistance, the coccus showing growth in at least twice the previously inhibitory concentration. In about nine weeks, with sub cultivation every few days, an approximately 30-fold adaptation was reached, and after a further 7 weeks the microbe was able to multiply in a concentration of penicillin a thousand times greater than that which inhibited the parent strain in a parallel test (...)'

Indeed, within a year after the clinical debut of penicillin, the first resistant bacterial isolates were reported. The rise in penicillin resistance was the most striking in *Staphylococcus aureus* (Medeiros 1997):

1941 clinical introduction of penicillin
1942 resistant isolates of *S. aureus* found in patients treated with penicillin
1944 resistant isolates of *S. aureus* found in patients not treated with penicillin
1946 14%
1947 38% } percentage of *S. aureus* isolates resistant to penicillin in an English hospital
1948 59%
1953 80% resistant *S. aureus* isolates in hospital strains worldwide

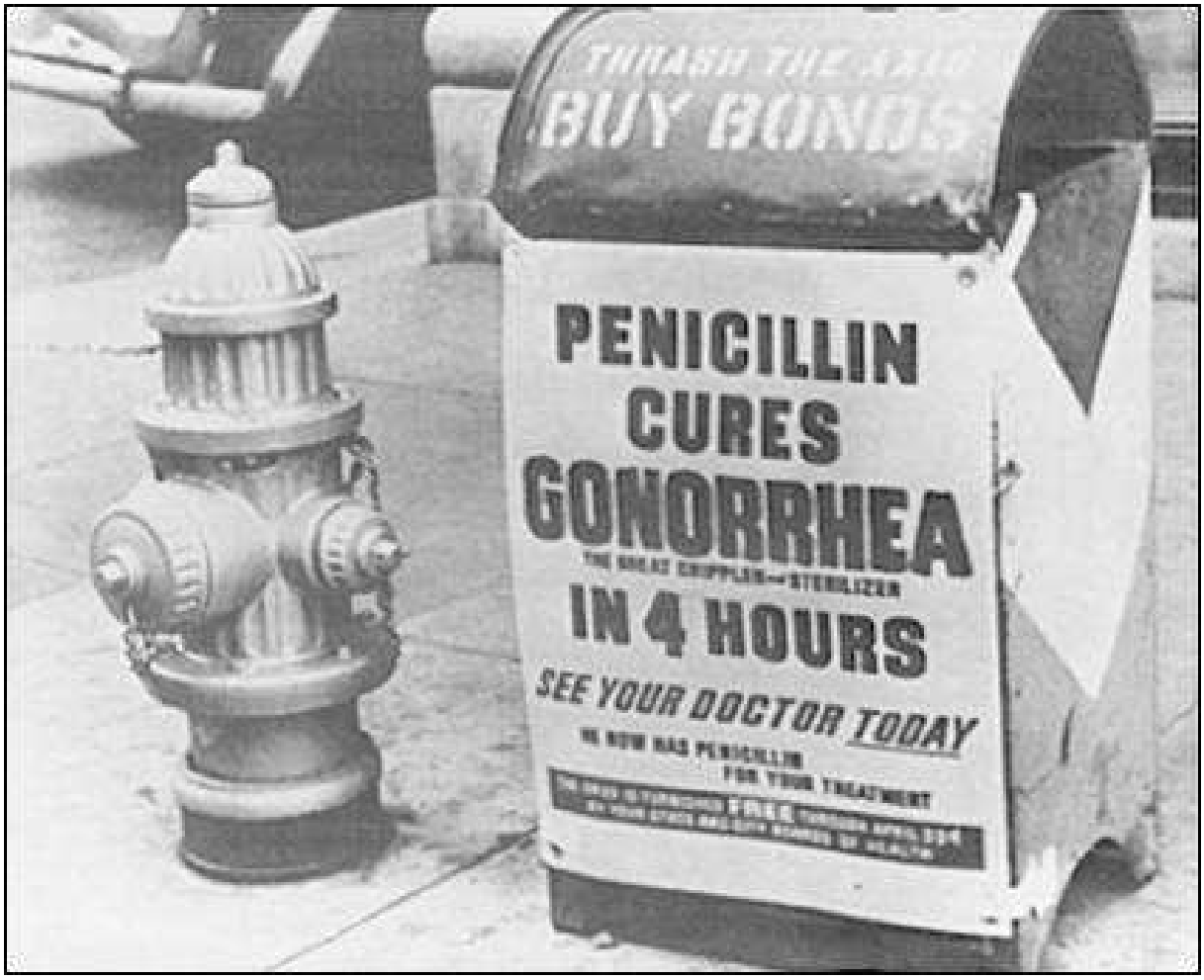


Figure 1.1: U.S. advertisement dating from the golden years of the 'wonder drug'.

The discovery and initial success of penicillin initiated a massive hunt for new antimicrobial agents. Despite of the many antibiotics that have been discovered ever since, it time and again turns out that 'once an antibiotic is proven to be effective and enters widespread human therapeutic use, its days are numbered' (WALSH 2000). With the introduction of new antibiotics and the study of bacterial resistance mechanisms, it became apparent that there are many ways in which bacteria are able to acquire and transmit resistance. Cells can lower their permeability, antibiotics can be actively pumped out of cells and antibiotic targets can be modified. Yet another mechanism is the enzymatic breakdown of antibiotics. The most notorious antibiotic resistance enzymes are the ß-lactamases, which can break down ß-lactam antibiotics. These antibiotics and their mode of action will first be addressed here.

ß-lactam antibiotics and their mode of action

ß-lactam antibiotics are characterized by the presence of a ß-lactam ring in their molecular structure. They include penicillin and derivatives (e.g. ampicillin), cephalosporins (including cefotaxime, the antibiotic studied in this thesis), monobactams, carbapenems and ß-lactamase inhibitors. ß-lactam antibiotics act by interfering with bacterial cell wall synthesis. The antibiotics irreversibly bind to transpeptidases, enzymes that are involved in the cross-linking of the peptidoglycan layer of bacterial cell walls. The initial hypothesis, which is still often referred to, is that this results in instability of the cell wall and lysis under osmotic stress (TIPPER and STROMING 1965). However, the true mechanism is much more sophisticated and acts at several different levels. It was found that bacterial lysis is actually independent of the weakening of the cell wall and is mediated by triggering of the major autolytic enzyme as a result of the drug-target interaction (TOMASZ *et al.* 1970; TOMASZ and WAKS 1975). However, even without this enzymatically regulated lysis antibiotics were found to have a strong bactericidal effect, suggesting that lysis and lethality are two separate bactericidal mechanisms caused by these antibiotics (MOREILLON *et al.* 1990). Indeed, a bactericidal effect of ß-lactam antibiotics that is independent of drug-target interaction and involves induction of a common mechanism of cellular death has recently been described (KOHANSKI *et al.* 2007). All in all, it seems that ß-lactams both disrupt the delicate balance of cell wall metabolism (weakening the cell wall structure and at the same time inducing autolysis) and induce programmed cell death.

Resistance to ß-lactam antibiotics

There are various ways in which bacteria can become resistant to ß-lactam antibiotics. In Gram-negative bacteria, resistance can evolve by means of alteration of the permeability of the outer membrane or the active efflux of ß-lactam antibiotics from the periplasmic space. Some bacteria, e.g. the infamous MRSA, have developed changes in their transpeptidases that cause low affinity for ß-lactams (ZAPUN *et al.* 2008). However, the most common mechanism of resistance is the production of ß-lactamases, enzymes that catalyze hydrolysis of the ß-lactam ring of ß-lactam antibiotics. A wide variety of these enzymes can be found in both gram-positive bacteria (where they are normally secreted) and gram-negative bacteria (where they are employed in the periplasmic space). This thesis deals with the natural and laboratory evolution of TEM-1, the most common of the ß-lactamases (BARLOW and HALL 2002).

System studied

TEM-1 ß-lactamase is one of the most notorious and widespread antibiotic resistance enzymes around. It was isolated as a plasmidic ampicillin resistance determinant in Greece in 1963 and named after the patient (Temoneira) providing the first sample (DATTA and KONTOMICHALOU 1965). Following the introduction of 3rd generation cephalosporins at the beginning of the 1980s, variants of the enzyme with subtle differences in their amino acid sequence were isolated. It had already been shown that a single amino acid substitution in a TEM-1 mutant obtained by *in vitro* mutagenesis could alter the catalytic activity (HALL and KNOWLES 1976) and it was therefore strongly suspected that the sequence alterations observed in clinical isolates could be a new mechanism of resistance (Bush 1989; Sougakoff *et al.* 1988). This was later confirmed using mutants created by site-directed mutagenesis (ZAFARALLA *et al.* 1992). Up until today, more than 150 derivatives of TEM-1 with aberrant amino acid sequences have been described (http://www.lahey.org/Studies/temtable.asp), and it is generally believed that they have evolved largely in response to selection imposed by the clinical use of more modern ß-lactam antibiotics (Barlow and Hall 2002; Medeiros 1997). An overview of the year of introduction of several ß-lactam antibiotics and the year in which a TEM-variant with increased resistance to this antibiotic was first isolated is shown in Figure 1.2.

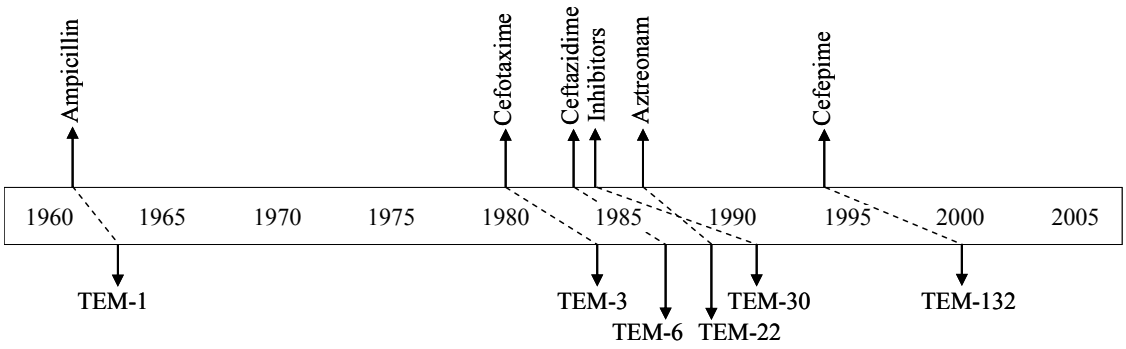


Figure 1.2: Overview of the year of clinical introduction of several ß-lactam antibiotics and the year in which the first TEM-variant conferring increased resistance to this antibiotic was isolated.

Directed evolution

The potential to use TEM-1 ß-lactamase as a model system for molecular evolution was recognized early on (HALL and KNOWLES 1976):

'To trace the evolutionary improvement of an enzyme in molecular terms and to watch the stepwise development of its catalytic apparatus, we need an enzyme susceptible to directed selective pressure towards an alternative catalytic function. So that mutational changes may be fully interpretable, the tertiary structure must be known. These criteria are fulfilled by the plasmid ß-lactamase from E. coli RTEM.'

The study by Hall and Knowles (1976) is in fact one of the first examples of directed evolution, where enzymes encoded by a gene that has been diversified *in vitro* are subjected to selection for gain of function *in vivo*. In this case, whole *E. coli* cells containing a plasmid carrying TEM-1 were subjected to chemical mutagenesis and subsequently selected for increased resistance to cephalosporin C (to which TEM-1 has a low resistance). The mutant TEM variant that was picked up conferred an increased resistance to cephalosporin C and it was later found that this was due to a single amino acid substitution (HEALEY *et al.* 1989).

The boom of different TEM-variants that were isolated in hospitals and clinics worldwide during the second half of the 1980s coincided with the rise of important techniques in molecular biology. Site-directed mutagenesis and easier sequencing methods greatly increased the knowledge of different TEM-variants and the effect of different substitutions on alteration of the resistance spectrum of the enzyme. Because of this and because of the ease of selecting for increased antibiotic resistance, TEM-1 was used as a genetic reporter in a study

on one of the most important inventions in enzyme engineering, DNA-shuffling (STEMMER 1994). A combination of this technique with *in vivo* selection for increased resistance to cefotaxime (a third generation cephalosporin against which TEM-1 has a very low resistance) resulted in the isolation of a TEM-variant with a ~100,000-fold increased resistance to this antibiotic caused by a handful of substitutions. The potential to use directed evolution to study the evolutionary potential of TEM-1 for resistance to novel antibiotics was recognized soon after (VAKULENKO *et al.* 1998). In order to adapt the conditions of *in vitro* evolution to those encountered *in vivo*, an experimental system was set up that both mimicked the mutagenic spectrum of *E. coli* and used multiple selective pressures (BARLOW and HALL 2002). An overview of this method, which is extensively used in this thesis, is shown in Figure 1.3.

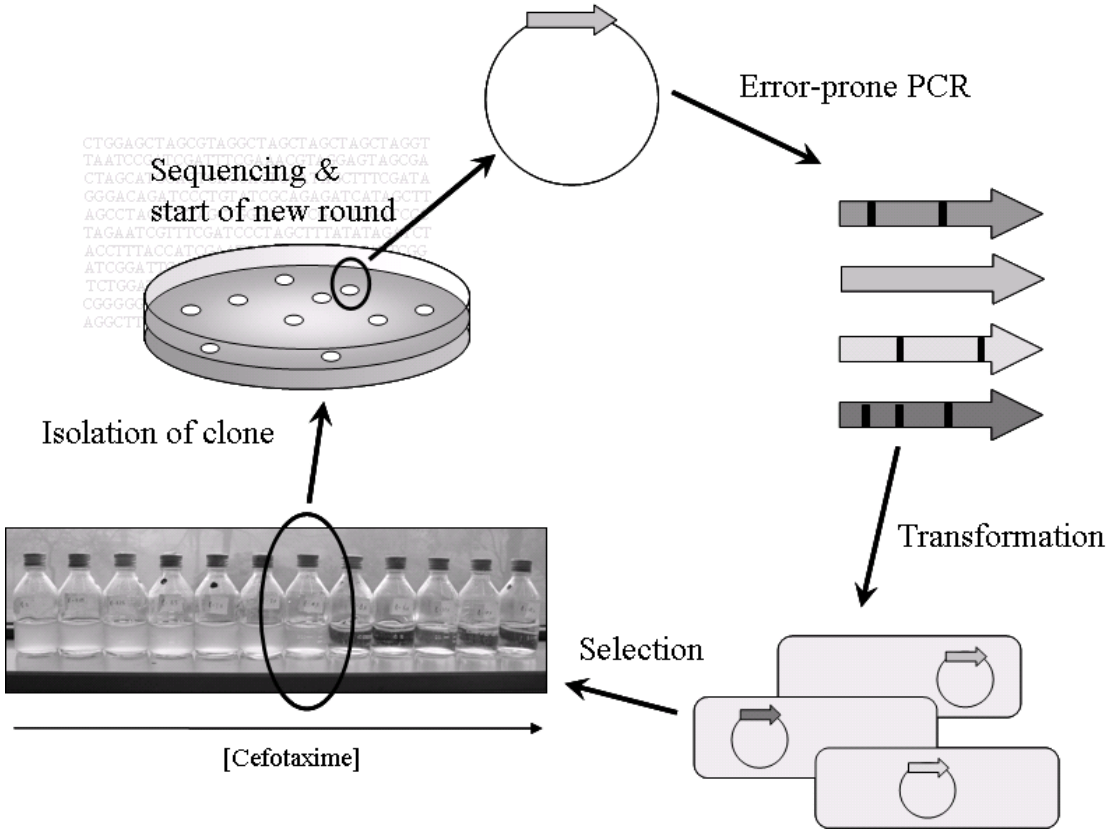


Figure 1.3: Overview of the *in vitro* evolution method used in this thesis and developed by Barlow and Hall (2002). Clockwise from top; an antibiotic resistance gene is located on a plasmid carrying an additional selectable marker. The antibiotic resistance gene is amplified using a sloppy polymerase, resulting in a pool of amplicons with various amounts of mutations. Mutated amplicons are then placed back into the plasmid, which is transformed into *E. coli* cells. Transformant cells are grown overnight in order to expand the library and the next day a series of bottles containing broth with increasing amounts of the antibiotic to be studied is inoculated with these cells. They are then grown again overnight and the next day cells from the bottle with the highest concentration of antibiotic still permitting bacterial growth are plated on agar. The following day, a single colony can be isolated and the altered resistance allele can be sequenced and used again in subsequent rounds of evolution.

TEM-1 as a model system for molecular evolution

TEM-1 ß-lactamase was one of the first antibiotic resistance enzymes for which it was demonstrated that amino acid substitutions could result in alteration of the resistance phenotype (HALL and KNOWLES 1976). Besides this, the clinical isolation of mutant TEM-alleles as a result of the introduction of novel antibiotics at the beginning of the 1980s yielded a 'free' database of amino acid substitutions that were very likely to be involved in alteration of TEM-1's catalytic properties. With this knowledge, TEM-1 has been used as a model system for the study of enzyme structure-function relationships, enzyme engineering, the *in vitro* evolution of antibiotic resistance and various fundamental evolutionary questions (Table 1.1).

Table 1.1: Important studies that have used TEM-1 ß-lactamase as a model system.

Reference	Topic
Hall and Knowles (1976)	Alteration of catalytic properties by *in vivo* mutagenesis
Dalbadie-McFarland *et al.* (1982)	Site directed mutagenesis for the study of protein function
Schultz and Richards *et al.* (1986)	Site-saturation mutagenesis
Oliphant and Struhl (1989)	Altering function by segment mutagenesis
Palzkill and Botstein (1992)	Structure probing by random replacement mutagenesis
Stemmer (1994)	DNA-shuffling
Zaccolo and Gherardi (1999)	High frequency random mutagenesis
Blazquez *et al.* (2000)	Effect of fluctuating ß-lactam pressure on *in vivo* evolution
Barlow and Hall (2002)	Reproduction of natural evolution by *in vitro* evolution
Hayes *et al.* (2002)	Optimization by computational and experimental screening
Osuna *et al.* (2002)	Regain of function after structural alteration
Hall (2002)	Reconstructing evolutionary pathways
Wang *et al.* (2002)	Stability and activity trade-offs
Bershtein *et al.* (2006)	Robustness-epistasis link
Weinreich *et al.* (2006)	Accessibility of evolutionary pathways
Bershtein *et al.* (2008)	Effect of neutral drift on evolvability

The evolution of resistance to cefotaxime (a 3rd generation cephalosporin) is by far the best-studied altered substrate specificity of TEM-1. Using DNA-shuffling, Stemmer (1994) isolated a mutant with a ~30,000-fold increased resistance to cefotaxime, apparently the result of six amino acid substitutions and a promoter mutation. Hall (2002), disregarding the

promoter mutation, tried to identify accessible evolutionary pathways to this six-fold mutant. Surprisingly, he found that only four of the six amino acid substitutions are needed to create an allele with even higher resistance to cefotaxime. Regarding the coding sequence, this mutant TEM-allele confers the highest cefotaxime resistance so far described. Weinreich and colleagues (2006) studied the accessibility of all trajectories leading to a quintuple mutant containing the four substitutions identified by Hall (2002) plus the promoter mutation found by Stemmer (1994). They found that, given that each mutation in such a trajectory must increase resistance, only a small fraction of al 120 possible trajectories leading to this mutant is accessible to Darwinian evolution, due to specific mutational interactions. These mutational interactions and the role they play in shaping the fitness landscape, an important metaphor in evolutionary dynamics, are discussed below.

The fitness landscape

The fitness landscape is a metaphor introduced by Wright (1932) to facilitate the discussion about mutation, selection and drift. The basic fitness landscape is a multi-dimensional graph mapping genotype to fitness. For illustrative purposes, often a three-dimensional picture is used with genotype space depicted in the xy-plane and fitness depicted along the z-axis. However, this simplistic metaphor is problematic, since it is impossible to depict all genotypes of an organism along two axes. Even for a single gene, although depicting the mutational neighbourhood at a distance of a single mutation might be possible, mutational neighbours that are two mutations away (let alone more) are virtually impossible to depict in an organized manner. The fitness landscape is thus a strong simplification of reality, but nevertheless it is a powerful metaphor for discussing evolutionary dynamics, as can be seen in the following example.

Imagine the black ball in Figure 1.4a is an isogenic population of bacteria. This population is optimally adapted to its environment, so it resides on a fitness peak (here shown in just two dimensions). Any genetic change occurring in individuals of this population will either be neutral (causing the ball to move on the peak-plateau, but not in height) or result in a decrease of fitness, causing the population to move downhill. However, since individuals with lower fitness tend to be removed from the population by selection, the population will tend to remain on the peak. Now imagine that the environment of this bacterial population undergoes a sudden change (e.g. the population is exposed to an antibiotic). This will change the

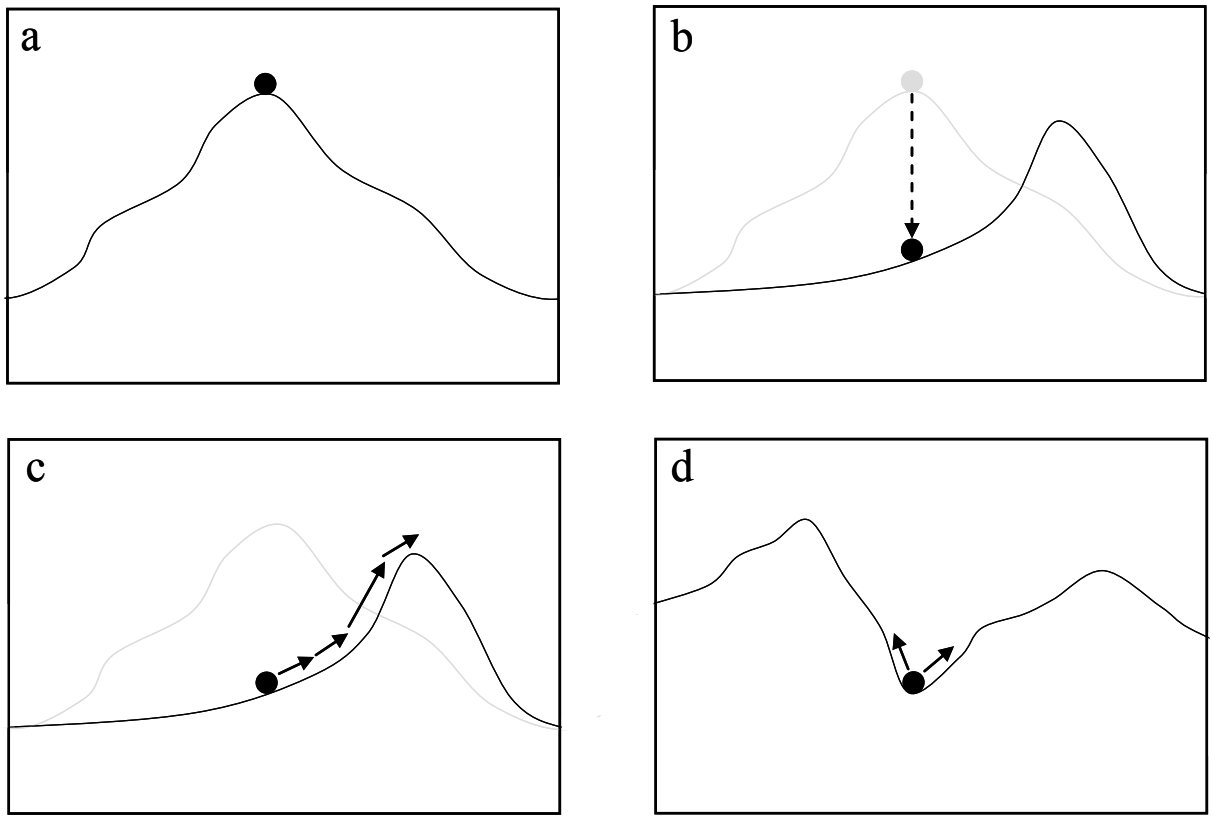


Figure 1.4: Population dynamics on two-dimensional fitness landscapes (fitness is on the y-axis, genotype on the x-axis, the black dot indicates a population of organisms). (a) The population is at a fitness peak and because all mutations decrease fitness, selection keeps the population at this peak. (b) An environmental change (for example exposure to antibiotics in a bacterial population) causes the landscape to change and fitness decreases. (c) Mutations can now give rise to a better adapted genotype that can take over the population, because it reproduces more successfully. Repetition of this process results in an 'adaptive walk' to the top of the fitness peak. (d) The population is in a fitness valley in a multi-peaked or 'rugged' fitness landscape. Alternative solutions to the same problem, involving different mutational pathways, are now available.

landscape radically; as the population is now no longer optimally adapted to its environment, it is all of a sudden at the foot of the hill (Figure 1.4b). Mechanisms to increase resistance exist, but the population simply has not had sufficient time to respond. Over the course of time, an individual with a single beneficial mutation may appear. Such an individual may pull the population upwards, since it can outcompete other individuals and may replace them eventually. In this way, step-by-step, the population can drag itself uphill again, until it is once again back at the fitness peak. Adaptive evolution is the motor behind this "hill climbing" and the process is also known as an "adaptive walk" (Figure 1.4c). An important assumption is that mutations occur one at a time and that an evolutionary trajectory must exist between the lower point and the peak, consisting of consecutive mutations that all increase fitness.

Now, the fitness landscape sketched above is extended by the addition of an extra fitness peak while the bacterial population is once again maladapted (Figure 1.4d, a so-called 'rugged' fitness landscape). The peaks surrounding the population can be seen as alternative solutions to the problem of survival (GAVRILETS 2004). Which peak will be ascended is largely a matter of chance when mutations occur and go to fixation one at a time. Because of this, the population does not necessarily end up at the highest peak. Adaptive evolution is 'blind' for the heights of the different peaks and simply forces the population upward along the steepest slope for which beneficial mutations are available. In this way, the population can end up at a suboptimal peak and selection will keep it at this peak. Escape from such a peak is only possible by genetic drift or by a change in the fitness landscape itself.

A mechanism that shapes the fitness landscape is epistasis, i.e. the interaction between mutations in their effect on phenotype or fitness. Consider two genotypes, ab and AB, with low and high fitness respectively (Figure 1.5). Without epistasis, the fitness effect of mutation a → A is independent of which allele is present at the b/B locus and vice versa (Figure 1.5, left). With sign epistasis, the fitness effect of mutation a → A is dependent on which allele is present at the b/B locus: in the background of b, the fitness is decreased and in the background of B fitness is increased. This does not account for mutation b → B, which increases fitness in both backgrounds (Figure 1.5, middle). Under a special form of epistasis, called 'reciprocal sign epistasis' (POELWIJK *et al.* 2007), the sign of the fitness effect of both mutations a → A and b → B is dependent on which allele is present at the other locus (Figure 1.5, right). In other words, when selection has to move from low fitness genotype ab to high fitness genotype AB, it has to cross a fitness valley, since both intermediate genotypes have an even lower fitness than ab. This type of sign epistasis causes constraints on protein evolution that give rise to multiple adaptive peaks (POELWIJK *et al.* 2007).

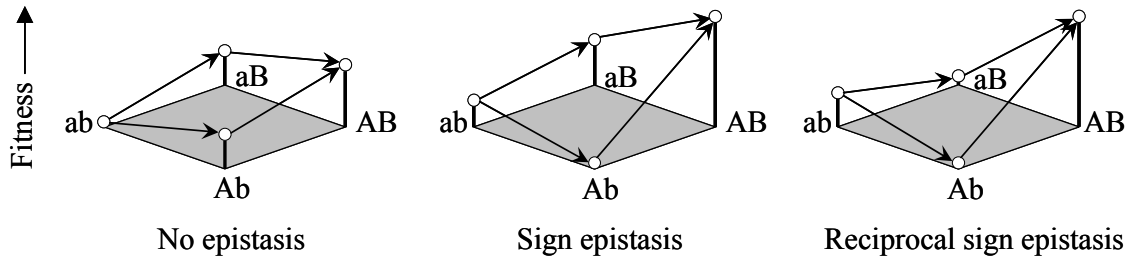


Figure 1.5: Epistatic interactions between two loci with two alleles, a/A and b/B. The picture design is based on Weinreich *et al.* (2005). Left: no epistasis, fitness effect of a mutation is independent of its genetic background. Middle: sign epistasis, sign of fitness effect of mutation a → A depends on what allele is present at the b/B locus. Right: reciprocal sign epistasis, sign of fitness effect of mutation a → A depends on what allele is present at the b/B locus and vice versa.

Outline of this thesis

This thesis deals with the *in vitro* and *in vivo* evolution of TEM-1 ß-lactamase. It is divided in two main parts; one that deals with fundamental questions about mutations and their interactions and one that deals with more practical questions about how antibiotic resistance enzymes evolve and to what extent laboratory techniques can be used to predict their evolutionary potential.

Chapter 2: In this chapter the *in vitro* evolution of TEM-1 ß-lactamase for increased resistance to the antibiotic cefotaxime was studied with the aim to answer fundamental questions about mutation interaction and its effect on the fitness landscape. It is shown that whatever adaptive mutation for cefotaxime resistance appears first can influence the direction and eventual level of adaptation. A special kind of mutational interaction called sign epistasis is demonstrated to be at the basis of this adaptive contingency.

Chapter 3: The potential to predict the natural evolution of antibiotic resistance genes using laboratory techniques is studied in this chapter. Substitutions found in all known laboratory studies in which the evolution of TEM-1 for increased or regained resistance to ß-lactam antibiotics were compared to a dataset that contains all variants of this gene that have been isolated in hospitals and clinics, with the aim to predict future resistant alleles that may be encountered in patients.

Chapter 4: Here, genetic mechanisms underlying the evolution of TEM-variants are examined. Variation in known TEM-variants has typically been attributed to the occurrence of point mutations. Examination of both coding and silent mutations present in a dataset of natural TEM-isolates and the distribution of co-occurring TEM-alleles in clinical microbial populations indicates that recombination is an important additional factor in the evolution of plasmidic TEM-alleles.

Chapter 5: Synonymous mutations present in clinical and laboratory TEM-isolates are more closely examined in this chapter. Although such mutations have long been thought to be functionally ‘silent’, evidence to the contrary has been accumulating. Evidence was found that at least one silent mutation found multiple independent times in laboratory isolates occurs

with a higher frequency than expected for a neutral mutation and is either the result of selection or a mutational hotspot. Another 'silent' mutation is shown to consistently increase resistance to cefotaxime compared to TEM-1. Indications exist that functional-yet-synonymous mutations occur in clinical isolates as well. Putative mechanisms for the functionality of such mutations are discussed.

Chapter 6: In the general discussion, I summarize the most important findings of this thesis and I elaborate on interesting areas for future research.

References

Abraham, E. P., and E. Chain, 1940 An enzyme from bacteria able to destroy penicillin. Nature **146:** 837.

Abraham, E. P., E. Chain, C. M. Fletcher, A. D. Gardner, N. G. Heatley *et al.*, 1941 Further observations on penicillin. The Lancet **238:** 177-206.

Barlow, M., and B. G. Hall, 2002 Predicting evolutionary potential: in vitro evolution accurately reproduces natural evolution of the tem beta-lactamase. Genetics **160:** 823-832.

Bershtein, S., K. Goldin and D. S. Tawfik, 2008 Intense neutral drifts yield robust and evolvable consensus proteins. Journal of Molecular Biology **379:** 1029-1044.

Bershtein, S., M. Segal, R. Bekerman, N. Tokuriki and D. S. Tawfik, 2006 Robustness-epistasis link shapes the fitness landscape of a randomly drifting protein. Nature **444:** 929-932.

Blazquez, J., M. I. Morosini, M. C. Negri and F. Baquero, 2000 Selection of naturally occurring extended-spectrum TEM beta-lactamase variants by fluctuating beta-lactam pressure. Antimicrobial Agents and Chemotherapy **44:** 2182-2184.

Bush, K., 1989 Excitement in the beta-lactamase arena. Journal of Antimicrobial Chemotherapy **24:** 831-840.

Dalbadie-McFarland, G., L. W. Cohen, A. D. Riggs, C. Morin, K. Itakura *et al.*, 1982 Oligonucleotide-Directed Mutagenesis as a General and Powerful Method for Studies of Protein Function. Proceedings of the National Academy of Sciences of the United States of America-Biological Sciences **79:** 6409-6413.

Datta, N., and P. Kontomichalou, 1965 Penicillinase Synthesis Controlled by Infectious R Factors in Enterobacteriaceae. Nature **208:** 239-241.

Gavrilets, S., 2004 *Fitness landscapes and the origin of species*. Princeton University Press, Princeton, NJ.

Hall, A., and J. R. Knowles, 1976 Directed Selective Pressure on a Beta-Lactamase to Analyze Molecular Changes Involved in Development of Enzyme Function. Nature **264:** 803-804.

Hall, B. G., 2002 Predicting evolution by in vitro evolution requires determining evolutionary pathways. Antimicrob Agents Chemother **46:** 3035-3038.

Hayes, R. J., J. Bentzien, M. L. Ary, M. Y. Hwang, J. M. Jacinto *et al.*, 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.

Healey, W. J., M. R. Labgold and J. H. Richards, 1989 Substrate Specificities in Class-a Beta-Lactamases - Preference for Penams Vs Cephems - the Role of Residue-237. Proteins-Structure Function and Genetics **6:** 275-283.

Kerridge, A., H. Lappin-Scott and J. R. Stevens, 2005 Antibacterial properties of larval secretions of the blowfly, Lucilia sericata. Medical and Veterinary Entomology **19:** 333-337.

Kohanski, M. A., D. J. Dwyer, B. Hayete, C. A. Lawrence and J. J. Collins, 2007 A common mechanism of cellular death induced by bactericidal antibiotics. Cell **130:** 797-810.

Medeiros, A. A., 1997 Evolution and dissemination of beta-lactamases accelerated by generations of beta-lactam antibiotics. Clinical Infectious Diseases **24:** S19-S45.

Moellering, R. C., 1995 Past, present, and future of antimicrobial agents. American Journal of Medicine **99:** S11-S18.

Moreillon, P., Z. Markiewicz, S. Nachman and A. Tomasz, 1990 2 Bactericidal Targets for Penicillin in Pneumococci - Autolysis-Dependent and Autolysis-Independent Killing Mechanisms. Antimicrobial Agents and Chemotherapy **34:** 33-39.

Oliphant, A. R., and K. Struhl, 1989 An Efficient Method for Generating Proteins with Altered Enzymatic-Properties - Application to Beta-Lactamase. Proceedings of the National Academy of Sciences of the United States of America **86:** 9094-9098.

Osuna, J., A. Perez-Blancas and X. Soberon, 2002 Improving a circularly permuted TEM-1 beta-lactamase by directed evolution. Protein Engineering **15:** 463-470.

Palzkill, T., and D. Botstein, 1992 Probing Beta-Lactamase Structure and Function Using Random Replacement Mutagenesis. Proteins-Structure Function and Genetics **14:** 29-44.

Poelwijk, F. J., D. J. Kiviet, D. M. Weinreich and S. J. Tans, 2007 Empirical fitness landscapes reveal accessible evolutionary paths. Nature **445:** 383-386.

Schultz, S. C., and J. H. Richards, 1986 Site-Saturation Studies of Beta-Lactamase - Production and Characterization of Mutant Beta-Lactamases with All Possible Amino-Acid Substitutions at Residue-71. Proceedings of the National Academy of Sciences of the United States of America **83:** 1588-1592.

Sherman, R. A., M. J. R. Hall and S. Thomas, 2000 Medicinal maggots: An ancient remedy for some contemporary afflictions. Annual Review of Entomology **45:** 55-81.

Sougakoff, W., S. Goussard and P. Courvalin, 1988 The Tem-3 Beta-Lactamase, Which Hydrolyzes Broad-Spectrum Cephalosporins, Is Derived from the Tem-2 Penicillinase by 2 Amino-Acid Substitutions. Fems Microbiology Letters **56:** 343-348.

Stemmer, W. P., 1994 Rapid evolution of a protein in vitro by DNA shuffling. Nature **370:** 389-391.

Tipper, D. J., and J. l. Stroming, 1965 Mechanism of Action of Penicillins - a Proposal Based on Their Structural Similarity to Acyl-D-Alanyl-D-Alanine. Proceedings of the National Academy of Sciences of the United States of America **54:** 1133-&.

Tomasz, A., A. Albino and E. Zanati, 1970 Multiple Antibiotic Resistance in a Bacterium with Suppressed Autolytic System. Nature **227:** 138-&.

Tomasz, A., and S. Waks, 1975 Mechanism of Action of Penicillin - Triggering of Pneumococcal Autolytic Enzyme by Inhibitors of Cell-Wall Synthesis. Proceedings of the National Academy of Sciences of the United States of America **72:** 4162-4166.

Vakulenko, S. B., B. Geryk, L. P. Kotra, S. Mobashery and S. A. Lerner, 1998 Selection and characterization of beta-lactam-beta-lactamase inactivator-resistant mutants following PCR mutagenesis of the TEM-1 beta-lactamase gene. Antimicrob Agents Chemother **42:** 1542-1548.

Walsh, C., 2000 Molecular mechanisms that confer antibacterial drug resistance. Nature **406:** 775-781.

Wang, X. J., G. Minasov and B. K. Shoichet, 2002 Evolution of an antibiotic resistance enzyme constrained by stability and activity trade-offs. Journal of Molecular Biology **320:** 85-95.

Weinreich, D. M., N. F. Delaney, M. A. DePristo and D. L. Hartl, 2006 Darwinian evolution can follow only very few mutational paths to fitter proteins. Science **312:** 111-114.

Weinreich, D. M., R. A. Watson and L. Chao, 2005 Perspective: Sign epistasis and genetic constraint on evolutionary trajectories. Evolution **59:** 1165-1174.

Wright, 1932 The roles of mutation, inbreeding, crossbreeding and selection in evolution. Proceedings of th sixth international congress of genetics **I:** 356-366.

Zaccolo, M., and E. Gherardi, 1999 The effect of high-frequency random mutagenesis on in vitro protein evolution: A study on TEM-1 beta-lactamase. Journal of Molecular Biology **285:** 775-783.

Zafaralla, G., E. K. Manavathu, S. A. Lerner and S. Mobashery, 1992 Elucidation of the Role of Arginine-244 in the Turnover Processes of Class-a Beta-Lactamases. Biochemistry **31:** 3847-3852.

Zapun, A., C. Contreras-Martel and T. Vernet, 2008 Penicillin-binding proteins and beta-lactam resistance. Fems Microbiology Reviews **32:** 361-385.

-Chapter 2-

Directed evolution of an antibiotic resistance gene reveals evolutionary contingency on early substitutions in a rugged fitness landscape

M. Salverda[1], A.J.M. Debets[1], R.F. Hoekstra[1], J. van der Oost[2] & J. A.G.M. de Visser[1]

[1] Laboratory of Genetics, Wageningen University, the Netherlands

[2] Laboratory of Microbiology, Wageningen University, the Netherlands

Abstract: The exact shape of the fitness landscape is of great importance for many questions in evolutionary biology. Whether such landscapes are smooth (with a single optimum) or rugged (with many optima separated by fitness valleys) remains an open question. Sign epistasis, a special form of mutational interaction where the sign of the fitness effect of a mutation depends on the genetic background the mutation occurs in, can give rise to rugged fitness landscapes. Using directed evolution, the adaptive evolution of an antibiotic resistance gene towards recognition of a novel antibiotic was studied in twelve replicate lines during three rounds of mutagenesis and selection. It was found that multiple sign epistatic interactions between mutations are present which give rise to a rugged fitness landscape and constrain the adaptive trajectory to a limited number of mutational pathways.

2.1 Introduction

The concept of a fitness landscape, i.e. the mapping from genotype to fitness (WRIGHT 1932), is an appealing and useful metaphor in evolutionary biology (COLEGRAVE and BUCKLING 2005). The exact shape of the fitness landscape has important implications for evolutionary theory, including the origin and maintenance of sex and recombination (KONDRASHOV and KONDRASHOV 2001), the nature and extent of evolutionary constraint (WHITLOCK *et al.* 1995; CAMPS *et al.* 2007) and the evolution of diversity and speciation (WHITLOCK *et al.* 1995; GAVRILETS 2004). However, whether this shape is 'smooth' (single peaked) or 'rugged' (multi peaked) remains an open question (LUNZER *et al.* 2005).

Wright (1932) pointed out that when the fitness effect of an allele present at one locus depends on which alleles are present at other loci (i.e. epistasis), such interactions can give the fitness landscape a rugged appearance with many local optima separated by fitness valleys. Fitness valleys, and hence ruggedness, are specifically caused by a special form of mutation interaction called 'sign epistasis', where the sign of the fitness effect of a mutation (i.e. positive or negative) depends on the genetic background in which it appears (WEINREICH *et al.* 2005). For example, consider genotypes ab and AB with low and high fitness respectively. Whenever either or both intermediate genotypes aB and Ab have a fitness lower than ab, the interaction between the a/A and b/B loci is said to be 'sign epistatic', since the sign of the fitness effect of changing from 'a' to 'A' is dependent on the allele present at the b/B locus. Such interactions may constrain evolution towards the genotype of highest fitness and give rise to multiple peaks on the fitness landscape (WEINREICH *et al.* 2005). The type of sign epistasis that most severely imposes adaptive constraints occurs when both intermediate genotypes have lower fitness and is called 'reciprocal sign epistasis' (POELWIJK *et al.* 2007).

Despite the appeal of the landscape metaphor, empirical information on real fitness landscapes is limited due to experimental limitations in obtaining fitness measurements of many known genotypes (COLEGRAVE and BUCKLING 2005). Nevertheless, recent experimental studies with simple systems, such as enzymes, viruses and bacteria, have begun to empirically explore the local fitness surface in a defined environment. Evolution experiments with bacteria (KORONA *et al.* 1994) and viruses (BURCH and CHAO 2000) have shown fitness divergence among replicate populations suggesting the approach of different peaks in a rugged fitness landscape. Other studies have reconstructed evolutionary pathways

from an ancestral to a particular evolved form of an enzyme or metabolic pathway by constructing all intermediate genotypes and measuring their fitness or enzymatic activity (LUNZER *et al.* 2005; MILLER *et al.* 2006; POELWIJK *et al.* 2006; WEINREICH *et al.* 2006). A general finding of these studies is that the fitness landscape shows many signs of ruggedness caused by sign epistatic interactions among the mutations involved. Although the presence of multiple peaks in a fitness landscape has been demonstrated experimentally (BURCH and CHAO 2000), the involvement of sign epistasis can only be inferred in this case due to lack of information about the underlying genetic changes.

In the present study, we use directed evolution experiments with TEM-1 ß-lactamase in order to characterize this enzyme's fitness landscape for the evolution of substrate specificity from ampicillin to cefotaxime (Ctx). Directed evolution (the creation of mutant enzymes *in vitro*, followed by selection *in vivo*) has been used to mimic the natural evolution of TEM-1 (ORENCIA *et al.* 2001; BARLOW and HALL 2002) and to explore its evolutionary potential for hydrolysis of a range of ß-lactam antibiotics (VAKULENKO *et al.* 1998; VAKULENKO and GOLEMI 2002; BARLOW and HALL 2003). Besides, TEM-1 has been used as a genetic reporter exploring the potency of novel directed evolution techniques (STEMMER 1994; ZACCOLO and GHERARDI 1999; CAMPS *et al.* 2003; FUJII *et al.* 2004; KOPSIDAS *et al.* 2007). Most of these studies have used Ctx (a 3rd generation cephalosporin that is poorly hydrolyzed by TEM-1) as selective agent, making this by far the best studied alteration of the enzyme's substrate specificity. The allele encoding the most efficient enzyme that has been identified in these studies has a ~100.000-fold increased minimal inhibitory concentration (MIC) for Ctx and carries a single mutation in the promoter and four non-synonymous mutations in the coding region (HALL 2002). It has been shown, based on MIC assays, that only a small fraction of the evolutionary pathways leading to this quintuple mutant is accessible to Darwinian selection due to sign epistasis between the mutations involved (WEINREICH *et al.* 2006). Moreover, it was shown in the same study that the two most likely pathways of fixation are G238S > E104K > A42G > M182T and E104K > G238S > A42G > M182T (the promoter mutation is left out here since the promoter is not part of the mutagenized sequence in our experiments). An important assumption made in these predictions is that population size and mutation rates are sufficiently low to prevent the selection of more than one beneficial mutation at a time. It is conceivable that under conditions where more than one beneficial mutation is selected at once, some of the constraints caused by sign epistasis may be overcome.

Here, we seek to characterize the fitness landscape during the evolution of TEM-1 towards Ctx hydrolysis. The emphasis is on testing the prediction that sign epistasis can give rise to multiple adaptive peaks, the ascent of which depends on early substitutions which exclude later switching to alternative peaks. A previously developed *in vitro* evolution method (BARLOW and HALL 2002) is used, where replicate lines with varying mutation and recombination rates are selected to increase the ability to detect unique pathways by expanding the amount of sequence space covered. Three consecutive rounds of mutagenesis and selection resulted in the repetitive evolution of (part of) a previously described route to increased Ctx resistance (STEMMER 1994; ZACCOLO and GHERARDI 1999; ORENCIA *et al.* 2001; HALL 2002). Next to this dominant route, several alternative pathways were observed that lack most of the substitutions of the main route. To assess whether these alternative routes represent separate peaks on the fitness landscape, early substitutions suspected to be responsible for blocking the main route to Ctx resistance were identified and built into TEM-1. Three further rounds of directed evolution starting with these mutants show that the identified substitutions indeed block the dominant route towards Ctx resistance and lead to alternative peaks. Sign epistatic interactions are shown to be at the basis of this rugged empirical fitness landscape.

2.2 Materials and methods

Bacterial strains and plasmids: *Escherichia coli* strain DH5αE was used as the host for all plasmids. Plasmid pACSE3 (BARLOW and HALL 2002) was used as the vector for cloning and expressing TEM alleles.

Media: LB-medium is 10g trypticase peptone, 5 g yeast extract and 10 g NaCl/ liter. LB-tetracycline medium is LB-medium containing 15 mg tetracycline/ liter. Mueller Hinton (Merck) and Mueller Hinton II (BD) medium were prepared according to the manufacturers' instructions. SOC medium is 20 g bacto-tryptone and 5 g yeast extract/ liter with 10 mM NaCl, 2.5 mM KCl, 10 mM $MgCl_2$ and 20 mM glucose. Solid media contained 16 g agar/ liter.

Recombinant DNA methodology: Plasmids were prepared from overnight cultures grown in LB-tetracycline medium. Plasmids were purified using GenElute™ Plasmid Miniprep Kit (Sigma) according to the manufacturer's instructions. TEM-1 was amplified from pBR322 (BOLIVAR *et al.* 1977) using *Pfu* polymerase (Stratagene), sense primer P1 (GGGGGGTCATGAGTATTCAACATTTCCGTGTCG) (*Bsp*HI site underlined) and antisense primer P2 (CCGAGCTCTTGGTCTGACAGTTACCAATGC) (*Sac*I site underlined), and the following cycling-protocol; denaturation at 95° for 30 sec, annealing at 61°C for 30 sec and extension at 72°C for 90 sec repeated for 30 cycles, followed by a single cycle of 72°C for 10 minutes.

The resulting amplicon was digested with *Bsp*HI (New England Biolabs) and *Sac*I (New England Biolabs). The pACSE3 plasmid was digested with the same restriction enzymes and dephosphorylated with Calf Intestinal Phosphatase (New England Biolabs). Digested amplicon and vector were purified using Sigma's GenElute™ PCR Clean-up Kit, ligated using T4 DNA Ligase (New England Biolabs), and transformed into DH5αE by electroporation.

Mutagenesis: Mutant TEM alleles were created by introducing random mutations into the TEM-1 allele using the Genemorph I and Genemorph II PCR Mutagenesis Kits (Stratagene) according to the manufacturer's instructions. The primers used for the PCR were P3 (TCATCCGGCTCGTATAATGTGGA) and P4 (ACTCTCTTCCGGGCGCTATCAT), which flank the multiple cloning site of pACSE3.

The mutation rate in a Mutazyme-based error-prone PCR (ep-PCR) is determined by the number of duplications. As the yield of each PCR is approximately the same, higher mutation frequencies can be achieved by lowering the initial amount of template in the reaction. Conditions were set to introduce on average 2 (low mutation rate) or 6 (high mutation rate) mutations per amplicon per PCR.

The resulting mutated amplicons were digested with *Bsp*HI and *Sac*I restriction enzymes, ligated into pACSE3 and electroporated into DH5αE. After recovery for 90 min in SOC medium at 37°C, the cells were diluted in 500 ml LB-tetracycline medium. An aliquot was taken out of this directly after mixing and plated on LB-tetracycline to determine the library-size, while the remainder of the culture was placed overnight at 37°C to expand the library. An aliquot of the expanded library was stored in 10% glycerol at -80°C.

Site-directed mutagenesis: Specific mutations were built into the TEM-1 gene using the QuickChange site-directed mutagenesis kit (Stratagene), according to the manufacturer's instructions.

Artificial recombination: Where used, recombination was introduced by digestion of the error-prone PCR-amplicons with *Sac*I, *Bsp*HI and *Pvu*I endonuclease. *Pvu*I has a single recognition site halfway the TEM-1 sequence and digestion of the amplicons results in two fragments of approximately the same length. The truncated amplicons were subsequently restored to full-length sequences and placed in vector pACSE3 by a single ligation step as described previously. Because ligation is random this mimics a single recombination event at a fixed spot approximately halfway the TEM-1 gene, which was confirmed by pilot experiments using restriction markers.

Expression of TEM alleles: Expression of cloned TEM alleles in pACSE3 is under control of the *pTac* promoter that is tightly regulated by the lac repressor that is encoded by the *lacI* gene on the plasmid. Expression of TEM alleles was induced by adding 50 μM isopropyl-ß-D-thiogalactopyranoside (IPTG), a concentration that has previously been shown to mimic a natural expression level (BARLOW and HALL 2002).

Selection for extended-spectrum TEM-alleles: A series of bottles containing 50 ml Mueller Hinton medium was inoculated with Ctx in two-fold increments (ranging from 0.125 μg/ ml (the MIC for TEM-1) to 1024 μg/ ml). Subsequently, each bottle was inoculated with a number of cells from the overnight-enriched library that was equivalent to 10 times the library-size. Cultures were incubated for 48 hours at 37°C. The culture that grew at the highest concentration of Ctx was plated on LB + Tet. Because of clonal fixation, a single genotype is expected to dominate this culture (BARLOW and HALL 2002). This was confirmed by sequencing of 10 clones from the bottle with the highest Ctx concentration that still permitted visible growth from a random selection series. The next day a single colony was picked from this plate and grown overnight in LB + Tet. Plasmid from this overnight culture was subsequently isolated using the GenElute™ Plasmid Miniprep Kit (Sigma) and sequenced.

Sequencing: The TEM-allele was amplified from isolated plasmids using primers P3 and P4 and BigDye™ (Perkin Elmer) or DYEnamic™ ET (AP-biotech) Terminator Cycle Sequencing kits according to the manufacturer's instructions.

Antibiotics and MIC-assay: The following ß-lactam antibiotics were used in this study: ampicillin, cefotaxime and tetracycline (all Sigma). Stock solutions were made in H_2O (ampicillin), 0.1M $NaPO_4$, pH 7.0 (Ctx) and 70% ethanol (tetracycline).
MICs were determined from 150 µl cultures at a titre of 10^5 cells/ ml in Mueller Hinton II medium containing 50 µm IPTG. Another 150 µl of twofold serial dilutions of antibiotic in MH II were added to this. MIC is defined as the lowest concentration of antibiotic that completely prevents visible growth. Cultures were grown for 24 hours at 37°C, after which MIC was determined by visual inspection. Prior to each MIC-assay, plasmid was isolated from all selected clones and transformed into isogenic *E. coli* DH5αE cells by electroporation in order to exclude any phenotypic variation in the bacteria.

Statistical analysis: Because MIC values are measured on a discontinuous scale (using two-fold increases of Ctx concentration), non-parametric tests were used to compare differences in evolved MIC values among treatments. Statistical analysis was done using SPSS, rel. 15.0.1.1 (SPSS Inc., 2007).

2.3 Results

Exploring TEM-1's evolutionary heterogeneity towards increased cefotaxime resistance

We first performed an evolution experiment to explore the different mutational trajectories of TEM-1 towards increased Ctx resistance. Six independent error-prone PCRs were started each with a low mutation rate (LMR; ~2.0 mutations per amplicon) and a high mutation rate (HMR; ~6.1 mutations per amplicon). Three LMR and three HMR lines were subjected to recombination before selection while the remaining six lines were not allowed to recombine, resulting in four different treatments of three replicates each; (i) LMR no recombination (ii) LMR recombination (iii) HMR no recombination and (iv) HMR recombination. Each of the twelve mutagenized libraries was subsequently subjected to selection for increased resistance in a gradient of twofold increasing Ctx concentrations, and a random clone from the highest

Ctx concentration that allowed growth was selected for the next round of evolution. Three rounds of evolution were carried out. The average library-size over the whole experiment was ~1.1 x 10^6 transformants.

Figure 2.1 shows the average log relative change in Ctx resistance for all lines per treatment over the three rounds of evolution. Three conclusions can be drawn from this pattern. First, evolution leads to substantial adaptation, with changes of more than three orders of magnitude in relative MIC. Second, mutation rate and recombination do not affect the final level of adaptation (mutation rate: Mann-Whitney $U = 24$, $P = 0.295$; recombination: $U = 18$, $P = 1$). Third, the MIC levels do not significantly increase from the second to the third round of evolution (Sign test comparing MIC between round 2 and 3, $P = 0.754$; comparison of MIC between round 1 and 2, $P < 0.001$), suggesting that most lines approach a fitness peak.

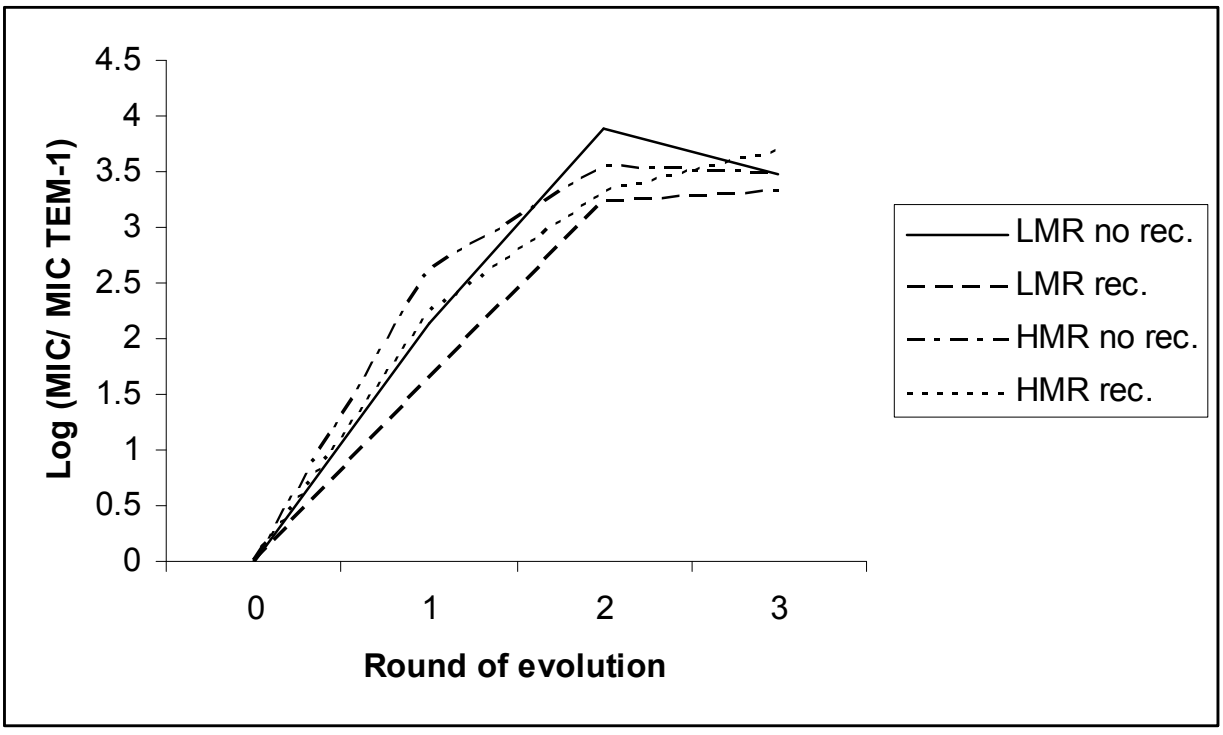


Figure 2.1: Average relative log change in cefotaxime resistance over the three evolutionary rounds.

The amino acid substitutions encountered in all twelve lines after the different rounds of evolution are shown in Table 2.1, together with the MIC values for all genotypes (see Figure 2.3 for the location, in the three-dimensional structure of TEM-1, of amino acid residues at which substitutions were observed with high frequency). Seven of the twelve lines seem to follow a preferred mutational pathway involving substitutions E104K, M182T and G238S

(from here on referred to as the 'G238S-based pathway'). The combination of these three substitutions has been reported as the evolutionary end-point in several other directed evolution studies on increased Ctx resistance of TEM-1 (ZACCOLO and GHERARDI 1999; ORENCIA *et al.* 2001; KOPSIDAS *et al.* 2007).

Five lines deviate from this pattern in that they lack one or two of the three substitutions present in the G238S-based pathway. The MIC levels of these lines are significantly lower than those of the seven lines that do contain the three substitutions (Mann-Whitney $U = 0$, $P = 0.002$), suggesting that they approach one or more peaks of lower fitness than the G238S-based pathway. We were interested whether sign epistatic interactions between early substitutions present in these lines and substitutions from the G238S-based pathway could be responsible for the observed deviations in mutational pathways and whether such substitutions could lead evolution to alternative adaptive peaks. We focussed on the three lines that contain only one of the three substitutions of the G238S-based pathway (lines 3, 4 and 7 in Table 2.1) and screened their first round alleles for candidate substitutions with such properties. Since lines 3 and 4 are the only ones lacking substitution G238S, we suspected that they harbour a substitution with a sign epistatic interaction with G238S. Substitution R164S, present in line 3, is expected to show such a sign epistatic interaction with G238S (ZACCOLO and GHERARDI 1999; GIAKKOUPI *et al.* 2000). At the same time, this substitution confers a considerable increased resistance when present alone (ZACCOLO and GHERARDI 1999) and therefore it might function as the first substitution of a putative alternative mutational pathway. In experimental line 4, substitution A237T is considered to be the most likely candidate, because the substitution (i) is physically close to G238S (ii) frequently occurs in clinical isolates and (iii) has never been found in combination with G238S in clinical isolates (http://www.lahey.org/studies/temtable.asp). A237T confers both a slightly increased MIC (GIAKKOUPI *et al.* 2001) and increased Kcat/ Km values for Ctx (CANTU *et al.* 1997) suggesting that this substitution is individually accessible to selection. We constructed the single mutants A237T and G238S as well as the double mutant A237T/ G238S. A comparison of the MICs of TEM-1 (0.0625 μg Ctx/ ml), the single mutants A237T (0.125 μg/ ml) and G238S (1.0 μg/ ml) and the double mutant A237T/ G238S (0.125 μg/ ml), suggests that A237T has a sign epistatic interaction with G238S while being individually adaptive. Like R164S, this substitution might therefore be a starting substitution of an alternative mutational pathway.

Table 2.1: Amino acid substitutions in evolved TEM alleles. Substitutions fixed after the first round are in **bold**, those fixed after the second round are in *italic* and those fixed after the third round are underlined.

DNA site[a]	Mutation	Amino acid substitution[b]											
		High mutation-frequency						Low mutation-frequency					
		No recombination			Recombination			No recombination			Recombination		
		1	2	3	4	5	6	7	8	9	10	11	12
22	G>A	**V10I**[d]			V10I		*V10I*						
49	T>A		*F19I*										
50	T>G						F19C						
51	T>A								F19L				
91	G>A										**V33I**[c]		
304	G>A	*E104K*	*E104K*			*E104K*	*E104K*		*E104K*	*E104K*	*E104K*	*E104K*	*E104K*
335	C>T												*T114M*
364	A>G											*S124G*	
428	C>A				*P145Q*								
439	A>G				**T149A**								
452	A>T	H153L											
459	G>A						*M155I*				**M155I**[c]		
484	C>A			**R164S**									
493	C>A			*P167T*									
509	C>T				**A172V**								
517	A>G				*N175D*								
530	A>G				*D179G*								
539	T>C	**M182T**	**M182T**	*M182T*	**M182T**	**M182T**	**M182T**		**M182T**	**M182T**	*M182T*		
545	C>T							**A184V**					
557	C>G												*T188R*
566	G>A							**R191H**					
593	C>G							*T200S*					
595	C>A				**L201I**								
697	T>A				S235T								
703	G>A				**A237T**								
706	G>A	**G238S**	**G238S**			**G238S**	**G238S**	**G238S**	**G238S**	**G238S**	**G238S**	**G238S**	**G238S**
709	G>A			**E240K**				*E240K*					
737	C>T								A249V				
782	C>T				*T265M*			*T265M*				*T265M*	**T265M**
788	G>C			G267A									
800	C>T				*T271I*								
829	G>T		**E281-**										
Round		**MIC value**[e]											
1		64	4	8	8	8	16	2	8	16	2	2	4
2		512	128	32	64	256	64	128	1024	256	128	128	64
3		256	256	64	128	512	256	64	256	256	256	64	64

[a] Position in the coding sequence, numbering according to Sutcliffe (1978).
[b] Substitutions are shown as the TEM-1 amino acid according to the IUPAC single-letter code (left), the position in the protein as numbered in Ambler *et al*. (1991) (middle), and the mutant amino acid (right). A dash at the position of the mutant amino acid indicates a (new) stop-codon.
[c] The mutation causing this substitution disappeared after the second round of evolution and selection
[d] The mutation causing this substitution disappeared after the third round of evolution and selection
[e] Median value across three replicates

Like line 3 and 4, line 7 deviates from the other lines in that it lacks both E104K and M182T. Because substitution G238S is present in line 7 in the first round of evolution, another first-round substitution must have been responsible for the deviating mutational pathway of this line. Of the two possible substitutions (A184V and R191H), we expected that A184V was

adaptive in the TEM-1 background carrying G238S, because we observed this combination after a single round of evolution in two pilot experiments before. To investigate whether the single mutants R164S and A237T or the double mutant A184V/ G238S would prevent the (other) substitutions of the G238S-based pathway from occurring, we constructed TEM-1 alleles carrying these substitutions and subjected them to three rounds of ep-PCR and selection for increased Ctx resistance with fivefold replication. The mutation rate was set to introduce on average ~2 mutations per amplicon per ep-PCR and extra recombination was omitted this time. The average library-size was ~5.9 x 10^4 for the libraries of mutant R164S, ~6.9 x 10^4 for the libraries of mutant A237T and ~1.0 x 10^5 for the libraries of double mutant A184V/ G238S. The average mutation rate was 2.23 (± 0.0) (R164S), 2.23 (± 0.0) (A237T) and 2.51 (±0.22) (A184V/ G238S). The amino acid substitutions and MIC-values recorded after each round of evolution and selection can be found in Tables 2.2 (R164S and A237T mutants) and 2.3 (A184V/ G238S double mutant).

MIC levels

Figure 2.2 shows the changes in mean MIC values for each of the three constructed mutants together with those of the seven lines following the G238S-based pathway of the previous experiment. Two conclusions can be drawn from these changes in MIC. First, after an initial rise (Sign test comparing all 15 lines for first two rounds, $P = 0.001$), MIC values level off (Sign test, $P = 0.453$), suggesting that they approach an adaptive peak. Second, despite their head start, the final resistance reached by these lines is significantly lower than that reached by the G238S-based lines of the previous experiment (R164S lines: Mann-Whitney $U = 35$, $P = 0.002$; A237T lines: $U = 35$, $P = 0.002$; A184V/G238S lines: $U = 35$, $P = 0.002$). The final MIC values do not vary significantly among the three types of lines (Kruskal-Wallis statistic = 3.404, $P = 0.182$), although the difference between the 10 single-substitution and the five double-substitution lines approaches significance ($U = 10.5$, $P = 0.065$). This suggests that evolution leads to one or more peaks that differ from the peak approached by the G238S-based lines. To verify whether these lines follow adaptive trajectories different from the G238S-based trajectories and leading to different local peaks, we then looked at the mutational pathways involved for each of the three constructed mutants.

R164S mutant

The absence of M182T and G238S in all five, and E104K in four of the five lines containing R164S (see Table 2.2) strongly suggests that R164S hampers the selection of these substitutions involved in the G238S-based pathway. The substitutions present in the final genotypes of the R164S mutant differ almost completely from the substitutions observed in the seven lines following the G238S-based pathway in the first experiment. The repeated observation in independent lines suggests that substitutions A237T, E240K, R241H and I173V are unique adaptive mutations in this background.

A237T mutant

None of the five lines carrying A237T acquires substitution G238S, while only in one line E104K is found after three rounds of evolution. However, different from the R164S background, M182T is found in four lines. A general pattern in all lines is that they contain substitutions in the functionally important Ω-loop (positions 161-179) after the initial round of evolution. Most of the substitutions again differ from the ones encountered in the lines following the G238S-based mutational pathway (see Table 2.1). Therefore, while the pathways observed here show some overlap with the G238S-based pathway and particularly with the R164S-based pathways (see Table 2.2) they also involve substitutions that are unique to this genetic background.

A184V/ G238S double mutant

Contrary to what was expected, E104K appears in three of the lines carrying A184V/ G238S. Substitution M182T remains absent in all experimental lines, indicating that this substitution has become selectively neutral or deleterious in the background of A184V and G238S, perhaps because A184V is functionally similar to M182T. When line 7 from Table 2.1 is included, three parallel changes can be observed next to E104K; I208M, E240K and T265M, none of which were observed in the seven lines following the G238S-based pathway (see Table 2.1). Therefore, although some mutations are shared with the G238S-based pathway, the mutational pathways followed by this background seem to lead to one or more adaptive peaks that are different from the one approached by the G238S-based pathway.

Table 2.2: Amino acid substitutions in evolved, mutant TEM-alleles containing R164S (left) and A237T (right). Substitutions present at the start of the experiment are in normal text and shaded, substitutions fixed after the first round are in **bold**, those fixed after the second round are in *italic* and those fixed after the third round are underlined.

DNA site[a]	Mutation	Amino acid substitution[b]									
		1	2	3	4	5	1	2	3	4	5
22	G > A					V10I					
65	T > G	*F24C*									
74	C > T										P27L
107	A > G		*D38G*								
110	A > G							*Q39R*			
257	A > T				*Q88L*						
284	A > T					**Y97F**					
294	T > G		*N100K*								
304	G > A				**E104K**						**E104K**
366	T > G	**S124R**									
379	A > T				M129L						
482	A > T								**D163V**		
484	C > T						**R164C**[c]				
484	C > A	R164S	R164S	R164S	R164S	R164S		**R164S**			
485	G > A						C164Y				**R164H**
487	T > G								*W165G*		
500	T > C									**L169P**	
511	A > G	*I173V*		I173V							
512	T > C							I173T			
517	A > T										*N175Y*
533	C > G		*T180S*								
539	T > C						**M182T**		**M182T**	**M182T**	*M182T*
544	G > A				A184T						
557	C > T			*T188I*							
568	A > G										*K192E*
570	A > T							K192N			
583	G > A	*E197K*									
596	T > C				*L201P*			L201P			
640	G > A				*V216I*					V216I	
643	G > T								**A217S**		
665	C > T				**A224V**		*A224V*		A224V		
697	T > A	*S235T*									
703	G > A		*A237T*	*A237T*		*A237T*	A237T	A237T	A237T	A237T	A237T
709	G > A	**E240K**	**E240K**	**E240K**		**E240K**		**E240K**			
713	G > A		*R241H*			*R241H*					
730	A > G				I247V						
737	C > T				A249V						
756	G > T			K256N							
782	C > T		*T265M*								
788	G > T				G267V						
790	A > G					S268G					
821	A > G				Q278R						
859	T > A	*end291K*									

Round	MIC value[d]									
1	4	8	4	16	8	4	32	1	8	4
2	16	64	32	16	64	32	32	8	8	64
3	16	32	32	16	128	64	32	16	16	64

[a] Position in the coding sequence, numbering according to Sutcliffe (1978)
[b] Substitutions are shown as the TEM-1 amino acid according to the IUPAC single-letter code (left), the position in the protein as numbered in Ambler *et al*. (1991) (middle), and the mutant amino acid (right)
[c] Changed into 164Y after the third round of evolution and selection
[d] Median value across three replicates in µg Ctx/ ml

Table 2.3: Amino acid substitutions in evolved, mutant TEM-alleles containing A184V and G238S. Substitutions present at the start of the experiment are in normal text and shaded, substitutions fixed after the first round are in **bold**, those fixed after the second round are in *italic* and those fixed after the third round are underlined.

		Amino acid substitution[b]				
DNA site[a]	Mutation	1	2	3	4	5
315	A > T					**D38V**
317	C > A				**Q39K**	
327	C > G				*A42G*	
345	A > T	**E48V**				
347	C > A					**L49M**
512	G > A	**E104K**	**E104K**			*E104K*
534	A > T		*K111M*			
660	A > G		*H153R*			
661	C > G			H153Q		
725	A > G					N175D
753	C > T	A184V	A184V	A184V	A184V	A184V
826	A > G		**I208M**	I208M		
839	G > C			**A213P**		
914	G > A	G238S	G238S	G238S	G238S	G238S
917	G > A			E240K	**E240K**	
990	C > T			**T265M**		

Round	MIC values[c]				
1	32	32	8	64	8
2	64	128	8	128	32
3	32	128	128	128	32

[a] Position in the coding sequence, numbering according to Sutcliffe (1978)
[b] Substitutions are shown as the TEM-1 amino acid according to the IUPAC single-letter code (left), the position in the protein as numbered in Ambler *et al*. (1991) (middle), and the mutant amino acid (right)
[c] Median value across three replicates in µg Ctx/ ml

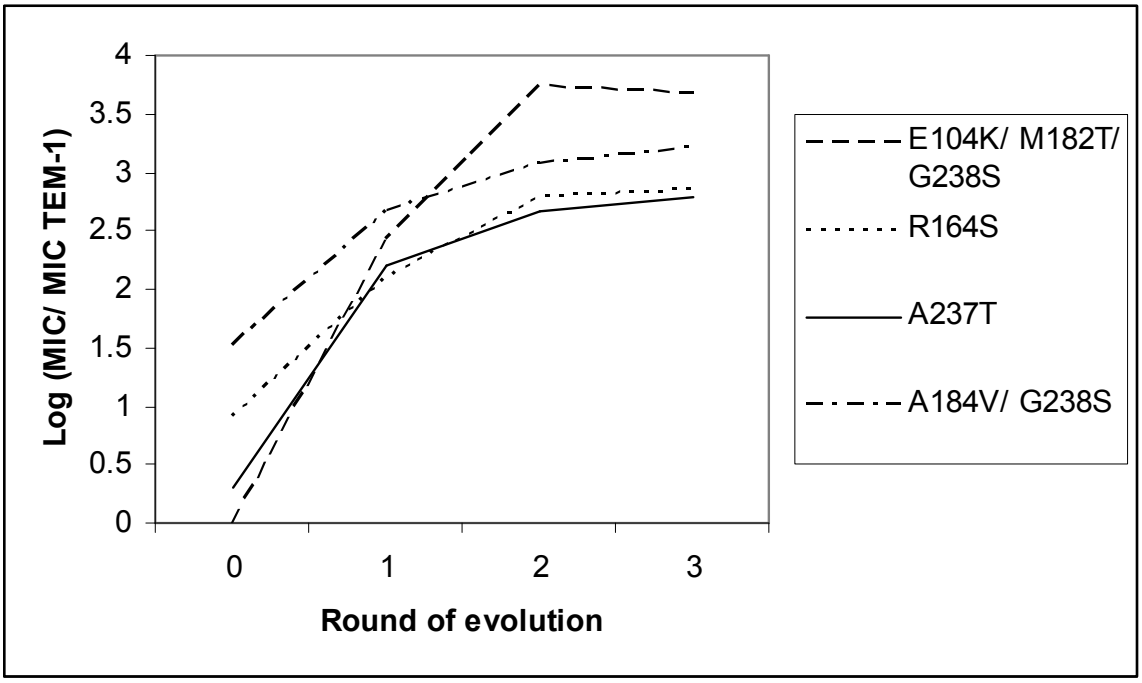


Figure 2.2: Average relative log change in cefotaxime resistance over the three evolutionary rounds for the seven E104K/ M182T/ G238S containing lines from table 2.1 and the mutants containing R164S, A237T and A184V/ G238S.

2.4 Discussion

The exact shape of the fitness landscape has important implications for many questions in evolutionary biology, including the causes of evolutionary constraints, the evolution of diversity and speciation, and the evolutionary significance of sex and recombination. Here we showed that, using *in vitro* evolution, the fitness landscape of TEM-1 ß-lactamase is rugged for adaptation to a novel antibiotic. We found that while most experimental lines acquire three amino acid substitutions leading to a high level of resistance, some lines acquire other substitutions and reach lower levels of adaptation. When we introduced some of the early substitutions observed in the lines with exceptional substitutions and selected these alleles for increased resistance, the three substitutions of the preferred pathway were avoided in most cases. These results therefore indicate that sign epistatic interactions between early mutations that individually increase resistance can make evolution contingent upon early substitutions, which can drive evolution towards different adaptive peaks. Which peak is approached depends on the epistatic properties of early substitutions. We found that escape from local peaks was not possible, despite our high mutation rate which allowed multiple mutations to fix in a single round of evolution. Finally, we showed that not only the first, but also later substitutions are involved in negative interactions with subsequent mutations.

What exactly is the highest fitness peak of TEM-1 for increased cefotaxime resistance?
As indicated before, the main difference between the most Ctx resistant mutants identified in this study and the most Ctx resistant mutant known from literature (which was identified by DNA-shuffling (STEMMER 1994) and subsequent step-by-step reconstruction of evolutionary intermediates (HALL 2002)) is the absence of substitution A42G in our study. According to literature, addition of A42G to an E104K/ M128T/ G238S background should increase Ctx MIC levels 8- to 32-fold (HALL 2002; WEINREICH *et al.* 2006). However, the A42G substitution does not occur in the background of these three substitutions in our experiments and even a single round of ep-PCR and selection using the triple mutant E104K/ M182T/ G238S as template in fivefold replicate did not yield this substitution. Apart from the DNA-shuffling experiment by Stemmer (1994), none of the experimental or directed evolution studies on increased Ctx resistance of TEM-1 has identified the fourfold mutant containing A42G, though several studies have identified the E104K/ M182T/ G238S triple mutant that

was frequently encountered in our experiments (ZACCOLO and GHERARDI 1999; ORENCIA *et al.* 2001; KOPSIDAS *et al.* 2007). All in all, it seems that in our experimental setup and in that of most others, the peak for Ctx resistance is the E104K/ M182T/ G238S triple mutant. Although minor adaptive substitutions may be added in this background (notice the parallel changes at residues 10 and 19 in the background of the triple mutants in Table 2.1 for example), they do not seem to noticeably increase Ctx resistance and therefore it seems reasonable to assume that E104K/ M182T/ G238S is approximately the evolutionary endpoint when selecting for increased Ctx resistance.

Height of the alternative fitness peaks

The MIC levels reached by both the R164S and A237T mutants (~32 μg/ ml) are considerably lower than those of the E104K/ M182T/ G238S-containing mutants in experiment I (~256 μg/ ml). This indicates that whatever substitution is selected first can not only force evolution onto different evolutionary trajectories, but also lead evolution to lower fitness peaks. Interestingly, clinical isolates of R164S/ A237T/ E240K (TEM-5) and E104K/ M182T/ G238S (TEM-52) exist. Although no direct comparison has been made between these two alleles, TEM-5 has been compared to TEM-3 (Q39K/ E104K/ G238S) in an isogenic *E. coli* background (BRADFORD and SANDERS 1995). Since Q39K is generally assumed to be neutral (KNOX 1995), TEM-3 is comparable to the E104K/ G238S double mutant and despite the fact that this mutant does not contain M182T, its MIC for Ctx is already much higher than that of the R164S/ A237T/ E240K mutant (32 vs 4 μg/ ml) (BRADFORD and SANDERS 1995). This strengthens our conclusion that the adaptive peak for Ctx resistance reached by the G238S-based route is considerably higher than those reached by the R164S- and A237T-based routes.

Accessibility of the alternative pathways

In nature, mutations most probably go to fixation one at a time, assuming that the time to fixation for a particular mutation is much lower than the time between the occurrence of successive mutations (GILLESPIE 1984). This is in contrast with our experimental setup where often two or three mutations go to fixation at the same time. In the presence of sign epistatic interactions between these mutations (for example when two mutations have a neutral or negative effect on fitness when individually present, but a positive effect when combined) this could uncover pathways that are not accessible to standard 'one-by-one' selection. Previous

analysis has shown that the G238S-based pathway, which was repeatedly found in our initial experiment, is accessible to Darwinian selection, although the order in which mutations appear is severely constrained (WEINREICH *et al.* 2006). Below, we briefly address the accessibility of the general pathway found for the R164S mutant, illustrated by data on evolutionary intermediates of this pathway that have previously been described.

Substitution R164S is known to increase Ctx resistance when present in the TEM-1 background (SOWEK *et al.* 1991; BLAZQUEZ *et al.* 1998; ZACCOLO and GHERARDI 1999; WANG *et al.* 2002), although the effect is less strong than that of G238S. It is known that addition of either E104K, M182T, A237T or E240K can further increase Ctx resistance in an R164S background (BLAZQUEZ *et al.* 1998; ZACCOLO and GHERARDI 1999). In agreement with this, E104K and E240K are observed in the background of R164S after the first round of evolution (once and four times respectively). Addition of A237T to either of these double mutants is also known to increase Ctx resistance (BLAZQUEZ *et al.* 1998; GIAKKOUPI *et al.* 2001) and is observed in three of the four experimental lines containing R164S and E240K. Although we stress that more adaptive substitutions are present in our experiments (e.g. substitutions R241H and T265M appear independently in multiple lines), it is unknown whether these substitutions are part of accessible evolutionary pathways. Similar evidence for accessible evolutionary pathways in the A237T- and A184V/ G238S-based routes exists, on which we will not further elaborate here.

R164S- and A237T-based pathways; one or two fitness peaks?

When comparing the sequence data of the R164S- and A237T-based evolutionary trajectories towards increased Ctx resistance (Table 2.3) it is clear that the substitutions do not exclude each other; A237T appears in the background of R164S three times, while R164S appears in the background of A237T once (notice that in the background of A237T, two alternative substitutions occur at residue 164 as well). Therefore, it seems reasonable to ask the question whether the R164S- and A237T-based pathways ascend a single adaptive peak. Experimental line 2 of the A237T mutant (Table 2.3) contains several of the substitutions observed in the experiment with the R164S mutant (R164S, E240K), which suggests that a similar peak is reached while only the sequence of mutational events differs. However, the substitutions found in the other experimental lines of mutant A237T differ markedly from those reported for the R164S-based genotypes. For example, they all contain M182T which is absent in the

R164S-mutant lines. In fact, even if we consider the alternative substitutions at residue 164 (lines 1 and 5) equal to R164S, experimental lines 3 and 4 still show totally different mutations in the Ω-loop. After three rounds of evolution, these experimental lines show a considerable increase in MIC, but the substitutions they contain are completely different from those observed for the R164S-based pathway. This suggests that more than one accessible A237T-based pathway towards increased Ctx resistance exists for TEM-1.

More constraints on TEM-1 evolution

Clear indications for yet another constraint at the level of the second mutation (as found in the experiment on the A184V/ G238S mutant) exist. In the dataset of the R164S-mutant, four experimental lines fix E240K after the first round of evolution. Experimental line 4 however, contains E104K after the first round but lacks E240K even after the third round. This is consistent with the previous observation that, even though both E104K and E240K enhance Ctx resistance in the background of R164S, the triple mutant E104K/ R164S/ E240K has a lower MIC for Ctx than both double-mutants E104K/ R164S and R164S/ E240K (SOWEK *et al.* 1991). Since the additional substitutions gathered by line 4 differ greatly from those observed in the other four experimental lines, this seems to be a clear example of yet another constraint on evolution, caused by a sign epistatic interaction, this time at the level of the second substitution.

Clinical relevance

Here, we have focussed exclusively on the evolution of TEM-1 for increased Ctx resistance. In the 'natural' setting of TEM-1 and its derivatives, hospitals and clinics worldwide, many more antibiotic pressures are encountered (MEDEIROS 1997). The great variability in mutational trajectories observed in our experiments creates the possibility that in a setting with variable ß-lactam pressure different mutants may be successful due to their different pleiotropic effects with respect to various antibiotics. For example, even though the triple mutant E104K/ M182T/ G238S confers the highest Ctx resistance encountered here, it may show much less of an increase or even a decrease in resistance when challenged with a different ß-lactam. Conversely, the triple-mutant R164S/ A237T/ E240K, though seemingly stuck at a suboptimal peak concerning Ctx resistance, may 'beat' the E104K/ M182T/ G238S mutant when it concerns another ß-lactam. Such pleiotropic effects are nicely illustrated by

the side-by-side comparison of the resistance spectra of the R164S and G238S mutants; while G238S indeed gives a higher increase in MIC for Ctx, the situation is reversed for ceftazidime (another 3rd generation cephalosporin) (GIAKKOUPI *et al.* 2000). Thus, temporal and spatial variation in the application of antibiotics may change the fate of TEM alleles with suboptimal adaptation to Ctx.

A structural explanation for the sign epistatic interaction between R164S and G238S

A previous analysis of the interaction between R164S and G238S (ZACCOLO and GHERARDI 1999) revealed that addition of R164S decreased the MIC for Ctx in a series of backgrounds containing G238S, while vice versa the addition of G238S appeared to be neutral in all these backgrounds except in the background of R164S itself (Table 2.4). These data suggest that R164S, presumably because it is translated first, dominates the folding of ß-lactamase, thereby neutralizing the effect of G238S.

Table 2.4: Effect of addition of R164S and G238S to various genetic backgrounds. Data reproduced from Zaccolo and Gherardi (1999). Additional data from Giakkoupi *et al.* (2000) are in parentheses.

			Relative increase in MIC compared to genotype reported on the left	
			Addition of R164S	*Addition of G238S*
TEM-1			20 (17)	78 (33)
E104K			4	16
R164S				-4 (-8)
M182T			625	1250
G238S			-16 (-17)	
E104K	R164S			0
E104K	G238S		-32	
M182T	R164S			0
M182T	G238S		-2	
E104K	R164S	M182T		0
E104K	M182T	G238S	-32	

Detailed analysis of the effect of various substitutions at residue 238 on kinetic parameters for different antibiotics indicates that the most likely effect of G238S is alteration of the active site configuration by repositioning of the Ω-loop. The larger Ser-238 side chain can be positioned towards the Ω-loop as a result of a new hydrogen bond of Ser-238 to Asn-170, causing perturbation of the loop and of the deacylation residue Glu-166 (CANTU and PALZKILL 1998). The most widely accepted explanation for the structural effect of R164S also involves perturbation of the Ω-loop and Asn-170; X-ray crystallography of a mutant TEM containing E104K, R164S and M182T has revealed that substitution R164S introduces a

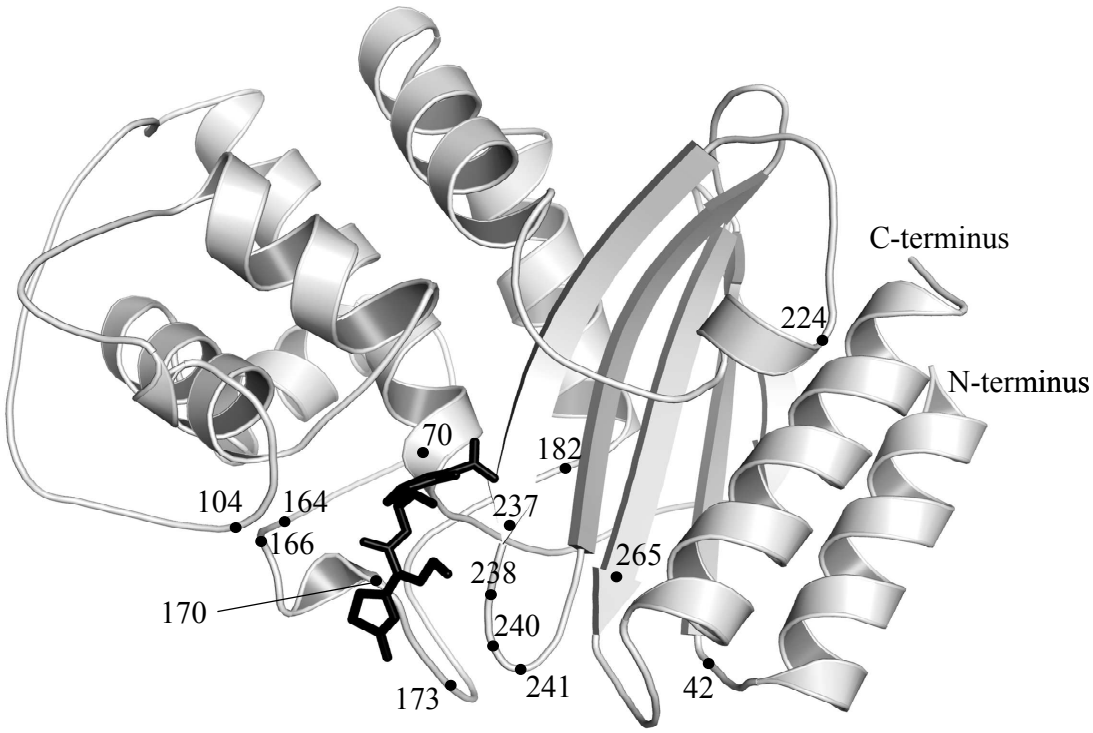


Figure 2.3: Three-dimensional structure of TEM-1 ß-lactamase (PDB-code: 1BTL) with a cefotaxime molecule modelled in the binding site. Some important amino acid residues mentioned in the text have been indicated. The active site residue at position 70 has also been indicated. Residue numbering according to Ambler *et al.* (1991).

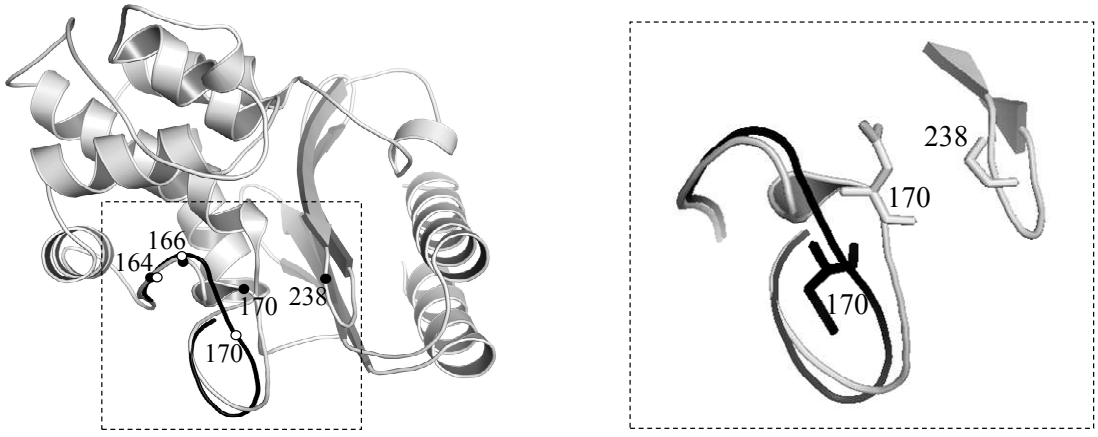


Figure 2.4: *Left*: the three-dimensional structure of TEM-1 ß-lactamase (PDB-code: 1BTL) with in black the Ω-loop of a TEM-mutant containing substitution R164S (PDB-code: 1JWZ, this mutant contains E104K and M182T additionally). *Right*: a magnification of the boxed area of the left figure showing the two Ω-loops and the repositioning of the asparagine at residue 170 compared to the glycine at residue 238. Residue numbering according to Ambler *et al.* (1991).

packing defect in the Ω-loop, due to loss of hydrogen bonds to Glu-171 and Asp-179 (Figure 2.4). As a result, Asn-170 flips by almost 180° and fills the cavity created by substitution of the large Arg-164 by the smaller Ser. In turn, the small 3_{10} helix (168-170) unwinds and leaves a cavity in the active site, which is thought to create space for the more bulky oxyimino side chains of third generation cephalosporins like Ctx (WANG *et al.* 2002). These observations provide an explanation for the negative epistatic interaction between R164S and G238S; if the structural changes brought about by substitutions of Gly-238 are not only the result of steric hindrance between Ser-238 and the Ω-loop, but also depend on formation of a new hydrogen bond between this residue and Asn-170 on the Ω-loop in order to stabilize the new configuration (as suggested in Cantu and Palzkill 1998), substitution G238S may lose its effect when the position of the Ω-loop and Asn-170 has already been altered by another substitution like R164S. This could either be due to loss of the stabilizing hydrogen bond between Ser-238 and Asn-170 as a result of repositioning of the Ω-loop, or because this loop is repositioned in such a way by R164S that substitution G238S no longer results in steric conflict.

In conclusion, we have shown that sign epistatic interactions can give rise to multiple peaks of different height in the fitness landscape of a single gene, which poses constraints on the mutational pathways followed by replicate lines. These interactions do not only occur at the level of the first substitution, but can also occur later on. The fact that such interactions do not seem to be rare suggests that the ruggedness of the fitness landscape is pervasive. A rugged fitness landscape has so far only been demonstrated using experimental evolution of a phage (BURCH and CHAO 2000), but indirect evidence exists indicating that ruggedness may be a more general feature of fitness landscapes (KORONA *et al.* 1994; ROZEN 2008). Although we can not extrapolate our findings to the level of organisms, contrary to the aforementioned studies, we can make a clear link between genotype and phenotype and in this way demonstrate that sign epistatic interactions are responsible for the observed ruggedness of the fitness landscape.

Acknowledgements

We thank Daniel M. Weinreich for critical comments to this manuscript and Bertha Koopmanschap for help with the molecular work and MIC-assays.

References

AMBLER, R. P., A. F. W. COULSON, J. M. FRERE, J. M. GHUYSEN, B. JORIS *et al.* 1991. A Standard Numbering Scheme for the Class-a Beta-Lactamases. Biochem. J. **276:** 269-270.

BARLOW, M., and B. G. HALL. 2002. Predicting evolutionary potential: in vitro evolution accurately reproduces natural evolution of the tem beta-lactamase. Genetics **160:** 823-832.

BARLOW, M., and B. G. HALL. 2003. Experimental prediction of the natural evolution of antibiotic resistance. Genetics **163:** 1237-1241.

BLAZQUEZ, J., M. C. NEGRI, M. I. MOROSINI, J. M. GOMEZ-GOMEZ and F. BAQUERO. 1998. A237T as a modulating mutation in naturally occurring extended-spectrum TEM-type beta-lactamases. Antimicrob. Agents Chemother. **42:** 1042-1044.

BOLIVAR, F., R. L. RODRIGUEZ, M. C. BETLACH and H. W. BOYER. 1977. Construction and Characterization of New Cloning Vehicles .1. Ampicillin-Resistant Derivatives of Plasmid Pmb9. Gene **2:** 75-93.

BRADFORD, P. A., and C. C. SANDERS. 1995. Development of Test Panel of Beta-Lactamases Expressed in a Common Escherichia-Coli Host Background for Evaluation of New Beta-Lactam Antibiotics. Antimicrob. Agents Chemother. **39:** 308-313.

BURCH, C. L., and L. CHAO. 2000. Evolvability of an RNA virus is determined by its mutational neighbourhood. Nature **406:** 625-628.

CAMPS, M., A. HERMAN, E. LOH and L. A. LOEB. 2007. Genetic constraints on protein evolution. Crit. Rev. Biochem. Mol. Biol. **42:** 313-326.

CAMPS, M., J. NAUKKARINEN, B. P. JOHNSON and L. A. LOEB. 2003. Targeted gene evolution in Escherichia coli using a highly error-prone DNA polymerase I. Proc. Natl. Acad. Sci. U. S. A. **100:** 9727-9732.

CANTU, C., W. Z. HUANG and T. PALZKILL. 1997. Cephalosporin substrate specificity determinants of TEM-1 beta-lactamase. J. Biol. Chem. **272:** 29144-29150.

CANTU, C., and T. PALZKILL. 1998. The role of residue 238 of TEM-1 beta-lactamase in the hydrolysis of extended-spectrum antibiotics. J. Biol. Chem. **273:** 26603-26609.

COLEGRAVE, N., and A. BUCKLING. 2005. Microbial experiments on adaptive landscapes. BioEssays **27:** 1167-1173.

FUJII, R., M. KITAOKA and K. HAYASHI. 2004. One-step random mutagenesis by error-prone rolling circle amplification. Nucleic Acids Res. **32:** -.

GAVRILETS, S. 2004. *Fitness landscapes and the origin of species*. Princeton University Press, Princeton, NJ.

GIAKKOUPI, P., A. M. HUJER, V. MIRIAGOU, E. TZELEPI, R. A. BONOMO *et al.* 2001. Substitution of Thr for Ala-237 in TEM-17, TEM-12 and TEM-26: alterations in beta-lactam resistance conferred on Escherichia coli. FEMS Microbiol. Lett. **201:** 37-40.

GIAKKOUPI, P., E. TZELEPI, P. T. TASSIOS, N. J. LEGAKIS and L. S. TZOUVELEKIS. 2000. Detrimental effect of the combination of R164S with G238S in TEM-1 beta-lactamase on the extended-spectrum activity conferred by each single mutation. J. Antimicrob. Chemother. **45:** 101-104.

GILLESPIE, J. H. 1984. Molecular Evolution over the Mutational Landscape. Evolution **38:** 1116-1129.

HALL, B. G. 2002. Predicting evolution by in vitro evolution requires determining evolutionary pathways. Antimicrob. Agents Chemother. **46:** 3035-3038.

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.

KONDRASHOV, F. A., and A. S. KONDRASHOV. 2001. Multidimensional epistasis and the disadvantage of sex. Proc. Natl. Acad. Sci. U. S. A. **98:** 12089-12092.

KOPSIDAS, G., R. K. CARMAN, E. L. STUTT, A. RAICEVIC, A. S. ROBERTS *et al.* 2007. RNA mutagenesis yields highly diverse mRAN libraries for *in vitro* protein evolution. BMC Biotechnol. **7**.

KORONA, R., C. H. NAKATSU, L. J. FORNEY and R. E. LENSKI. 1994. Evidence for Multiple Adaptive Peaks from Populations of Bacteria Evolving in a Structured Habitat. Proc. Natl. Acad. Sci. U. S. A. **91:** 9037-9041.

LUNZER, M., S. P. MILTER, R. FELSHEIM and A. M. DEAN. 2005. The biochemical architecture of an ancient adaptive landscape. Science **310:** 499-501.

MEDEIROS, A. A. 1997. Evolution and dissemination of beta-lactamases accelerated by generations of beta-lactam antibiotics. Clin. Infect. Dis. **24:** S19-S45.

MILLER, S. P., M. LUNZER and A. M. DEAN. 2006. Direct demonstration of an adaptive constraint. Science **314:** 458-461.

ORENCIA, M. C., J. S. YOON, J. E. NESS, W. P. STEMMER and R. C. STEVENS. 2001. Predicting the emergence of antibiotic resistance by directed evolution and structural analysis. Nat. Struct. Biol. **8:** 238-242.

POELWIJK, F. J., D. J. KIVIET and S. J. TANS. 2006. Evolutionary potential of a duplicated repressor-operator pair: Simulating pathways using mutation data. Plos Computational Biology **2:** 467-475.

POELWIJK, F. J., D. J. KIVIET, D. M. WEINREICH and S. J. TANS. 2007. Empirical fitness landscapes reveal accessible evolutionary paths. Nature **445:** 383-386.

ROZEN, D. E., M.G.J.L. HABETS, A. HANDEL AND J.A.G.M. DE VISSER. 2008. Heterogeneous adaptive trajectories of small populations on complex fitness landscapes. Plos One **3:** e1715.

SOWEK, J. A., S. B. SINGER, S. OHRINGER, M. F. MALLEY, T. J. DOUGHERTY *et al.* 1991. Substitution of Lysine at Position 104 or 240 of Tem-1ptz18r Beta-Lactamase Enhances the Effect of Serine-164 Substitution on Hydrolysis or Affinity for Cephalosporins and the Monobactam Aztreonam. Biochemistry **30:** 3179-3188.

STEMMER, W. P. 1994. Rapid evolution of a protein in vitro by DNA shuffling. Nature **370:** 389-391.

SUTCLIFFE, J. G. 1978. Nucleotide-Sequence of Ampicillin Resistance Gene of Escherichia-Coli Plasmid Pbr322. Proc. Natl. Acad. Sci. U. S. A. **75:** 3737-3741.

VAKULENKO, S., and D. GOLEMI. 2002. Mutant TEM beta-lactamase producing resistance to ceftazidime, ampicillins, and beta-lactamase inhibitors. Antimicrob. Agents Chemother. **46:** 646-653.

VAKULENKO, S. B., B. GERYK, L. P. KOTRA, S. MOBASHERY and S. A. LERNER. 1998. Selection and characterization of beta-lactam-beta-lactamase inactivator-resistant mutants following PCR mutagenesis of the TEM-1 beta-lactamase gene. Antimicrob. Agents Chemother. **42:** 1542-1548.

WANG, X. J., G. MINASOV and B. K. SHOICHET. 2002. Evolution of an antibiotic resistance enzyme constrained by stability and activity trade-offs. J. Mol. Biol. **320:** 85-95.

WEINREICH, D. M., N. F. DELANEY, M. A. DEPRISTO and D. L. HARTL. 2006. Darwinian evolution can follow only very few mutational paths to fitter proteins. Science **312:** 111-114.

WEINREICH, D. M., R. A. WATSON and L. CHAO. 2005. Perspective: Sign epistasis and genetic constraint on evolutionary trajectories. Evolution **59:** 1165-1174.

WHITLOCK, M. C., P. C. PHILLIPS, F. B. G. MOORE and S. J. TONSOR. 1995. Multiple Fitness Peaks and Epistasis. Annu. Rev. Ecol. Syst. **26:** 601-629.

WRIGHT. 1932. The roles of mutation, inbreeding, crossbreeding and selection in evolution. Proceedings of th sixth international congress of genetics **I:** 356-366.

ZACCOLO, M., and E. GHERARDI. 1999. The effect of high-frequency random mutagenesis on in vitro protein evolution: A study on TEM-1 beta-lactamase. J. Mol. Biol. **285:** 775-783.

-Chapter 3-

A comparison of the natural and laboratory evolution of TEM-1 ß-lactamase

M. Salverda[1], J.A.G.M. de Visser[1] & M. Barlow[2]

[1]Laboratory of Genetics, Wageningen University, the Netherlands

[2]School of Natural Sciences, University of California Merced, U.S.A.

Abstract: TEM-1 ß-lactamase is one of the most notorious and widespread antibiotic resistance determinants around. It confers resistance to penicillins and early cephalosporins and has shown an astonishing functional plasticity in response to the introduction of novel drugs derived from these antibiotics. Ever since its discovery in the 1960s, over 150 variants of TEM-1 - differing in their amino acid sequence and often in their resistance phenotype - have been isolated in hospitals and clinics worldwide. Next to this well-documented 'natural' evolution, the *in vitro* evolution of TEM-1 has been the focus of attention of many experimental and directed evolution studies. In this review we compare the substitutions found in clinical TEM-isolates to those found in laboratory studies in order to address the question to what extent the evolution of antibiotic resistance can be repeated or even predicted under laboratory conditions. We find that laboratory evolution has very accurately repeated the natural evolution of TEM-1 at many residues, and we present a list of substitutions found frequently in experimental isolates that have not (yet) been found in clinical isolates, but may predict their future evolution.

3.1 Introduction

The TEM ß-lactamases are among the best-studied antibiotic resistance enzymes around. They are the main cause of bacterial resistance to penicillins, cephalosporins and related antibiotics and they are found at high frequencies in hospitals and clinics around the world (MEDEIROS 1997; MATAGNE *et al.* 1998). TEM-1, the ancestor of this family, was first isolated from penicillin-resistant bacteria in the 1960s (DATTA and KONTOMICHALOU 1965). In response to the introduction of new antibiotics, TEM-1 has given rise to over 150 descendants ever since. All of these descendants differ in amino acid sequence and in most cases they also differ in resistance phenotype (the degree of resistance to different antibiotics). While TEM-1 only confers resistance to penicillins and early cephalosporins, the resistance spectrum of the TEM-family of ß-lactamases has fanned out to 2nd and 3rd generation cephalosporins, monobactams and penicillin-inhibitor combinations. Typically, within three years after the introduction of a new antibiotic, TEM-1 descendants have evolved the ability to hydrolyze the new drug (MEDEIROS 1997).

The detailed description of the natural evolution of TEM-alleles combined with the experimental ease of selecting for increased antibiotic resistance has made TEM-1 an interesting tool for directed evolution studies, where it has been used to explore the potential of novel *in vitro* evolution methods for protein engineering (STEMMER 1994; ZACCOLO and GHERARDI 1999; LONG-MCGIE *et al.* 2000; CAMPS *et al.* 2003; FUJII *et al.* 2004; KOPSIDAS *et al.* 2007). The remarkable results of the first directed evolution experiment on TEM-1 (STEMMER 1994) did not go unnoticed, and the potential of using *in vitro* evolution to repeat or even predict the natural evolution of TEM-1 was recognized soon after (VAKULENKO *et al.* 1998). At a later stage, *in vitro* evolution protocols were further fine-tuned in order to mimic natural evolution more accurately (BARLOW and HALL 2002).

A question that remains is to what extent laboratory evolution can repeat or predict the natural evolution of antibiotic resistance. One of the main problems in addressing this question is that, contrary to the controlled conditions in the laboratory, the exact selective conditions in nature are almost always unknown. In recent years however, the spectrum of selective agents used in experimental studies has broadened to such an extent that a direct comparison between clinical and laboratory isolates of TEM-alleles seems legitimate (see Table 3.2). Moreover, both the number of clinically isolated TEM-alleles described and the

amount of data on the laboratory evolution of TEM-1 has grown enormously.

Here, we compare the amino acid substitutions found in all known clinical TEM-isolates to those found in 17 different laboratory experiments that have looked at the directed or experimental evolution of TEM-1. We analyze to what extent laboratory evolution can be used to repeat the natural evolution of TEM-1 ß-lactamase. In addition, we also try to identify substitutions that are frequently encountered in laboratory isolates but that have not been identified in clinical isolates, in order to predict substitutions that might be encountered in clinical isolates yet to be identified. Since a lot of substitutions occur in just a few clinical isolates and since their effect is often unknown, we also use both datasets to make an inventory of exactly which substitutions are involved the in extension of the resistance phenotype of TEM-1.

Terminology

TEM-variants that differ in their amino acid sequence are referred to as 'TEM-alleles' here, although this does not necessarily imply a difference in resistance phenotype. We use the term 'clinical isolates' to describe the ~150 different TEM-alleles that have been isolated from patients in hospitals and clinics worldwide. Since the variation present in clinical isolates of TEM-alleles is the direct result of genetic drift and natural selection for resistance to (novel) antibiotics, we here use the term 'natural evolution' to describe the process by which these clinical isolates have arisen. Mutant TEM-alleles have also been identified in studies that mimic this process under laboratory conditions, for which we here use the term 'laboratory evolution'. We distinguish two types of laboratory evolution; 'experimental evolution' (which uses *in vivo* evolution by growing bacteria carrying TEM-1 in the presence of increasing amounts of antibiotics) and 'directed evolution' (which relies on *in vitro* mutagenesis techniques to create variation in TEM-1 and subsequently subjects the mutant alleles to *in vivo* selection). Specific substitutions are here referred to by the TEM-1 amino acid according to the IUPAC single-letter code, the number of the amino acid residue (AMBLER *et al.* 1991) and the respective mutant amino acid (e.g.; 'G238S' indicates that the glycine residue at position 238 was substituted for a serine residue).

Data included

Clinical isolates

The data on clinical TEM-isolates were collected from the Lahey database (http://www.lahey.org/studies/temtable.asp, as updated on June 10th 2008), which contains an up to date list of the amino acid substitutions present in all known (~150) clinical TEM-isolates with a unique amino acid sequence. Two alleles (TEM-41 and TEM-69) were left out of the analysis since their sequences have been withdrawn. Three other alleles (TEM-116, 157 and 162) contain substitutions V84I and A184V, which have been introduced into TEM-1 by site directed mutagenesis in order to remove restriction sites (SHORT *et al.* 1988; CHAIBI *et al.* 1997). The TEM-allele that contains both substitutions has been identified as a contamination in Taq-polymerase (KONCAN *et al.* 2007). Both substitutions were therefore ignored in the analysis, but other substitutions occurring in their background in TEM-157 and TEM-162 have been included.

Laboratory isolates

The data of seventeen laboratory studies have been included in this review (Table 3.1), and they are divided in four different categories: (i) twelve studies selecting for increased resistance to one or more ß-lactam antibiotics, (ii) two studies selecting for increased resistance to a ß-lactam - inhibitor combination, (iii) one study selecting for retention of ampicillin resistance and regain of this resistance after the introduction of a single deleterious mutation and (iv) two studies selecting for restored ampicillin resistance after structural perturbation.

Not all studies have used TEM-1 as the starting point of evolution and selection. Some of the studies have used the previously described TEM-variant that contains substitutions V84I and A184V. Reverse substitutions at these amino acid residues have been ignored in this case. One study has partially used TEM-variants containing a single adaptive substitution (Salverda and de Visser unpublished), while another study partially used a TEM-variant containing a single maladaptive substitution(BERSHTEIN *et al.* 2008). In both cases, these substitutions were not counted. The two structural perturbation studies differ from the other studies analyzed here in that they first used circular permutation (OSUNA *et al.* 2002) and N-terminal truncation (HECKY and MULLER 2005) to create crippled TEM-1 variants that had lost their native ampicillin resistance. Directed evolution was subsequently used to select for

Table 3.1: Laboratory studies included.

Reference	# Alleles [a]	Selective agent(s) [b]	Experimental technique
Antibiotic resistance			
Stemmer 1994	1	Ctx	Directed evolution; DNA-shuffling
Zaccolo and Gherardi 1999	7	Ctx	Directed evolution; high-frequency random mutagenesis
Long-McGie *et al.* 2000	2	Ctx	Experimental evolution; phagemid in hypermutator *E. coli* strains KH1366 and VS03
Blazquez *et al.* 2000	18	Caz, Amx	Experimental evolution; *E. coli* strain K12
Orencia *et al.* 2001	1	Ctx	Directed evolution; hypermutator *E. coli* combined with DNA-shuffling
Barlow and Hall 2002	9	Caz, Ctx, Cxm, Atm [c]	Directed evolution; error-prone PCR (Mutazyme I)
Barlow and Hall 2003	8	Fep	Directed evolution; error-prone PCR (Mutazyme I)
Camps *et al.* 2003	2	Atm	Directed evolution; *in vivo* error-prone DNA-replication
Fujii *et al.* 2004	3	Caz	Directed evolution; error-prone Rolling Circle Amplification
Holloway *et al.* 2007	4	Amp, Ctx	Experimental evolution; *E. coli* strain W3110
Kopsidas *et al.* 2007	1	Ctx	Directed evolution; error-prone RNA-replication
Salverda *et al.* unpublished	64	Ctx	Directed evolution; error-prone PCR (Mutazyme I & II)
Antibiotic-inhibitor resistance			
Vakulenko *et al.* 1998	1	Amp, Clavulanate [d]	Directed evolution; error-prone PCR ($MgCl_2$- and $MnCl_2$-based)
Vakulenko and Golemi 2002	1	Caz, Clavulanate [d]	Directed evolution; error-prone PCR ($MgCl_2$- and $MnCl_2$-based) followed by DNA-shuffling
Retention and regain of resistance			
Bershtein *et al.* 2008	2	Amp	Directed evolution; error-prone PCR (wobble base analogues)
Structural perturbation			
Osuna *et al.* 2002	2	Amp	Directed evolution; DNA-shuffling
Hecky and Müller 2005	1	Amp	Directed evolution; DNA-shuffling followed by error-prone PCR ($MgCl_2$- and $MnCl_2$-based)

[a] Number of independent alleles present in the respective study (see text).
[b] Abbreviations used; Amx, amoxicillin; Atm, aztreonam; Caz, ceftazidime; Ctx, cefotaxime; Cxm, cefuroxime; Fep, cefepime.
[c] In addition to resistance to these drugs, mutated alleles were submitted to selection for retention of resistance to ampicillin, pipericillin and cephalotin.
[d] ß-lactam inhibitor.

substitutions restoring this resistance. Structural perturbation followed by directed evolution is an efficient way of analyzing thermostability and finding thermo-stabilizing mutations (HECKY and MULLER 2005). Since protein evolution is considered to be prone to activity-stability tradeoffs, stabilizing mutations are thought to play an important role by compensating the destabilizing effects of primary mutations that increase activity but decrease stability (DePristo *et al.* 2005). In TEM-1, such interactions have been demonstrated for mutations that increase cefotaxime resistance (WANG *et al.* 2002). Therefore, despite the fact that the structural perturbation studies have used TEM-variants with a rather aberrant amino acid sequence, we consider it informative to include these studies in the review presented here, as they posses the potential to localize a class of mutations that clearly plays an important role in the evolution of TEM-1 mediated antibiotic resistance.

Most of the studies analyzed here use truly independent libraries (directed evolution) or selection lines (experimental evolution), and therefore substitutions found in different libraries or lines can be considered independent. However, in some directed evolution studies a single pool of mutagenized amplicons was screened by plating transformants on selective agar immediately after transformation, thus preventing competition between clones. As a consequence, the isolated clones, though differing in the exact substitutions they carry, are not independent products of evolution, since a substitution present in two clones can originate from one and the same mutational event. Therefore, substitutions were scored only once, no matter how many different clones were found to contain a particular substitution in such studies (VAKULENKO *et al.* 1998; ZACCOLO and GHERARDI 1999; OSUNA *et al.* 2002; HECKY and MULLER 2005; BERSHTEIN *et al.* 2008). Finally, whenever genotypes were characterized after each cycle of evolution and selection in an experiment, substitutions absent in the final genotype but present in a genotype characterized after one of the preceding cycles were included in the dataset.

Clinical and laboratory selection-spectra

The majority of clinical TEM-isolates that have been phenotypically characterized are so-called ‘extended spectrum ß-lactamases’, meaning that they have evolved the ability to break down more modern ß-lactam antibiotics or ß-lactam - inhibitor combinations, to which TEM-1 is highly sensitive. The exact selection regimes that have preceded the appearance of these alleles are almost always unknown. The antibiotics used in the 17 experimental studies

reviewed here include several modern cephalosporins, a monobactam and two ß-lactam - inhibitor combinations (ampicillin - clavulanate and ceftazidime - clavulanate). An overview of the antibiotics and inhibitors used in the laboratory studies is given in Table 3.2.

Table 3.2: Antibiotics used in laboratory studies and their classification.

ß-lactam family	ß-lactam sub-family	ß-lactams used in laboratory studies
Penicillins	Narrow spectrum	-
	Narrow spectrum penicillinase-resistant	-
	Narrow spectrum ß-lactamase-resistant	-
	Moderate spectrum	Ampicillin , amoxicillin
	Broad spectrum	-
	Extended spectrum	Pipericillin
Cephalosporins	First generation	Cephalotin
	Second generation	Cefuroxime
	Third generation	Cefotaxime, ceftazidime
	Fourth generation	Cefepime
Carbapenems	-	-
Monobactams	-	Aztreonam
ß-lactamase inhibitors	-	Clavulanate

3.2 Results and discussion

The primary objective of this study is to analyze to what extent laboratory evolution can identify amino acid substitutions found in clinical isolates of TEM-alleles that are known to contribute to an extended resistance phenotype. Three factors hamper the straightforward comparison of amino acid substitutions found in nature and experiments.

First, the functional significance of alternative amino acid substitutions at the same position is unclear. Whenever substitutions at a certain amino acid residue occur both in nature and in experiments but the exact substitutions differ between the two, it is important to know whether or not this implies a functional difference. Laboratory evolution shows that various alternative substitutions at the same position can be selected when a single selective agent is used, indicating that alternative substitutions can have qualitatively similar phenotypic effects (BARLOW and HALL 2003; FUJII *et al.* 2004). On the other hand, a detailed phenotypic analysis of all possible substitutions at two amino acid residues in TEM-1 shows that alternative substitutions at one and the same locus can have quantitatively different effects on the degree of resistance (VAKULENKO *et al.* 1995; VAKULENKO *et al.* 1999).

However, since we have found no evidence for alternative substitutions at the same amino acid residue that have a completely different effect on resistance phenotype, we here assume that such alternative substitutions are functionally equivalent. The primary focus will therefore be on amino acid residues, while the specific substitutions at these residues will be addressed secondarily.

Second, it is not clear whether all substitutions that can be found in the dataset of clinical isolates have arisen by natural selection. The majority of randomly introduced substitutions in the TEM-1 background is thought to be functionally neutral (PALZKILL and BOTSTEIN 1992; HUANG *et al.* 1996) and such neutral substitutions could occasionally end up in clinical isolates by genetic hitchhiking (BAQUERO *et al.* 1998). Despite the fact that most of the clinical TEM-isolates have been phenotypically characterized, the contribution of individual amino acid substitutions to the resistance phenotype is often unknown and as a consequence it can be unclear whether substitutions have truly arisen by selection.

Third, methodological errors have led to the erroneous report of novel substitutions in the past and it can not be excluded that some unnoticed errors persist in the dataset of clinical isolates.

An overview of the number of times substitutions have been found at different positions within the TEM-1 amino acid sequence in the datasets of clinical and laboratory isolates is shown in Figure 3.1. In total, 73 different substitutions were recorded at 50 different amino acid residues in clinical isolates while 245 different substitutions were recorded at 135 amino acid residues in laboratory isolates (see supplementary Table 3.1).

It is clear from Figure 3.1 that for a substantial fraction of the amino acid residues at which substitutions have been recorded in clinical isolates only one or two substitutions have been found, rendering these substitutions the suspicion of being adaptively neutral. The analysis of the 50 amino acid residues harbouring substitutions in clinical isolates is therefore split up in two main parts. We first analyze the group of 18 residues at which substitutions with a demonstrated effect on resistance phenotype have been found, i.e. increased catalytic efficiency values or MICs compared to TEM-1 (the location of these residues in the three-dimensional structure of TEM-1 is shown in Figure 3.2a). Based on the type of effect on resistance phenotype, we have split this group in three subgroups of residues harbouring (i) substitutions with an effect on ß-lactam resistance, (ii) substitutions with an effect on

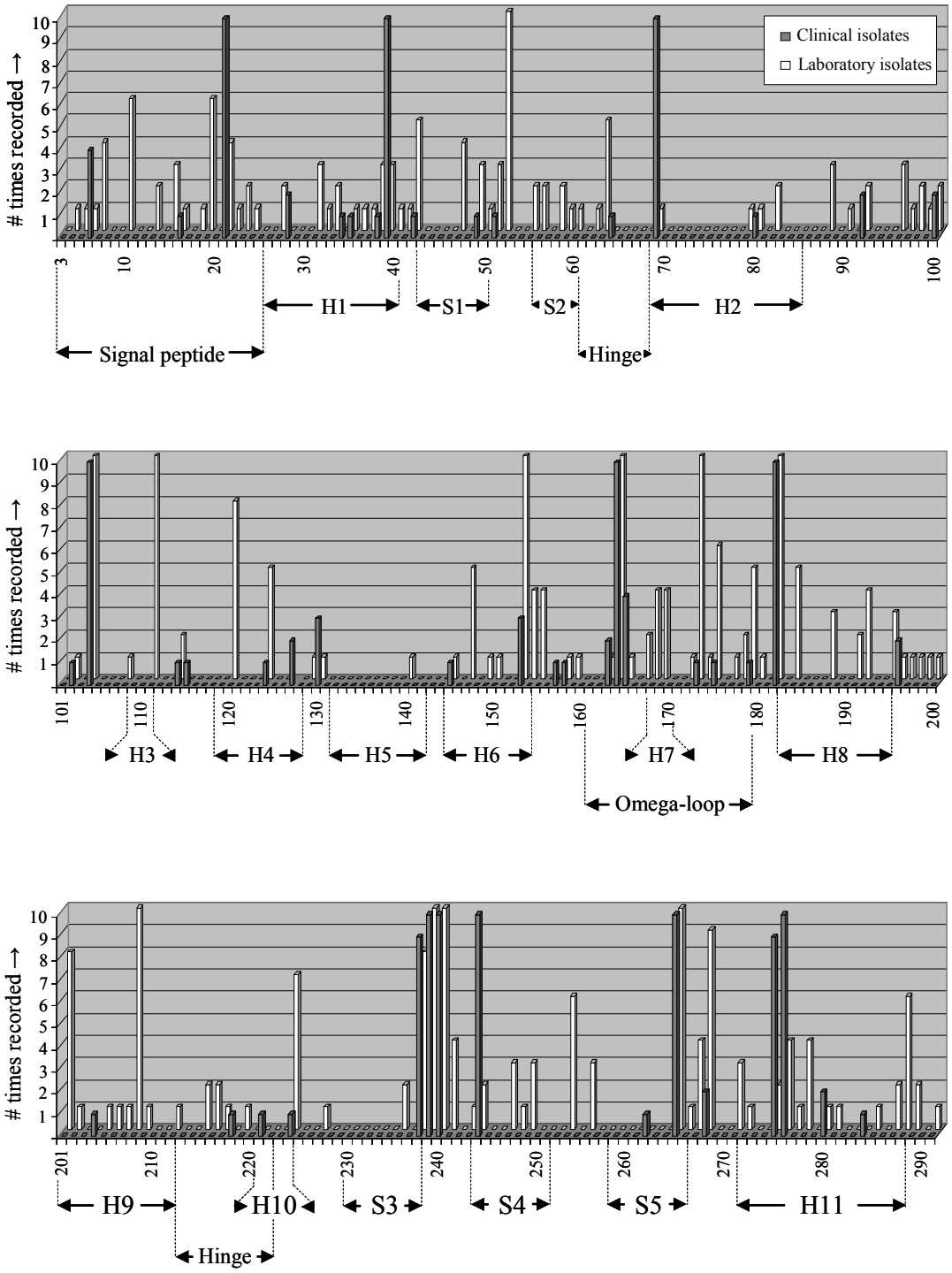


Figure 3.1: Occurrence of substitutions in clinical and laboratory isolates of TEM ß-lactamase. The amino acid numbering according to Ambler (1991) is on the x-axis (note that residues 1, 2, 239 and 253 do not exist in this numbering-scheme). Structural features of the enzyme according to Jelsch *et al.* (1993) are shown below these numbers (H = α-helix, S = ß-sheet). Substitutions occurring more than ten times in either of the two datasets are shown as occurring ten times here (see supplementary table 3.1 for exact frequencies).

resistance to ß-lactamase inhibitors and (iii) substitutions with an effect on both. Second, we analyze the group of 32 residues at which substitutions with unknown effect on resistance phenotype have been recorded in clinical isolates. Since several of these residues have been substituted in a single clinical isolate only, indicating that the respective substitutions might be functionally neutral (or erroneously reported), we have split this group up in two subgroups of residues substituted in (i) multiple clinical isolates and (ii) a single clinical isolate.

Substitutions with demonstrated effect on resistance phenotype

Most of the amino acid residues belonging to this group harbour substitutions that are found at high frequency in the dataset of clinical isolates and whose effect on resistance phenotype has been extensively studied (see Table 3.3 for an overview). While most of these substitutions play a role in extension of the resistance spectrum to more modern ß-lactam antibiotics, some have been shown to play a specific role in resistance to ß-lactamase inhibitors (drugs that bind to but do not hydrolyze ß-lactamases and that are often administered in combination with ß-lactam antibiotics). A third group of substitutions confers a more general effect and can increase resistance to both ß-lactams and inhibitors. We analyze to what extent these three types of substitutions were identified in laboratory isolates below.

Substitutions involved in extension of the ß-lactam resistance spectrum

Exactly which amino acid substitutions are involved in extension of the resistance spectrum of TEM-1 and how such substitutions influence the structure, function and stability of the enzyme has been intensively studied (KNOX 1995). Almost all substitutions that have been found with high frequency in clinical isolates and that have been shown to be involved in increasing resistance to more modern ß-lactam antibiotics (E104K, R164C, R164H, R164S, A237T, G238S and E240K) were identified in multiple experimental studies at high frequency (see Table 3.3). The only exception is substitution Q39K. This substitution has long been thought to be an example of neutral variation present in the TEM-1 background, though it was later shown to slightly but consistently increase the minimal inhibitory concentration (MIC) of ceftazidime and aztreonam (BLAZQUEZ *et al.* 1995), while it enlarges the catalytic efficiency values of a whole series of related antibiotics (CHAIBI *et al.* 1996). Q39K is present in 32 clinical isolates and in a single experimental isolate that was identified

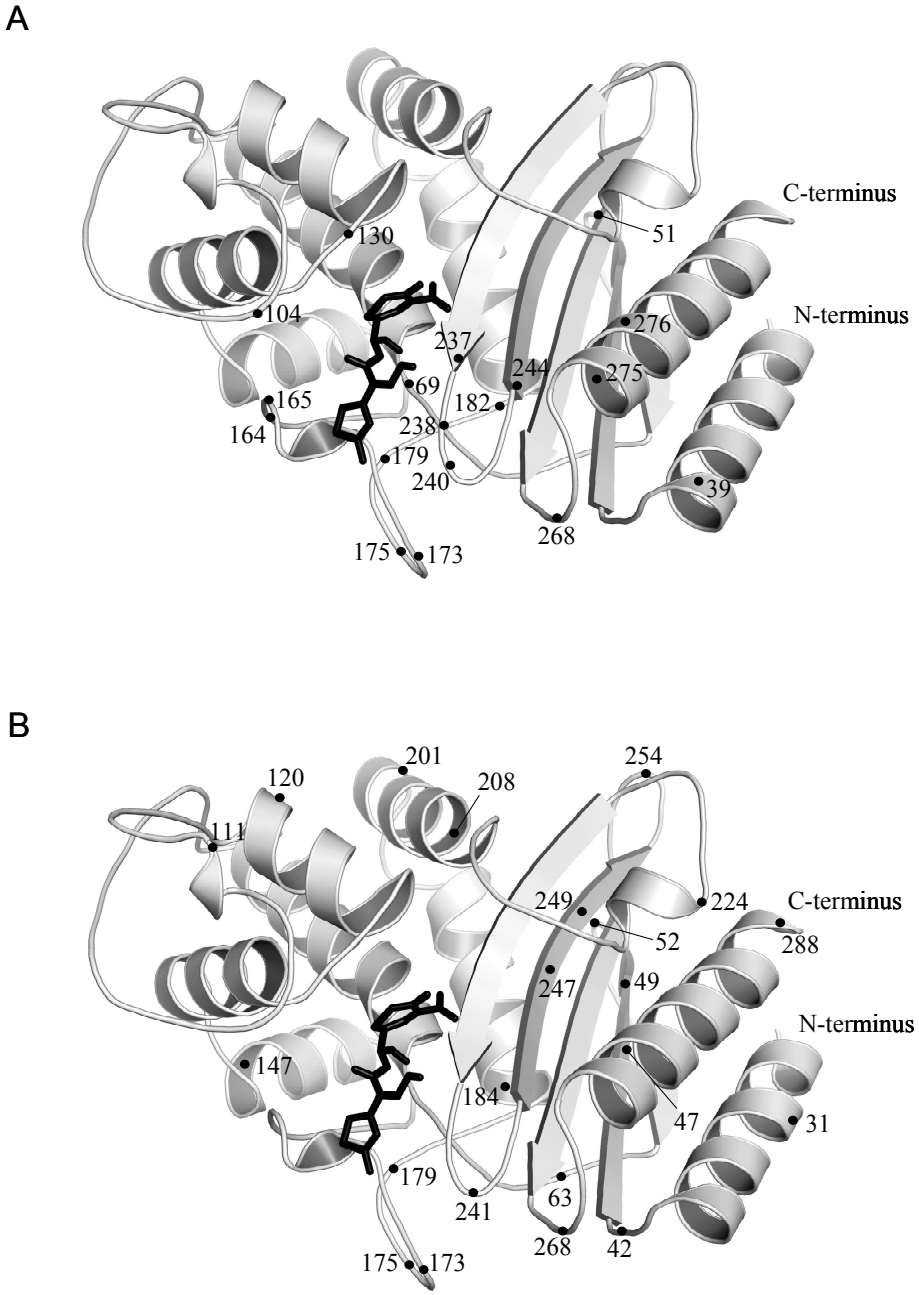


Figure 3.2: Three dimensional structures of TEM-1 ß-lactamase (PDB-code: 1BTL). (A) Amino acid residues at which substitutions with a demonstrated effect on resistance phenotype have been found (listed in Table 3.3). (B) Amino acid residues at which substitutions are found with high frequency in experimental isolates but not in clinical isolates (listed in Table 3.6).

Table 3.3: The 18 amino acid residues at which substitutions with a demonstrated effect on resistance phenotype have been recorded in clinical isolates. Substitutions found at these residues in clinical and laboratory isolates are also shown.

Amino acid residue	*Substitution*	*# clinical isolates*	*# laboratory isolates*
39	Q39K	32	1
	Q39R	0	2
51	L51P	1	0
	L51I	0	1
	L51F	0	2
69	M69I	4	0
	M69L	12	1
	M69V	19	0
104	E104K	42	36
	E104A	0	2
	E104G	0	2
	E104V	0	1
130	S130G	3	1
164	R164C	4	2
	R164H	18	15
	R164S	26	21
	R164G	0	2
	R164N	0	1
	R164Y	0	1
165	W165C	1	0
	W165R	3	0
	W165G	0	1
173	I173V	1	12
	I173L	0	1
	I173R	0	1
	I173T	0	2
175	N175I	1	0
	N175D	0	5
	N175Y	0	1
179	D179E	1	0
	D179G	0	4
	D179Y	0	1
182	M182K	0	1
	M182L	0	1
	M182R	0	1
	M182T	21	26
	M182V	0	1
237	A237G	1	0
	A237T	8	6
	A237S	0	2
238	G238D	1	0
	G238N	1	0
	G238S	32	57
	G238A	0	3
240	E240K	25	13
	E240R	1	0
	E240V	1	1
	E240G	0	4
244	R244G	1	0
	R244H	3	0
	R244L	1	0
	R244C	4	1
	R244S	6	1
268	S268G	2	7
	S268N	0	1
	S268T	0	1
275	R275A	2	0
	R275L	3	1
	R275Q	4	1
276	N276D	12	3
	N276R	1	0
	N276S	1	1

after selection for increased cefotaxime resistance (Salverda and de Visser unpublished), while an alternative substitution at this position (Q39R) was recorded in the same experiment (Salverda and de Visser unpublished). The low frequency of TEM-2 (which differs from TEM-1 by the presence Q39K only) in clinical *E. coli* isolates compared to the frequent presence of Q39K in extended-spectrum TEM-alleles has lead to the suggestion that Q39K might provide host strains with better survival at low concentrations of certain cephalosporins or monobactams, thus increasing the possibility of acquisition of new substitutions (BLAZQUEZ *et al.* 1995). This effect might explain the low frequency of occurrence of this substitution in experimental studies where, due to more stringent selective conditions and

higher mutation rates (when directed evolution is involved), substitutions with large effects tend to dominate.

What remains to be discussed are substitutions at four amino acid residues that have been shown to be involved in extending the resistance spectrum of TEM-1, but that nevertheless occur at low frequency in clinical isolates. These substitutions and substitutions found at these residues in laboratory isolates are discussed below.

L51P

Reversion of substitution L51P in the genetic background of the only clinical isolate it occurs in (TEM-60) was found to decrease both activity towards certain antibiotics and enzyme-stability, which lead to the conclusion that the substitution plays a role in both activity and stability of the enzyme (CAPORALE *et al.* 2004). Alternative substitutions L51F and L51I have been found in experiments twice (Osuna *et al.* 2002; Salverda and de Visser unpublished) and once (OSUNA *et al.* 2002), respectively. It is interesting that two of the three alternative substitutions found at this position in experiments were identified in one of the structural perturbation studies, where they persisted in the clones characterized after multiple rounds of mutagenesis and selection (OSUNA *et al.* 2002). This suggests that L51F and L51I affect enzyme stability and are thus functionally synonymous to L51P. Interestingly, a wide variety of substitutions has been found at the adjacent residue 52 as well and it seems that these substitutions are also involved in increasing enzyme stability (HECKY and MULLER 2005; BERSHTEIN *et al.* 2008).

I173V

This substitution was found in two *in vitro* evolution studies (BARLOW and HALL 2002; BARLOW and HALL 2003) prior to the report of the first and only clinical isolate carrying this substitution (ZARNAYOVA *et al.* 2005). Interestingly, the substitution was later identified in two more experimental studies (Fujii *et al.* 2004; Salverda and de Visser unpublished), resulting in a total of 11 independent observations in laboratory studies. Alternative substitutions I173R (OSUNA *et al.* 2002) and I173T (Barlow and Hall 2002; Salverda and de Visser unpublished) were identified once and twice in experiments, respectively.

The effect of I173V on resistance phenotype was already extensively tested after its identification in laboratory experiments, and although a number of ß-lactam antibiotics were

screened, the substitution was found to only increase resistance to the fourth generation cephalosporin cefepime (BARLOW and HALL 2003). The substitution was later identified in experimental studies using ceftazidime (FUJII *et al.* 2004) and cefotaxime (Salverda and de Visser unpublished) as sole selective agents, indicating a broader phenotypic effect. Therefore, this substitution may be expected to appear in future clinical isolates.

N175I

Despite the presence of N175I in a single clinical isolate only (TEM-138), the substitution is thought to be involved in increased ceftazidime resistance (CHOUCHANI *et al.* 2006). In experimental studies, alternative substitutions N175Y and N175D were found once and five times, respectively, the latter in three different studies (Barlow and Hall 2003; Bershtein *et al.* 2008; Salverda and de Visser unpublished). Although the exact mechanism remains obscure, the high frequency at which these substitutions occur strongly suggests that substitutions at position 175 are involved in extending the resistance phenotype. The Ω-loop on which N175 is present is involved in the formation of the active site pocket (JELSCH *et al.* 1993) and numerous substitutions within this loop are known to alter resistance phenotype. Substitutions at this residue could therefore appear in future clinical isolates.

D179E

Substitution D179E, reported in a single clinical isolate (TEM-126), was found to improve the catalytic efficiency against ceftazidime (DELMAS *et al.* 2005). Although D179E was not found in experiments, alternative substitutions D179G (Blazquez *et al.* 2000; Fujii *et al.* 2004; Salverda and de Visser unpublished) and D179Y (BLAZQUEZ *et al.* 2000) were found at this position. Because of its anchoring function in the Ω-loop, substitutions of residue 179 have been extensively studied. It was found that both D179G and D179Y increase the MIC for ceftazidime 16-fold (VAKULENKO *et al.* 1995). In agreement with this, D179G and D179Y were recorded in laboratory studies that use ceftazidime as selective agent (BLAZQUEZ *et al.* 2000; FUJII *et al.* 2004). Interestingly, although no effect on cefotaxime resistance could be measured for a mutant created *in vitro* (VAKULENKO *et al.* 1999), substitution D179G was isolated twice in an experiment that used cefotaxime as single selective agent (SALVERDA and DE VISSER unpublished), indicating that the effect on resistance phenotype might be broader than previously assumed.

A detailed study of the effect of all possible amino acid substitutions at this residue on resistance phenotype shows that although nearly all substitutions increase ceftazidime resistance, at the same time they almost all dramatically worsen resistance to ampicillin (VAKULENKO *et al.* 1995). Indeed, in a laboratory evolution experiment focussing on fluctuating ß-lactam pressure, substitutions D179G and D179Y were only found in the absence of fluctuating selection-pressure (BLAZQUEZ *et al.* 2000). Pleiotropic effects of substitutions at this residue might thus have prevented their appearance in clinical isolates so far.

Substitutions involved in inhibitor resistance

Substitutions at residues 69, 130, 165, 244, 275 and 276 have been shown to play an exclusive role in inhibitor resistance (CHAIBI *et al.* 1999) and are found at high frequency in the dataset of clinical isolates. Substitutions at residues 69, 130, 244 and 276 were also identified in an experimental study selecting for increased resistance to the ß-lactam - inhibitor combination ampicillin - clavulanate (VAKULENKO *et al.* 1998). The only other experimental study of inhibitor resistance used a combination of ceftazidime - clavulanate and also identified the most frequently found substitution at residue 276 (VAKULENKO and GOLEMI 2002). There is evidence for the involvement of substitutions at the remaining two residues (165 and 275) in increasing the resistance to inactivation by clavulanate, although the effects are small (LENFANT *et al.* 1993; CHAIBI *et al.* 1999; FIETT *et al.* 2000). This might explain why substitutions at these residues have not been found in the two experimental studies that have used clavulanate as a selective agent. Substitutions at residues 165 and 275 have been found in other experiments, but under selection regimes without inhibitors; W165G was found in a study selecting for increased cefotaxime resistance (Salverda and de Visser unpublished) while R275L and R275Q were both found in one of the structural perturbation studies (OSUNA *et al.* 2002), suggesting that substitutions at these residues are able to increase resistance to novel antibiotics and inhibitors at the same time.

Amino acid residues harbouring substitutions involved in both ß-lactam and inhibitor resistance

Substitutions M182T and S268G have been shown to increase resistance to both ß-lactams and ß-lactam – inhibitor combinations (HUANG and PALZKILL 1997; VAKULENKO and GOLEMI

2002). Typical for these modulating substitutions is that their effect is only apparent in the background of other substitutions that affect resistance phenotype. Substitution M182T, present in 21 clinical isolates, was found in the majority of the experimental studies. Substitution S268G is far less common among clinical isolates (present only in TEM-49 and TEM-136). This substitution was found in four different experimental studies, three using cefotaxime as selective agent (Holloway *et al.* 2007; Kopsidas *et al.* 2007; Salverda and de Visser unpublished) and one using a combination of ceftazidime and clavulanate (VAKULENKO and GOLEMI 2002). In agreement with evidence from mutant TEM-variants with and without this substitution, this indicates that the substitution plays a role in increased resistance to both ß-lactams and inhibitors and may very well have a similar mode of action as that described for M182T (HUANG and PALZKILL 1997; SIDERAKI *et al.* 2001; WANG *et al.* 2002).

Substitutions with unknown or no effect on resistance phenotype

There are 32 residues in clinical isolates at which substitutions have been identified that either have not been tested for their effect on resistance phenotype or that failed to give any measurable results when tested. For 11 of these residues, substitutions were identified in multiple clinical isolates (Table 3.4), while for the remaining 21 residues a substitution was found in a single clinical isolate only (Table 3.5). We analyze to what extent clinical and laboratory data overlap for both groups in order to identify to what extent laboratory data support adaptive effects of substitutions found at these residues.

Substitutions found in multiple clinical isolates

At eleven amino acid residues, multiple substitutions without known effect on resistance phenotype have been identified in clinical isolates. They are discussed below one by one.

Q6K

This substitution, located in the signal sequence (residue 3-25), was found in four different clinical isolates (TEM-87, 92, 123 and 124). In experiments an alternative substitution (Q6R) has been identified in a single sequenced clone (BARLOW and HALL 2002).

Although it has been extensively tested, no phenotypic effect of Q6K has been described (PERILLI *et al.* 2002). It has been argued that since the positively charged amino-terminal

region of the signal peptide plays an important role in efficient protein secretion across the cytoplasmic membrane (INOUYE *et al.* 1982), substitution of the polar glutamine residue at position 6 for a positively charged lysine residue could result in an effect on resistance phenotype (PERILLI *et al.* 2002). Interestingly, the alternative substitution Q6R that was found in a single experiment (BARLOW and HALL 2002) would result in a similar effect.

L21F

Substitution L21F, again present in the signal sequence, was found in as many as 21 clinical TEM-alleles. An alternative substitution, L21I, has been reported in a single allele (TEM-67) as well as in a single experimental study (Salverda and de Visser unpublished). Substitution L21P was identified three times in two other experimental studies (ZACCOLO and GHERARDI 1999; BERSHTEIN *et al.* 2008).

Because of its location in the signal sequence, substitution L21F has been speculated to be functionally neutral (SOUGAKOFF *et al.* 1989). Nevertheless, the frequent occurrence of L21F in clinical isolates as well as the presence of alternative substitutions at the same position in a clinical isolate and in four experimentally isolated alleles strongly suggests a phenotypic effect. This is supported by the observation that the substitution occurs in an otherwise unaltered TEM-1 background in TEM-117, excluding genetic hitchhiking as a factor that may have contributed to the frequent occurrence of the substitution. Furthermore, L21F occurs both in the background of substitutions known to be involved in extension of the resistance spectrum and in the background of substitutions involved in inhibitor resistance, indicating a general effect of the substitution (e.g. at the level of protein expression), which would be expected for substitutions with functional effect in the signal peptide. Although the exact effect remains to be identified, it seems that substitutions at this residue can alter resistance phenotype.

E28D and E28K

Both substitutions have been found in a single clinical isolate (TEM-150 and TEM-162 respectively). No substitutions were recorded at this residue in experiments and it therefore remains unclear whether the substitutions found in the two clinical isolates confer an adaptive advantage.

Table 3.4: The 11 amino acid residues at which substitutions with unknown effect on resistance phenotype have been found in multiple clinical isolates. Substitutions found at these residues in clinical and laboratory isolates are also shown.

Amino acid residue	*Substitution*	*# clinical isolates*	*# laboratory isolates*
6	Q6K	4	0
	Q6R	0	1
21	L21F	21	0
	L21I	1	1
	L21P	0	3
28	E28D	1	0
	E28K	1	0
92	G92D	2	1
	G92S	0	1
100	N100S	2	0
	N100G	0	1
	N100K	0	1
127	I127V	2	0
153	H153R	3	6
	H153D	0	1
	H153L	0	1
	H153Q	0	1
	H153Y	0	1
163	D163G	1	0
	D163H	1	0
	D163V	0	1
196	G196D	1	0
	G196S	1	1
265	T265M	20	12
	T265A	0	2
280	A280V	2	0
	A280T	0	1

Table 3.5: The 21 amino acid residues at which a substitution has been found in a single clinical isolate only. Substitutions found at these residues in laboratory isolates are also shown.

Amino acid residue	*Substitution*	*# clinical isolates*	*# laboratory isolates*
16	F16L	1	0
	F16C	0	1
34	K34E	1	0
35	D35P	1	0
	D35A	0	1
38	D38N	1	1
	D38G	0	1
	D38V	0	1
	D38Y	0	1
42	A42V	1	0
	A42G	0	6
49	L49M	1	3
	L49V	0	1
64	E64K	1	0
80	V80E	1	0
	V80I	0	1
102	L102V	1	0
	L102M	0	1
114	T114P	1	0
	T114A	0	1
	T114M	0	1
115	D115G	1	0
124	S124N	1	1
	S124G	0	2
	S124R	0	2
145	P145A	1	0
	P145L	0	1
	P145Q	0	1
157	D157E	1	0
158	H158N	1	0
	H158Y	0	1
204	R204Q	1	0
218	G218E	1	0
221	L221M	1	0
224	A224V	1	7
262	V262I	1	0
284	A284G	1	0

G92D

This substitution was found in two clinical isolates (TEM-57 and TEM-66). It has been recorded in experiments once (Salverda and de Visser unpublished), as has the alternative substitution G92S (STEMMER 1994).

Effects of G92D on resistance spectrum or enzyme kinetics have been tested but were not found (BONNET *et al.* 1999). The recovery of two substitutions at residue 92 in two laboratory studies provides some evidence that substitutions at this position can contribute to an altered resistance phenotype.

N100S

This substitution was only found in two very recently identified clinical isolates (TEM-157 and TEM-162). Alternative substitutions N100G and N100K have both been found in experiments once (Bershtein *et al.* 2008; Salverda and de Visser unpublished). It remains unclear whether or not substitutions at this residue confer an adaptive advantage.

I127V

Substitution I127V has been identified in two clinical isolates (TEM-80 and TEM-81). In both cases the substitution is present together with substitutions involved in inhibitor resistance (M69L and N276D in TEM-80 and M69L in TEM-81). No substitutions were recorded at this position in experiments.

The precise effect of I127V is unknown, although it has been found to increase Km values (and therefore decrease affinity) for a number of substrates compared to TEM-1 (ARPIN *et al.* 2002). The absence of substitution I127V in experimental studies may therefore either lie in the neutrality of this substitution or, considering the fact that it was found in inhibitor resistant backgrounds, in the limited scope of experimental studies that focus on inhibitor resistance.

H153R

This substitution was recorded in three clinical isolates (TEM-21, TEM-56 and TEM-112), always in the background of substitutions involved in extended cephalosporin resistance. No phenotypic effect of the substitution has been described. In experiments, this substitution was found in two studies selecting for increased cefotaxime resistance (Kopsidas *et al.* 2007; Salverda and de Visser unpublished), in a study on the retention and regain of ampicillin resistance (BERSHTEIN *et al.* 2008) and in both structural perturbation studies (OSUNA *et al.* 2002; HECKY and MULLER 2005), clearly suggesting that the substitution confers a selectable advantage. The presence of H153R in both structural perturbation studies and the fact that it has only been observed in the background of other substitutions in clinical and laboratory isolates may indicate that the substitution compensates for the pleiotropic effects of other adaptive substitutions. Alternative substitutions H153D, H153L, H153Q and H153Y have been recorded in experiments as well, suggesting that a variety of substitutions at this position can result in selectable effects.

D163G and D163H

These substitutions were found in TEM-96 and TEM-87 respectively. No phenotypic effect of these substitutions has ever been described, although there are indications that D163G might increase resistance to ceftazidime (PALZKILL *et al.* 1994). A single substitution at the same position (D163V) was found in an experiment selecting for increased cefotaxime resistance (Salverda and de Visser unpublished).

Residue 163 is located in the Ω-loop, which is involved in the formation of the active site pocket. The identification of substitutions at this position, both in clinical isolates and in experiments, suggests that perturbation of the Ω-loop at this position might alter the native resistance phenotype of TEM-1. In fact, it has been shown that simultaneous substitution of residue 161 and 163 can result in increased resistance to ceftazidime (PALZKILL *et al.* 1994).

G196D and G196S

Substitutions G196D and G196S were found in TEM-88 and TEM-108 respectively. In experiments, G196S was found once (Salverda and de Visser unpublished). Concerning G196D, no differences in kinetic properties were found when TEM-88 was compared to TEM-52, which lacks G196D but is otherwise identical (PAI *et al.* 2001). In a mutagenesis study, it was shown that a variety of substitutions at residue 196 are tolerated (HUANG *et al.* 1996), leading to the speculation that G196D is functionally silent (PAI *et al.* 2001). No further information is available on substitution G196S, except that it occurs in an unusual allele that also contains substitutions V80E and N276S that are not found in any other TEM-allele described. With only a single substitution recorded at residue 196 in experiments, it remains unclear whether or not substitutions at this position can alter resistance phenotype.

T265M

Although found in 21 different clinical TEM-isolates and extensively tested, no effect on enzyme catalysis or expression was found for this substitution (HUANG *et al.* 1994). However, since the substitution has been identified no less than twelve times in five independent experimental studies, it is clear that it does confer an adaptive advantage. The fact that T265M is frequently encountered in one of the two structural perturbation studies (HECKY and MULLER 2005) suggests that the substitution has an effect on enzyme stability rather than enzyme activity, which is in agreement with its location far away from the active site. This is

also in agreement with the fact that T265M has only been found in the background of other substitutions involved in both ß-lactam or inhibitor resistance. All in all, it seems that substitution T265M is a suppressor of defects introduced by other substitutions, with a mode of action similar to that of M182T (HUANG and PALZKILL 1997; SIDERAKI *et al.* 2001; WANG *et al.* 2002).

A280V

This obscure substitution has been described in two clinical isolates (TEM-101 and TEM-104). No phenotypic effect has been described and although a single alternative substitution (A280T) was found in one of the structural perturbation studies (OSUNA *et al.* 2002), it remains unclear whether or not any of these substitutions confers a selective advantage. Note that in the sequence of both clinical isolates substitution A280V is caused by a double point mutation (the codon changes from 'G**CT'** to 'G**TC'**).

Substitution found in a single clinical isolate

As mentioned before, there are 21 amino acid residues that have been altered in a single clinical isolate only and of which no phenotypic effect is known (see Table 3.5). At nine of the 21 residues, no substitutions were found in experiments and at five other residues a single alternative substitution was found in experiments, which we do not regard as sufficient evidence to support a putative adaptive advantage. At the remaining seven residues, multiple substitutions were found in experiments, suggesting a role of substitutions at these positions in extending resistance phenotype. The substitutions found at these positions in clinical and in experimental isolates are discussed in detail below.

D38N

This substitution was identified in TEM-162 and its effect on resistance phenotype is unknown. The substitution was found in a single experimental study on increased cefotaxime resistance (Salverda and de Visser unpublished) in which two alternative substitutions (D38G and D38V) were identified as well, while D38Y was found in yet another study (BARLOW and HALL 2003). Thus, indications exist that substitutions of this amino acid can confer a selectable advantage.

A42V

No effect on resistance phenotype was found for A42V in TEM-42, the only clinical isolate containing this substitution (MUGNIER *et al.* 1996). An alternative substitution (A42G) was identified in as many as six different experimental studies. The A42G substitution has been particularly well studied as, combined with substitutions E104K, M182T and G238S, it gives rise to the TEM-derivative with the highest cefotaxime resistance described so far (HALL 2002). Interestingly, all of the later studies that have used selection for increased cefotaxime resistance were able to identify the triple mutant E104K/ M182T/ G238S (ZACCOLO and GHERARDI 1999; LONG-MCGIE *et al.* 2000; ORENCIA *et al.* 2001; KOPSIDAS *et al.* 2007; SALVERDA and DE VISSER unpublished), but none of them identified the quadruple mutant containing A42G. Detailed MIC analyses of different combinations of the four substitutions however indicates that addition of A42G to the triple mutant should greatly increase cefotaxime resistance (HALL 2002; WEINREICH *et al.* 2006).

The identification of A42G in studies with various selective conditions (antibiotics, inhibitors and structural perturbation), its location (far away from the active site and binding pocket), and the fact that the substitution has no effect on resistance in a TEM-1 background (HALL 2002) suggest that this substitution has a modulating effect comparable to that of M182T. Whether substitution A42V that is found in a single clinical isolate confers a similar effect remains unclear.

L49M

Substitution L49M was found in TEM-139. The allele was found in isolates originating from several Bulgarian clinics, and since it was phenotypically indistinguishable from TEM-3 (from which it differs by L49M only) it has been suggested to have no clinical relevance (SCHNEIDER *et al.* 2007). L49M was identified three times in a single experimental study, while an alternative substitution (L49V) was found in the same study (Salverda and de Visser unpublished). The identification of L49M in three independent experiments within a single study indicates that the substitution is likely to have an effect on resistance phenotype.

T114P

This substitution was found in TEM-146. Alternative substitutions T114A (BERSHTEIN *et al.* 2008) and T114M (Salverda and de Visser unpublished) have both been found in a single

experimental isolate. It is unclear whether and to what extent substitutions at this position can alter the level or spectrum of resistance.

S124N

TEM-105 is the only clinical TEM-isolate known to carry substitution S124N, with no other substitutions present in its background. No information is available on the phenotypic effect of the substitution, although the fact that it occurs all by itself could imply some effect on extended spectrum activity.

Substitution S124N was found in experiments once, while alternative substitutions S124G and S124R were both found twice (Bershtein *et al.* 2008; Salverda and de Visser unpublished), indicating that substitutions at this residue may have effects on resistance phenotype, possibly by secondary, stabilizing effects.

P145A

The only clinical isolate containing P145A (TEM-95) was found to have a resistance phenotype similar to TEM-1 (BRINAS *et al.* 2002). Alternative substitutions P145L and P145Q were both identified once in an experiment selecting for increased cefotaxime resistance (Salverda and de Visser unpublished), indicating that substitutions at this position may be involved in extension of the resistance phenotype.

A224V

This substitution is present in a single TEM-allele only (TEM-147) and no phenotypic effect of the substitution has been described. In experiments, the substitution was found in studies using cefepime (BARLOW and HALL 2003) and cefotaxime (Salverda and de Visser unpublished) as selective agents. Furthermore, the substitution was recorded in all of the clones that were sequenced after the final round of mutagenesis and selection in one of the structural perturbation studies (HECKY and MULLER 2005), while it was also present in one of the clones analyzed in the other structural perturbation study (OSUNA *et al.* 2002). The presence of A224V in both structural perturbation studies strongly suggests a stabilizing effect. Both the position of residue 224 (on the small helix H10 just after the second hinge-region) and the fact that substitution A224V has never been found by itself in clinical or experimental isolates are in agreement with this conclusion. The substitution may be expected

to appear in future clinical TEM-isolates, due to its possible modulating function (similar to that of M182T).

Table 3.6: Amino acid substitutions found at high frequency in laboratory isolates that have not been found (or that are found at very low frequency) in clinical isolates.

Structure [a]	*AA* [b]	*Substitution*	*Clinical isolates*	*Laboratory isolates*
Signal peptide	10	V10I	0	6
	19	F19C	0	2
		F19I	0	1
		F19L	0	2
		F19V	0	1
H1	31	V31A	0	3
Turn H1/ S1	42	A42G	0	6
		A42V	1	0
S1	47	I47V	0	4
	49	L49M	1	3
		L49V	0	1
Turn S1/ S2	52	N52D	0	3
		N52H	0	1
		N52S	0	3
		N52T	0	2
		N52Y	0	1
Hinge I	63	E63A	0	1
		E63D	0	1
		E63G	0	1
		E63K	0	2
H3/ Loop	111	K111E	0	1
		K111M	0	2
		K111Q	0	2
		K111R	0	3
		K111T	0	2
H4	120	R120E	0	1
		R120G	0	5
		R120S	0	2
H6	147	E147A	0	1
		E147G	0	2
		E147K	0	2
Loop	155	M155I	0	4
Ω-loop	173	I173L	0	1
		I173R	0	1
		I173T	0	2
		I173V	1	12
	175	N175D	0	5
		N175I	1	0
		N175Y	0	1
	179	D179E	1	0
		D179G	0	4
		D179Y	0	1
H8	184	A184V	(3) [c]	4
		A184T	0	1
H9	201	L201I	0	1
		L201P	0	4
		L201Q	0	1
		L201R	0	1
	208	I208K	0	1
		I208L	0	1
		I208M	0	5
		I208T	0	1
		I208V	0	2
H10/ Hinge II	224	A224V	1	7
Turn S3/ S4	241	R241H	0	4
S4	247	I247V	0	3
	249	A249V	0	3
Turn S4/ S5	254	D254G	0	3
		D254H	0	1
		D254N	0	2
Turn S5/ H11	268	S268G	2	7
		S268N	0	1
		S268T	0	1
H11	288	K288E	0	2
		K288Q	0	1
		K288R	0	2
		K288T	0	1

[a] Structure of the part of TEM-1 the respective amino acid residue is localized in according to Jelsch *et al.* (1993).
[b] Amino acid residue number.
[c] Present in three clinical isolates, but together with V84I and therefore not the result of natural selection (see text).

Prediction of substitutions that may be encountered in future clinical isolates

Amino acid residues at which no substitutions have been found in clinical isolates, but at which multiple substitutions have been recorded in experiments are of special interest, since they may reveal the evolutionary potential of TEM-1 that has not been used by natural selection yet. In order to make a conservative (albeit somewhat arbitrary) estimate of this group of substitutions, we here consider a particular substitution to have arisen by selection

whenever it was found more than twice in laboratory isolates and we consider an amino acid residue to be under selection whenever more than four substitutions have been found at a single residue in laboratory isolates. The results of this analysis are shown in Table 3.6. Data on amino acid residues at which less than three substitutions were found in clinical isolates but at which substitutions were recorded with high frequency in experiments have been added to this table for completeness. An overview of the location of all amino acid residues in the three-dimensional structure of TEM-1 is shown in Figure 3.2b.

3.3 Conclusions

Summarizing, we have made an inventory of all amino acid substitutions present in the ~150 unique clinical TEM-isolates that have so far been described and in a collection of 17 laboratory studies on the experimental and directed evolution of TEM-mediated antibiotic resistance. Since we have reasons to believe that alternative substitutions at one and the same amino acid residue are, at least qualitatively, equivalent, we have focussed primarily on the 50 amino acid *residues* at which substitutions have been found in clinical isolates. Within this group, we have first analyzed the 18 residues at which substitutions have been found that have been shown to contribute to an extension of the resistance phenotype. Substitutions at residues 104, 164, 182, 237, 238 and 240, all frequently encountered in clinical isolates, are typically associated with resistance to extended-spectrum ß-lactams (Knox 1995). Laboratory evolution studies identified substitutions at all these residues, and, moreover, they correctly identified the substitutions at these positions that were found in multiple clinical isolates (E104K, R164C, R164H, R164S, M182T, A237T, G238S and E240K). In addition, several substitutions have been described that occur in a single clinical isolate only, but that nevertheless do confer an increased resistance to one or more ß-lactams (L51P, I173V, N175I, D179E). Again, laboratory evolution identified multiple substitutions at all of these residues, while in the case of I173V it also identified the correct substitution.

Substitutions at residues 69, 130, 165, 244, 275 and 276 are associated with inhibitor resistance (YANG *et al.* 1999). Although laboratory evolution again identified substitutions at all these residues, in this case the identification of specific amino acid substitutions was not perfect. Amino acid substitutions M69I, M69L, M69V, S130G, W165R, R244C, R244H, R244S, R275A, R275L, R275Q and N276D have all been found in multiple clinical isolates.

Laboratory evolution failed to find M69I, M69V, W165R, R244H and R275A, while R275L and R275Q were only found in one of the structural perturbation studies (OSUNA *et al.* 2002) where they may serve a different function. The two studies on ß-lactam - inhibitor combinations used the same inhibitor (clavulanate) and both isolated clones after a single round of mutagenesis and selection only. The resulting limited quantitative and qualitative scope of these studies is likely to explain the limited identification of substitutions involved in inhibitor resistance. Yet, even with this very limited scope, the two studies identified substitutions at four of the six residues associated with inhibitor resistance.

Besides the accurate identification of almost all substitutions known to be involved in increased resistance to extended-spectrum ß-lactams and inhibitors, a large number of substitutions present in multiple clinical isolates exists that have no known individual phenotypic effect. Most of these substitutions occur in a limited number of clinical isolates, but two of them occur in at least 20 different isolates (L21F and T265M). Interestingly, L21F has not been identified in laboratory studies, although three alternative substitutions were found at this residue. It seems that the failure to detect an effect of this substitution on resistance phenotype coincides with the failure to detect the same substitution in laboratory evolution experiments.

The case is different for substitution T265M, which has been identified in four different laboratory studies, strongly suggesting that, contrary to the results of extensive testing (HUANG *et al.* 1994), this substitution does alter resistance phenotype. The fact that this substitution is one of the most frequently encountered substitutions in one of the two structural perturbation studies (HECKY and MULLER 2005) indicates that it may have an effect on enzyme stability rather than enzyme activity, which is in agreement with the observation that the substitution is located far away from the active site (HUANG *et al.* 1994). This is also in agreement with the fact that in clinical isolates T265M occurs in the background of substitutions involved in both ß-lactam and inhibitor resistance. All in all, it seems that substitution T265M is a suppressor of defects introduced by other substitutions, with a mode of action similar to that of M182T (HUANG and PALZKILL 1997; SIDERAKI *et al.* 2001; WANG *et al.* 2002).

Next to tracing substitutions involved in altering the resistance phenotype of TEM-1, independent laboratory studies have identified a number of substitutions that have not been recorded in the clinical isolates so far described. The predictions that can be inferred from

laboratory studies concern substitutions in the signal peptide, in structurally important groups in or near the binding pocket, and away from the binding pocket that probably suppress disadvantageous effects of other substitutions. These substitutions illustrate the remarkable plasticity of TEM-1. In a number of cases, especially where it concerns substitutions in the Ω-loop, pleiotropic effects of substitutions might be responsible; such substitutions tend to surface in laboratory studies under the selective pressure of a single antibiotic, while in a clinical setting where often more antibiotics are used at the same time their selective advantage vanishes. For example, almost all of the 19 possible substitutions at residue 179 increase resistance to ceftazidime compared to TEM-1, but at the same time all these substitutions result in a dramatic decrease of resistance to ampicillin (VAKULENKO *et al.* 1995). Similarly, it has been found that exposing bacteria carrying a plasmidic TEM-1 to the alternating pressure of two ß-lactams eventually yields substitutions that are frequently found in clinical isolates, while continued exposure to only one of these antibiotics resulted in the identification of substitutions not previously found in clinical isolates (BLAZQUEZ *et al.* 2000).

The selective pressures applied in the different laboratory studies can be distinguished in conditions scanning for a better fit (increased antibiotic resistance), scanning for a worse fit (increased antibiotic - inhibitor resistance) and scanning for a regained fit (structural perturbation). Combining these three experimental approaches seems especially promising, as it provides a thorough scan of the evolutionary potential as well as robustness of an antibiotic resistance gene while at the same time it characterizes the substitutions involved.

References

AMBLER, R. P., A. F. W. COULSON, J. M. FRERE, J. M. GHUYSEN, B. JORIS *et al.* 1991. A Standard Numbering Scheme for the Class-a Beta-Lactamases. Biochem. J. **276:** 269-270.

ARPIN, C., R. LABIA, W. DUBOIS, P. NOURY, M. SOUQUET *et al.* 2002. TEM-80, a novel inhibitor-resistant beta-lactamase in a clinical isolate of Enterobacter cloacae. Antimicrob. Agents Chemother. **46:** 1183-1189.

BAQUERO, F., M. C. NEGRI, M. I. MOROSINI and J. BLAZQUEZ. 1998. Antibiotic-selective environments. Clin. Infect. Dis. **27:** S5-S11.

BARLOW, M., and B. G. HALL. 2002. Predicting evolutionary potential: in vitro evolution accurately reproduces natural evolution of the tem beta-lactamase. Genetics **160:** 823-832.

BARLOW, M., and B. G. HALL. 2003. Experimental prediction of the natural evolution of antibiotic resistance. Genetics **163:** 1237-1241.

BERSHTEIN, S., K. GOLDIN and D. S. TAWFIK. 2008. Intense neutral drifts yield robust and evolvable consensus proteins. J. Mol. Biol. **379:** 1029-1044.

BLAZQUEZ, J., M. I. MOROSINI, M. C. NEGRI and F. BAQUERO. 2000. Selection of naturally occurring extended-spectrum TEM beta-lactamase variants by fluctuating beta-lactam pressure. Antimicrob. Agents Chemother. **44:** 2182-2184.

BLAZQUEZ, J., M. I. MOROSINI, M. C. NEGRI, M. GONZALEZLEIZA and F. BAQUERO. 1995. Single Amino-Acid Replacements at Positions Altered in Naturally-Occurring Extended-Spectrum Tem Beta-Lactamases. Antimicrob. Agents Chemother. **39:** 145-149.

BONNET, R., C. DE CHAMPS, D. SIROT, C. CHANAL, R. LABIA *et al.* 1999. Diversity of TEM mutants in Proteus mirabilis. Antimicrob. Agents Chemother. **43:** 2671-2677.

BRINAS, L., M. ZARAZAGA, Y. SAENZ, F. RUIZ-LARREA and C. TORRES. 2002. beta-lactamases in ampicillin-resistant Escherichia coli isolates from foods, humans, and healthy animals. Antimicrob. Agents Chemother. **46:** 3156-3163.

CAMPS, M., J. NAUKKARINEN, B. P. JOHNSON and L. A. LOEB. 2003. Targeted gene evolution in Escherichia coli using a highly error-prone DNA polymerase I. Proc. Natl. Acad. Sci. U. S. A. **100:** 9727-9732.

CAPORALE, B., N. FRANCESCHINI, M. PERILLI, B. SEGATORE, G. M. ROSSOLINI *et al.* 2004. Biochemical characterization of laboratory mutants of extended-spectrum beta-lactamase TEM-60. Antimicrob. Agents Chemother. **48:** 3579-3582.

CHAIBI, E. B., S. FARZANEH, J. PEDUZZI, M. BARTHELEMY and R. LABIA. 1996. An additional ionic bond suggested by molecular modelling of TEM-2 might induce a slight discrepancy between catalytic properties of TEM-1 and TEM-2 beta-lactamases. FEMS Microbiol. Lett. **143:** 121-125.

CHAIBI, E. B., J. PEDUZZI, M. BARTHELEMY and R. LABIA. 1997. Are TEM beta-lactamases encoded by pBR322 and Bluescript plasmids enzymatically indistinguishable? J. Antimicrob. Chemother. **39:** 668-669.

CHAIBI, E. B., D. SIROT, G. PAUL and R. LABIA. 1999. Inhibitor-resistant TEM beta-lactamases: phenotypic, genetic and biochemical characteristics. J. Antimicrob. Chemother. **43:** 447-458.

CHOUCHANI, C., R. BERLEMONT, A. MASMOUDI, M. GALLENI, J. M. FRERE *et al.* 2006. A novel extended-spectrum TEM-type beta-lactamase, TEM-138, from Salmonella enterica serovar infantis. Antimicrob. Agents Chemother. **50:** 3183-3185.

DATTA, N., and P. KONTOMICHALOU. 1965. Penicillinase Synthesis Controlled by Infectious R Factors in Enterobacteriaceae. Nature **208:** 239-241.

DELMAS, J., F. ROBIN, F. BITTAR, C. CHANAL and R. BONNET. 2005. Unexpected enzyme TEM-126: Role of mutation Asp179Glu. Antimicrob. Agents Chemother. **49:** 4280-4287.

DEPRISTO, M. A., D. M. WEINREICH and D. L. HARTL. 2005. Missense meanderings in sequence space: A biophysical view of protein evolution. Nat. Rev. Genet. **6:** 678-687.

FIETT, J., A. PALUCHA, B. MIACZYNSKA, M. STANKIEWICZ, H. PRZONDO-MORDARSKA *et al.* 2000. A novel complex mutant beta-lactamase, TEM-68, identified in a Klebsiella pneumoniae isolate from an outbreak of extended-spectrum beta-lactamase-producing klebsiellae. Antimicrob. Agents Chemother. **44:** 1499-1505.

FUJII, R., M. KITAOKA and K. HAYASHI. 2004. One-step random mutagenesis by error-prone rolling circle amplification. Nucleic Acids Res. **32:** -.

HALL, B. G. 2002. Predicting evolution by in vitro evolution requires determining evolutionary pathways. Antimicrob. Agents Chemother. **46:** 3035-3038.

HECKY, J., and K. M. MULLER. 2005. Structural perturbation and compensation by directed evolution at physiological temperature leads to thermostabilization of beta-lactamase. Biochemistry **44:** 12640-12654.

HOLLOWAY, A. K., T. PALZKILL and J. J. BULL. 2007. Experimental evolution of gene duplicates in a bacterial plasmid model. J. Mol. Evol. **64:** 215-222.

HUANG, W. Z., Q. Q. LE, M. LAROCCO and T. PALZKILL. 1994. Effect of Threonine-to-Methionine Substitution at Position-265 on Structure and Function of Tem-1 Beta-Lactamase. Antimicrob. Agents Chemother. **38:** 2266-2269.

HUANG, W. Z., and T. PALZKILL. 1997. A natural polymorphism in beta-lactamase is a global suppressor. Proc. Natl. Acad. Sci. U. S. A. **94:** 8801-8806.

HUANG, W. Z., J. PETROSINO, M. HIRSCH, P. S. SHENKIN and T. PALZKILL. 1996. Amino acid sequence determinants of beta-lactamase structure and activity. J. Mol. Biol. **258:** 688-703.

INOUYE, S., X. SOBERON, T. FRANCESCHINI, K. NAKAMURA, K. ITAKURA *et al.* 1982. Role of Positive Charge on the Amino-Terminal Region of the Signal Peptide in Protein Secretion across the Membrane. Proceedings of the National Academy of Sciences of the United States of America-Biological Sciences **79:** 3438-3441.

JELSCH, C., L. MOUREY, J. M. MASSON and J. P. SAMAMA. 1993. Crystal-Structure of Escherichia-Coli Tem1 Beta-Lactamase at 1.8-Angstrom Resolution. Proteins-Structure Function and Genetics **16:** 364-383.

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.

KONCAN, R., A. VALVERDE, M. I. MOROSINI, M. GARCIA-CASTILLO, R. CANTON *et al.* 2007. Learning from mistakes: Taq polymerase contaminated with beta-lactamase sequences results in false emergence of Streptococcus pneumoniae containing TEM. J. Antimicrob. Chemother. **60:** 702-703.

KOPSIDAS, G., R. K. CARMAN, E. L. STUTT, A. RAICEVIC, A. S. ROBERTS *et al.* 2007. RNA mutagenesis yields highly diverse mRAN libraries for *in vitro* protein evolution. BMC Biotechnol. **7**.

LENFANT, F., A. PETIT, R. LABIA, L. MAVEYRAUD, J. P. SAMAMA *et al.* 1993. Site-Directed Mutagenesis of Beta-Lactamase Tem-1 - Investigating the Potential Role of Specific Residues on the Activity of Pseudomonas-Specific Enzymes. Eur. J. Biochem. **217:** 939-946.

LONG-MCGIE, J., A. D. LIU and V. SCHELLENBERGER. 2000. Rapid in vivo evolution of a beta-lactamase using phagemids. Biotechnol. Bioeng. **68:** 121-125.

MATAGNE, A., J. LAMOTTE-BRASSEUR and J. M. FRERE. 1998. Catalytic properties of class A beta-lactamases: efficiency and diversity (vol 330, pg 581, 1998). Biochem. J. **331:** 975-975.

MEDEIROS, A. A. 1997. Evolution and dissemination of beta-lactamases accelerated by generations of beta-lactam antibiotics. Clin. Infect. Dis. **24:** S19-S45.

MUGNIER, P., P. DUBROUS, I. CASIN, G. ARLET and E. COLLATZ. 1996. A TEM-derived extended-spectrum beta-lactamase in Pseudomonas aeruginosa. Antimicrob. Agents Chemother. **40:** 2488-2493.

ORENCIA, M. C., J. S. YOON, J. E. NESS, W. P. STEMMER and R. C. STEVENS. 2001. Predicting the emergence of antibiotic resistance by directed evolution and structural analysis. Nat. Struct. Biol. **8:** 238-242.

OSUNA, J., A. PEREZ-BLANCAS and X. SOBERON. 2002. Improving a circularly permuted TEM-1 beta-lactamase by directed evolution. Protein Eng. **15:** 463-470.

PAI, H. J., H. J. LEE, E. H. CHOI, J. KIM and G. A. JACOBY. 2001. Evolution of TEM-related extended-spectrum beta-lactamases in Korea. Antimicrob. Agents Chemother. **45:** 3651-3653.

PALZKILL, T., and D. BOTSTEIN. 1992. Probing Beta-Lactamase Structure and Function Using Random Replacement Mutagenesis. Proteins-Structure Function and Genetics **14:** 29-44.

PALZKILL, T., Q. Q. LE, K. V. VENKATACHALAM, M. LAROCCO and H. OCERA. 1994. Evolution of Antibiotic-Resistance - Several Different Amino-Acid Substitutions in an Active-Site Loop Alter the Substrate Profile of Beta-Lactamase. Mol. Microbiol. **12:** 217-229.

PERILLI, M., B. SEGATORE, M. R. DE MASSIS, L. PAGANI, F. LUZZARO *et al.* 2002. Biochemical characterization of TEM-92 extended- spectrum beta-lactamase, a protein differing from TEM-52 in the signal peptide. Antimicrob. Agents Chemother. **46:** 3981-3983.

SALVERDA, M., and J. A. G. M. DE VISSER. unpublished.

SCHNEIDER, I., R. MARKOVSKA, E. KEULEYAN, M. SREDKOVA, K. RACHKOVA *et al.* 2007. Dissemination and persistence of a plasmid-mediated TEM-3-like beta-lactamase, TEM-139, among Enterobacteriaceae in Bulgaria. Int. J. Antimicrob. Agents **29:** 710-714.

SHORT, J. M., J. M. FERNANDEZ, J. A. SORGE and W. D. HUSE. 1988. Lambda-Zap - a Bacteriophage Lambda-Expression Vector with Invivo Excision Properties. Nucleic Acids Res. **16:** 7583-7600.

SIDERAKI, V., W. Z. HUANG, T. PALZKILL and H. F. GILBERT. 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.

SOUGAKOFF, W., A. PETIT, S. GOUSSARD, D. SIROT, A. BURE *et al.* 1989. Characterization of the Plasmid Genes Blat-4 and Blat-5 Which Encode the Broad-Spectrum Beta-Lactamases Tem-4 and Tem-5 in Enterobacteriaceae. Gene **78:** 339-348.

STEMMER, W. P. 1994. Rapid evolution of a protein in vitro by DNA shuffling. Nature **370:** 389-391.

VAKULENKO, S., and D. GOLEMI. 2002. Mutant TEM beta-lactamase producing resistance to ceftazidime, ampicillins, and beta-lactamase inhibitors. Antimicrob. Agents Chemother. **46:** 646-653.

VAKULENKO, S. B., B. GERYK, L. P. KOTRA, S. MOBASHERY and S. A. LERNER. 1998. Selection and characterization of beta-lactam-beta-lactamase inactivator-resistant mutants following PCR mutagenesis of the TEM-1 beta-lactamase gene. Antimicrob. Agents Chemother. **42:** 1542-1548.

VAKULENKO, S. B., P. TAIBI-TRONCHE, M. TOTH, I. MASSOVA, S. A. LERNER *et al.* 1999. Effects on substrate profile by mutational substitutions at positions 164 and 179 of the class A TEMpUC19 beta-lactamase from Escherichia coli. J. Biol. Chem. **274:** 23052-23060.

VAKULENKO, S. B., M. TOTH, P. TAIBI, S. MOBASHERY and S. A. LERNER. 1995. Effects of Asp-179 Mutations in Tem(Puc19) Beta-Lactamase on Susceptibility to Beta-Lactams. Antimicrob. Agents Chemother. **39:** 1878-1880.

WANG, X. J., G. MINASOV and B. K. SHOICHET. 2002. Evolution of an antibiotic resistance enzyme constrained by stability and activity trade-offs. J. Mol. Biol. **320:** 85-95.

WEINREICH, D. M., N. F. DELANEY, M. A. DEPRISTO and D. L. HARTL. 2006. Darwinian evolution can follow only very few mutational paths to fitter proteins. Science **312:** 111-114.

YANG, Y. J., B. A. RASMUSSEN and D. M. SHLAES. 1999. Class A beta-lactamases - enzyme-inhibitor interactions and resistance. Pharmacol. Ther. **83:** 141-151.

ZACCOLO, M., and E. GHERARDI. 1999. The effect of high-frequency random mutagenesis on in vitro protein evolution: A study on TEM-1 beta-lactamase. J. Mol. Biol. **285:** 775-783.

ZARNAYOVA, M., E. SIEBOR, A. PECHINOT, J. M. DUEZ, H. BUJDAKOVA *et al.* 2005. Survey of Enterobacteriaceae producing extended-spectrum beta-lactamases in a Slovak hospital: Dominance of SHV-2a and characterization of TEM-132. Antimicrob. Agents Chemother. **49:** 3066-3069.

-Chapter 4-

Evidence for recombination among plasmidic TEM-alleles

M. Barlow[1] *& M. Salverda*[2]

[1]School of Natural Sciences, University of California Merced, U.S.A.

[2]Labortory of Genetics, Wageningen University, the Netherlands

Abstract: The TEM ß-lactamases are among the most widely dispersed and frequently encountered resistance determinants in clinical microbial populations. There are currently ~160 variants of the TEM ß-lactamase that vary in amino acid sequence. Typically, the variation among TEMs has been attributed to the occurrence of point mutations. Therefore recombination has not been considered an important or regularly occurring process among TEM-alleles. We present evidence for recombination that is based on phylogenetic interpretation of the occurrence of silent mutations in clinical isolates of TEM variants. We also show that the distribution of co-occurring TEM-alleles in clinical microbial populations is consistent with the regular occurrence of recombination among TEM-alleles. Our findings indicate that recombination has had an important effect on the sequence evolution and population distribution of TEM-alleles.

4.1 Introduction

The TEM ß-lactamases are one of the best-studied families of resistance determinants. They are encountered throughout the world at high frequency among clinical microbial populations, and they confer resistance to the heavily used ß-lactam antimicrobials (PATERSON and BONOMO 2005). The progenitor of this family, TEM-1 was first isolated in 1963 as a plasmidic penicillin resistance determinant (DATTA and KONTOMICHALOU 1965). Since that time, TEM-1 has given rise to ~160 variants (http://www.lahey.org/Studies/temtable.asp) that differ in amino acid sequence and that often differ in the phenotype they confer (PATERSON and BONOMO 2005). The TEM enzymes have exhibited considerable functional plasticity and are able to confer resistance to most ß-lactams. Extended spectrum TEMs (ESBLs) confer resistance to penicillins, 2nd and 3rd generation cephalosporins, and monobactams while inhibitor resistant TEMs confer resistance to penicillins and penicillin / inhibitor combinations (HELFAND *et al.* 2003; PATERSON and BONOMO 2005). The functional plasticity of the TEMs and the numerous publications documenting the evolution of the TEMs has made them an ideal model system for studying molecular evolution, phenotypic trade-offs and evolutionary potential. The TEMs are also an ideal model for studying evolutionary and population processes in microbes because of numerous hospital surveillance studies, which have documented their frequencies and distributions. Despite the numerous studies of TEM molecular evolution, there are few reports of sequence evolution occurring through homologous recombination (BARANIAK *et al.* 2005) of TEM-alleles. Therefore, recombination is not generally considered as a mechanism of importance in the evolution of TEM-alleles.

Recent evidence that recombination occurs regularly among plasmidic *qnr* resistance alleles indicates that recombination may be a regular occurrence among allelic variants of plasmidic resistance genes (Baquirin and Barlow 2007 (submitted)). Typically, recombination in microbes is studied in the chromosome with respect to clonality and parasexual exchange (MILKMAN 1997; KRAFT *et al.* 2006). Recombination of plasmidic DNA has also been studied with respect to increasing the stability of artificial plasmidic constructs (CASALI 2003). Occasional incidences of recombination among plasmidic alleles have been reported, however the occurrence of recombination among plasmids is not generally considered a process that has important consequences upon microbial populations (BARANIAK

et al. 2005). We have used the extensive data available about the synonymous variation and population distribution of plasmidic TEM-alleles to study the occurrence and the effects of recombination among TEM-alleles.

4.2 Materials and methods

Phylogenetic reconstruction: A maximum likelihood phylogeny of TEM-alleles was generated with Phyml v2.4.5 using a general time reversible substitution matrix, a discrete gamma model with 4 categories and an estimated shape parameter of 0.338. The proportion of invariant sites was estimated as 0.448. The tree was rooted using a $bla_{\text{SHV-2}}$ allele as the outgroup. A summary of that phylogeny is shown in Figure 4.1.

Analysis of silent mutations: There are 253 silent sites within TEM-alleles. In our phylogeny, 61 silent mutations occurred throughout the TEM phylogeny. Among all TEM-alleles, there is an average of 0.2411 silent mutations per silent site. This mean was used to compute the expected number of times that mutations would occur at a single site according to the Poisson distribution (Table 4.1). The observed number of times that point mutations occurred at each site was inferred from the phylogeny and the observed and expected values were compared using a G-test.

Population analysis: Data were compiled from ten surveillance studies (DECRE *et al.* 2004; KIM *et al.* 2004; KRUGER *et al.* 2004; LAVIGNE *et al.* 2004; BARANIAK *et al.* 2005; BRINAS *et al.* 2005; HO *et al.* 2005a; HO *et al.* 2005b; SCHLESINGER *et al.* 2005; TASLI and BAHAR 2005) in which bacteria exhibiting the extended-spectrum ß-lactamase (ESBL) phenotype were screened by isoelectric focusing or sequence analysis to identify genes conferring the ESBL phenotype. While the entire data set includes frequencies of Class A, Class C and Class D ß-lactamases, our analysis included only Class A ß-lactamases. The entire aggregated dataset used for this analysis contained 291 strains. Of those, 163 strains expressed a single Class A enzyme, 102 strains expressed two Class A enzymes, 25 strains expressed more than two Class A enzymes, and one strain expressed no Class A enzymes. The 25 strains that expressed more than two Class A enzymes were excluded from the analysis because sample size was insufficient for significant statistical analysis.

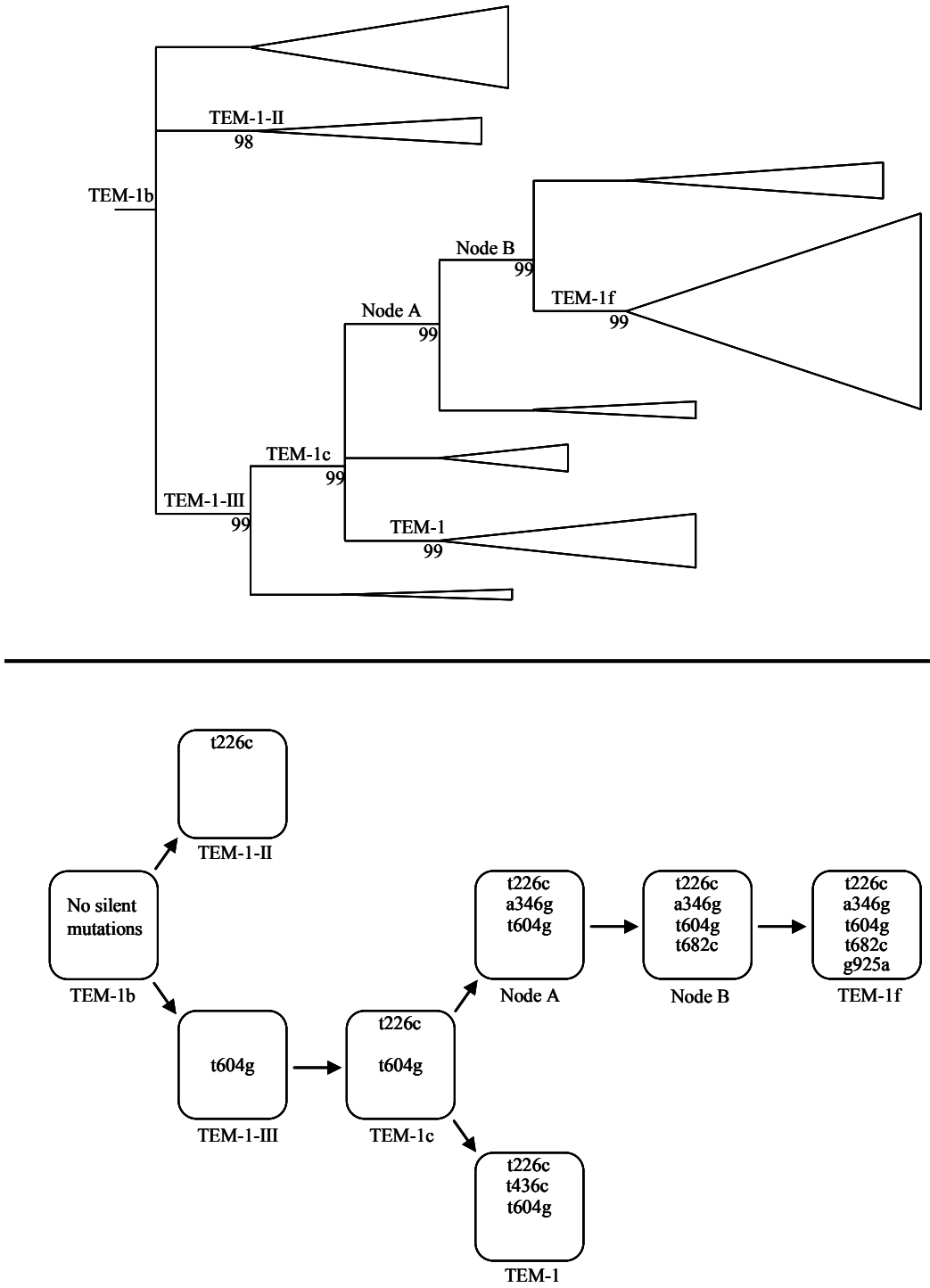


Figure 4.1: The occurrence of synonymous mutations in TEM-alleles. *Top panel*: phylogeny showing order of descent of alleles in which the silent mutations t226c, a346g, t604g, t682c, and g925a arose. Chi square likelihood ratios showing the confidence of the node are shown at each node where a synonymous mutation occurred. *Lower panel*: pathways of silent mutation occurrence in TEM-alleles. This diagram shows the order in which synonymous mutations arose. The nodes of the phylogeny at which synonymous mutations arose are represented with boxes. Inside each box is a list of the mutations that occur at the node.

Three families of Class A enzymes were identified among the isolates included in this analysis: TEM, SHV, and CTX-M. The frequency of each family was computed among all isolates included in the analysis. The frequencies of each family were then multiplied to predict the frequency of co-occurring alleles from the Class A families. Chi square analysis was used to compute the probability that the co-occurrence of alleles results from their frequencies in populations.

Table 4.1: Probability* that observed distribution of silent variation resulted from spontaneous point mutations.

Number of mutations	Poisson probability	Expected number of sites	Observed number of sites
0	0.7858	198.8074	235
1	0.18945	47.93085	10
2	2.28×10^{-2}	5.77852	1
3	1.80×10^{-3}	0.4554	1
4	1.10×10^{-4}	0.02783	0
5	5.34×10^{-6}	0.00134976	1
6	2.14×10^{-7}	5.4243×10^{-5}	2
7	7.39×10^{-9}	1.8684×10^{-6}	1
8	2.23×10^{-10}	5.6318×10^{-8}	0
9	5.96×10^{-12}	1.5084×10^{-9}	0
10	1.44×10^{-13}	3.6432×10^{-11}	1
11	3.15×10^{-15}	7.972×10^{-13}	0
12	6.33×10^{-17}	1.6017×10^{-14}	1

*χ^2 p-value = 6.82×10^{-45}

4.3 Results and Discussion

We analyzed the silent variation of TEM-alleles to determine whether recombination is detectable. Silent mutations usually occur randomly and they usually are not subject to selection. Therefore, independent occurrences of silent mutations that result from spontaneous point mutations should be Poisson distributed. Non-synonymous variation was excluded from our analysis because some recurrent amino acid replacement could be selected for as a means of adapting to new substrates and could therefore appear to arise more often than expected by

chance alone, while silent mutations are essentially neutral and their occurrence is expected to be random.

The allele TEM-1b is the ancestor of the TEM-alleles and was used as a reference sequence for identifying silent mutations. There are six silent mutations that occur repeatedly among TEM-alleles. They are t226c, a346g, t436c, t604g, t682c, g925a. The allele TEM-1b is the ancestor of the TEM-alleles and was used as a reference sequence for identifying silent mutations. The order in which those mutations arose is presented in Figure 4.1. Chi square likelihood ratio testing supports this order of occurrence as all nodes at which these mutations occurred had likelihood ratios of $\geq$ .98. The mutation t226c arose twice and could have resulted either from independent point mutations or transfer of the mutation by recombination into an additional TEM lineage.

By inferring the order in which silent mutations have occurred among TEM-alleles, it is also possible to infer combinations of mutations that are inconsistent with sequential acquisition of point mutations and that are more likely to occur through a recombination event. The silent mutations found in the TEM-68 allele would have required two simultaneous reversion mutations at sites 226 and 346 of a TEM-1f descendent, but they are consistent with a single recombination event between an allele descended directly from TEM-1b and another descended from TEM-1f. Assuming that a single recombination event is more likely than two simultaneous reversions, the silent mutations found in the TEM-68 allele provide evidence that recombination within the TEM-alleles does occur.

The pattern of mutations observed in 32 additional TEM-alleles could be explained either by recombination or by mutation/reversion of a single site among the six sites listed previously. For any single TEM allele, it would be impossible to distinguish whether recombination or the occurrence of a point mutation led to the pattern of silent mutations. However, when the entire phylogeny is taken into consideration, it becomes possible to detect whether the occurrence of silent mutations is consistent with spontaneous point mutations.

The occurrence of point mutations is consistent with the Poisson distribution because point mutations occur randomly and infrequently. A total of 61 silent mutations occurred throughout the TEM phylogeny. There are 253 silent sites within TEM-alleles. Among all TEM-alleles, there is an average of 0.2411 silent mutations per silent site. This mean was used to compute the expected number of times that mutations would occur at a single site according to the Poisson distribution (Table 4.1). The observed number of times that point

mutations occurred at each site was inferred from the phylogeny and the observed and expected values were compared using a G-test. The observed and expected distributions of point mutations were significantly different ($p=6.82 \times 10^{-45}$) which indicates that the observed pattern of silent mutations is not consistent with the occurrence of spontaneous point mutations. To eliminate the possibility that the difference may be an artifact caused by under-sampling silent mutations due to biases against reporting TEM-alleles that contain only silent variation, the mean number of silent mutations per silent site used for computing the expected values was increased by an order of magnitude. This resulted in the observed pattern of silent mutations becoming even less likely to have occurred by silent point mutations ($p=2.04 \times 10^{-219}$). These results indicate that the pattern of silent variation among TEM-alleles cannot be explained by spontaneous point mutations. The most reasonable alternative is that recombination is responsible for the seemingly independent occurrence of pre-existing point mutations among independent lineages of the TEM-alleles. These results show that recombination has occurred regularly among TEM-alleles.

To determine the amount of variation in TEM-alleles that has resulted from recombination, we estimated the frequency of alleles that have arisen by spontaneous point mutations and the frequency of those that have arisen by recombination. There are 61 independent occurrences of silent mutations that occur at 18 sites in the TEM gene. There are 7 sites at which 49 of the 61 silent mutations occur (Table 4.1). The difference between the expected and observed number of occurrences of mutations at those sites exceeds two orders of magnitude. By assuming that all variation at those sites, except for the first mutation, arose by recombination we can infer that 42 mutations or 69% of silent variation resulted from recombination occurring between variant TEM-alleles. It is reasonable to assume that recombination has also resulted in disseminating non-synonymous variation among the TEM-alleles and that a similar proportion of amino acid substitutions may have likewise resulted from recombination. Conservatively, we can conclude that recombination contributes as much to variation as does mutation.

Phylogenetic evidence indicates that recombination occurs regularly among TEM-alleles. Therefore, recombination may result in other detectable patterns among microbial populations that carry TEM-alleles. If two TEM-alleles co-occur on a single plasmid and are oriented in the same direction, a single recombination event between them will remove one allele from the plasmid and one will remain. Depending on the recombination site, the remaining allele

may be a chimera derived from the two alleles that previously co-occurred. If the co-occurring TEM-alleles are oriented in opposite directions, a single recombination event between them will result in the rotation of the DNA segment between the recombination sites. Assuming that TEM-alleles orientation is random and that there is an equal probability of either orientation, and that recombination occurs regularly among TEM-alleles, one would predict that approximately half of all co-occurrences of TEM-alleles would be subject to elimination by recombination.

To test this expectation, we compiled data from ten surveillance studies (DECRE *et al.* 2004; KIM *et al.* 2004; KRUGER *et al.* 2004; LAVIGNE *et al.* 2004; BARANIAK *et al.* 2005; BRINAS *et al.* 2005; HO *et al.* 2005a; HO *et al.* 2005b; SCHLESINGER *et al.* 2005; TASLI and BAHAR 2005) in which bacteria exhibiting the extended-spectrum ß-lactamase (ESBL) phenotype were screened by isoelectric focusing or sequence analysis to identify genes conferring the ESBL phenotype. Three families of Class A enzymes were identified among the isolates included in this analysis. They are the TEM, SHV, and CTX-M families. The frequency of each family was computed among all isolates included in the analysis. The frequencies of each family were then multiplied to predict the frequency of co-occurring alleles from the Class A families. Chi square analysis of the co-occurrence of Class A ß-lactamases revealed that the observed frequencies of co-occurring alleles from a single family are significantly lower than expected and that the frequencies of co-occurring alleles from different families are significantly higher than expected ($p=1.64 \times 10^{-7}$) (Table 4.2). These results indicate suppression of two allelic variants from one family co-occurring in a single strain. This result rejects the hypothesis that the co-occurrences of Class A ß-lactamases depends upon their frequencies in microbial populations and indicates that recombination affects both sequence evolution and population distributions of TEM-alleles. We predicted that ~55 isolates should co-express two TEM enzymes and found that only 31 (56.4%) did. This finding is consistent with random orientation of co-occurring TEM-alleles and removal of a TEM allele when the co-occurring alleles are oriented in the same direction.

It is interesting to note that in the aggregated bacterial population used for this study, the TEM-1 protein was the most frequently recovered Class A ß-lactamase, despite the fact that inclusion in the aggregated dataset was based on resistance to extended spectrum ß-lactams. The frequent expression of TEM-1 in clinical bacterial populations and the less than expected frequency of co-occurring TEMs in bacterial isolates suggests that there may be an advantage

Table 4.2: Probability** that co-expression of Class-A ß-lactamases results from their frequencies in bacterial populations.

	Observed number of occurrences	Expected number of occurrences
Single TEM	140	
Single CTX-M	15	
Single SHV	8	
TEM + TEM	31	55.21
CTX-M + CTX-M	0	1.04
SHV + SHV	0	2.73
TEM + CTX-M	19	15.13
TEM + SHV	49	24.54
SHV + CTX-M	3	3.36

**$\chi 2$ p-value = 1.64×10^{-7}

for bacteria that express TEM-1 rather than an extended spectrum TEM enzyme. It has been shown that expression of TEM-1 confers a fitness advantage relative to strains expressing extended spectrum TEMs when bacteria are exposed to ampicillin (MROCZKOWSKA and BARLOW 2008). In a recent report of antimicrobial consumption in European countries, penicillin consumption consistently exceeded cephalosporin consumption. This pattern of antimicrobial consumption supports the explanation that the prevalence of TEM-1 expression in microbial populations has resulted from a selective advantage conferred by TEM-1.

These results further indicate that the combined effects of selection and recombination may actually suppress the occurrence of extended spectrum TEMs in microbial populations. By suppressing the co-occurrence of multiple TEM-alleles in a bacterial strain, recombination causes strains to carry, in general, only one TEM allele. A selective advantage of strains expressing TEM-1 over strains expressing extended-spectrum TEMs may cause the strains that express TEM-1 to dominate clinical microbial populations (MROCZKOWSKA and BARLOW 2008). Although extended spectrum TEMs are frequently encountered, none have been encountered as frequently as TEM-1 (DECRE *et al.* 2004; KIM *et al.* 2004; KRUGER *et al.* 2004; LAVIGNE *et al.* 2004; BARANIAK *et al.* 2005; BRINAS *et al.* 2005; HO *et al.* 2005a; HO *et al.* 2005b; SCHLESINGER *et al.* 2005; TASLI and BAHAR 2005). While cephalosporin resistance is common, it is not as common as penicillin resistance. It is possible that resistance to cephalosporins has been somewhat inhibited by the combined effects of selection and

recombination. It is likewise possible that an increased understanding of these effects may lead to novel strategies to extend the useful period of antimicrobials.

References

BAQUIRIN, M. H., and M. BARLOW. 2007 (submitted). Evolution and recombination of the plasmidic qnr alleles Submitted.

BARANIAK, A., J. FIETT, A. MROWKA, J. WALORY, W. HRYNIEWICZ *et al.* 2005. Evolution of TEM-type extended-spectrum beta-lactamases in clinical Enterobacteriaceae strains in Poland. Antimicrob. Agents Chemother. **49:** 1872-1880.

BRINAS, L., M. LANTERO, I. DE DIEGO, M. ALVAREZ, M. ZARAZAGA *et al.* 2005. Mechanisms of resistance to expanded-spectrum cephalosporins in Escherichia coli isolates recovered in a Spanish hospital. J. Antimicrob. Chemother. **56:** 1107-1110.

CASALI, N. 2003. Escherichia coli host strains. Methods Mol. Biol. **235:** 27-48.

DATTA, N., and P. KONTOMICHALOU. 1965. Penicillinase Synthesis Controlled by Infectious R Factors in Enterobacteriaceae. Nature **208:** 239-241.

DECRE, D., B. BURGHOFFER, V. GAUTIER, J. C. PETIT and G. ARLET. 2004. Outbreak of multi-resistant Klebsiella oxytoca involving strains with extended-spectrum beta-lactamases and strains with extended-spectrum activity of the chromosomal beta-lactamase. J. Antimicrob. Chemother. **54:** 881-888.

HELFAND, M. S., M. A. TOTIR, M. P. CAREY, A. M. HUJER, R. A. BONOMO *et al.* 2003. Following the reactions of mechanism-based inhibitors with beta-lactamase by Raman crystallography. Biochemistry **42:** 13386-13392.

HO, P. L., A. Y. M. HO, K. H. CHOW, R. C. W. WONG, R. S. DUAN *et al.* 2005a. Occurrence and molecular analysis of extended-spectrum beta-lactamase-producing Proteus mirabilis in Hong Kong, 1999-2002. J. Antimicrob. Chemother. **55:** 840-845.

HO, P. L., R. H. L. SHEK, K. H. CHOW, R. S. DUAN, G. C. MAK *et al.* 2005b. Detection and characterization of extended-spectrum beta-lactamases among bloodstream isolates of Enterobacter spp. in Hong Kong, 2000-2002. J. Antimicrob. Chemother. **55:** 326-332.

KIM, S., J. KIM, Y. H. KANG, Y. PARK and B. LEE. 2004. Occurrence of extended-spectrum genus Shigella in the beta-lactamases in members of the Republic of Korea. Journal of Clinical Microbiology **42:** 5264-5269.

KRAFT, C., A. STACK, C. JOSENHANS, E. NIEHUS, G. DIETRICH *et al.* 2006. Genomic changes during chronic Helicobacter pylori infection. J. Bacteriol. **188:** 249-254.

KRUGER, T., D. SZABO, K. H. KEDDY, K. DEELEY, J. W. MARSH *et al.* 2004. Infections with nontyphoidal Salmonella species producing TEM-63 or a novel TEM enzyme, TEM-131, in South Africa. Antimicrob. Agents Chemother. **48:** 4263-4270.

LAVIGNE, J. P., N. BOUZIGES, C. CHANAL, A. MAHAMAT, S. MICHAUX-CHARACHON *et al.* 2004. Molecular epidemiology of Enterobacteriaceae isolates producing extended-spectrum beta-lactamases in a French hospital. Journal of Clinical Microbiology **42:** 3805-3808.

MILKMAN, R. 1997. Recombination and population structure in Escherichia coli. Genetics **146:** 745-750.

MROCZKOWSKA, J. E., and M. BARLOW. 2008. Fitness trade-offs in bla(TEM) evolution. Antimicrob. Agents Chemother. **52:** 2340-2345.

PATERSON, D. L., and R. A. BONOMO. 2005. Extended-spectrum beta-lactamases: a clinical update. Clin. Microbiol. Rev. **18:** 657-+.

SCHLESINGER, J., S. NAVON-VENEZIA, I. CHMELNITSKV, O. HAMMER-MUNZ, A. LEAVITT *et al.* 2005. Extended-spectrum beta-lactamases among Enterobacter isolates obtained in Tel Aviv, Israel. Antimicrob. Agents Chemother. **49:** 1150-1156.

TASLI, H., and I. H. BAHAR. 2005. Molecular characterization of TEM- and SHV-derived extended-spectrum beta-lactamases in hospital-based Enterobacteriaceae in Turkey. Jpn. J. Infect. Dis. **58:** 162-167.

-Chapter 5-

Adaptive significance of silent mutations in an antibiotic resistance gene

M. Salverda[1], B. van Putten[2], B. Koopmanschap[1], R.F. Hoekstra[1], A.J.M. Debets[1] & J.A.G.M. de Visser[1]

[1]Laboratory of Genetics, Wageningen University, the Netherlands

[2]Mathematical and Statistical Methods Group, Wageningen University, the Netherlands

Abstract: Silent or synonymous mutations have long been thought to be functionally and adaptively neutral. However, several mechanisms that can cause synonymous mutations to have functional effects have recently been identified. The availability of many antibiotic resistance genes from clinical isolates and the ease with which they can be used in laboratory experiments offer an excellent opportunity for studying the adaptive significance of silent mutations. Here, we examine synonymous mutations present in clinical isolates of the antibiotic resistance gene TEM-1, as well as synonymous mutations present in laboratory isolates of this gene that were isolated after directed evolution for increased resistance to the antibiotic cefotaxime. It was found that among laboratory isolates at least one synonymous mutation occurs at a frequency that is too high to be explained by mere chance, while another mutation conferred a direct adaptive benefit at low cefotaxime concentrations. Among clinical TEM-isolates, indications exist that similar functional-yet-synonymous mutations may be present. Putative mechanisms underlying the adaptive benefit are explored using information on mRNA stability and codon usage.

5.1 Introduction

Owing to the redundancy of the genetic code, most amino acids are encoded by more than one codon (a triplet of nucleotides). Codons encoding the same amino acid are called synonymous, because they use a different 'word' (the nucleotides) to deliver exactly the same 'message' (the amino acid encoded). Synonymous mutations are also known as 'silent' mutations, because of the once widely accepted notion that mutations cannot alter gene function or phenotype when they do not change the amino acid sequence (ANFINSEN 1973). However, in the past decades evidence has accumulated that silent mutations are not necessarily functionally silent; silent mutations have been found to cause functional effects by affecting translation efficiency, splicing, miRNA binding, mRNA folding and translational timing (PARMLEY and HURST 2007).

The evolution of antibiotic resistance genes offers an excellent opportunity for tracing and studying functional-yet-silent mutations, since both the natural and *in vitro* evolution of some of these genes in response to challenges with novel antibiotics has been recorded in great detail. Here, we study the silent mutations present in two datasets, one of clinical isolates and one of laboratory isolates of the TEM family of antibiotic resistance genes.

TEM-1 is one of the best studied and widespread antibiotic resistance genes around. It encodes the enzyme TEM-1 ß-lactamase, which confers resistance to most ß-lactam antibiotics like penicillins and early cephalosporins (MATAGNE *et al.* 1998). Mutations in the gene can expand this resistance spectrum to more modern cephalosporins, monobactams and penicillin/ inhibitor combinations. It is generally thought that the many different TEM-alleles that have been isolated in the past decades are the result of increased and novel antibiotic pressures (BARLOW and HALL 2002). Since the first TEM-allele was isolated in 1963 (DATTA and KONTOMICHALOU 1965), more than 160 TEM-alleles have been identified, each with a unique amino acid sequence. TEM-1 has also been used to predict the evolution of antibiotic resistance in the laboratory (VAKULENKO *et al.* 1998; BARLOW and HALL 2002; BARLOW and HALL 2003).

We have sequenced 65 TEM-alleles isolated in separate experiments of *in vitro* mutagenesis followed by *in vivo* selection for increased resistance to the antibiotic cefotaxime (Salverda and de Visser unpublished). Next to a large number of nonsynonymous mutations, 85 silent mutations were identified in this dataset. Several of these silent mutations were

found in two, three or even four independently evolved alleles, suggesting that they may have been selected. Here, we provide direct experimental evidence that at least one of these frequently occurring silent mutations increases resistance to cefotaxime. We also present statistical evidence that another silent mutation with beneficial effect is present in our dataset. Apart from this, we have analysed the silent variation present in 175 unique clinical isolates of TEM-alleles and we find indications that at least some of these silent mutations could also have been selected for their improved functional effects.

5.2 Materials and methods

Clinical dataset: Sequence data of 175 clinical isolates of TEM that differ in their nucleotide sequence were gathered using the references in the online database of clinically isolated TEM-variants (http://www.lahey.org/studies/temtable.asp) and GenBank.

Experimental dataset: The experimental dataset consists of the sequences of 65 TEM-alleles isolated after *in vitro* mutagenesis of TEM-1 followed by *in vivo* selection for increased cefotaxime resistance. The exact experimental procedures used in these experiments are explained in detail below. It is important to note that all these alleles have been isolated under identical mutagenic and selective conditions, which are detailed below.

Bacterial strains and plasmids: *Escherichia coli* strain DH5αE (Invitrogen) was used as the host for all plasmids. Plasmid pACSE3 (BARLOW and HALL 2002) was used as the vector for cloning and expressing TEM alleles.

Media: LB-broth is 10g trypticase peptone, 5 g yeast extract and 10 g NaCl/ litre. LB-Tet medium is LB-broth containing 15 mg tetracycline/ litre. Mueller Hinton broth (Merck) was prepared according to the manufacturer's instructions. SOC medium is 20 g bacto-tryptone and 5 g yeast extract/ litre with 10 mM NaCl, 2.5 mM KCl, 10 mM $MgCl_2$ and 20 mM glucose. Solid media contained 16 g agar/ litre.

Mutagenesis: TEM-1 alleles were mutagenized using the Genemorph I and Genemorph II PCR Mutagenesis Kits (Stratagene) according to the manufacturer's instructions. Primers P3

(TCATCCGGCTCGTATAATGTGGA) and P4 (ACTCTCTTCCGGGCGCTATCAT), were used for PCR. Mutated amplicons were digested with *Bsp*HI and *Sac*I restriction enzymes, ligated into plasmid pACSE3 (BARLOW and HALL 2002) and electroporated into *Escherichia coli* strain DH5αE. After recovery for 90 min in SOC medium at 37°C, the cells were diluted in 500 ml LB-tetracycline broth. An aliquot was taken out of this directly after mixing and plated out on LB-tetracycline to determine the library-size, while the remainder of the culture was placed overnight at 37°C to expand the library.

Selection for extended-spectrum TEM-alleles: A series of bottles containing 50 ml Mueller Hinton broth was inoculated with 50 mM IPTG and cefotaxime in two-fold increments (ranging from 0.125 µg/ ml, the minimal inhibitory concentration of TEM-1, to 1024 µg/ ml). Subsequently, each bottle was inoculated with a number of cells from the overnight-enriched library that was equivalent to 10 times the library-size. Cultures were incubated for 48 hours at 37°C. The culture that grew at the highest concentration of cefotaxime was plated on LB + Tet to select for maintenance of the plasmid, which contains a Tet^r gene. A single genotype is expected to dominate this culture (BARLOW and HALL 2002). The next day a single colony was picked from this plate and grown overnight in LB + Tet. Plasmid from this overnight culture was subsequently isolated using the GenElute™ Plasmid Miniprep Kit (Sigma) and sequenced.

Sequencing: The TEM-allele was amplified from isolated plasmids using primers P3 and P4 and BigDye™ (Perkin Elmer) or DYEnamic™ ET (AP-biotech) Terminator Cycle Sequencing kits according to the manufacturer's instructions.

Antibiotics and MIC-assay: The antibiotics cefotaxime (Ctx) and tetracycline (both Sigma) were used in this study. Stock solutions were made in 0.1M $NaPO_4$, pH 7.0 (cefotaxime) and 70% ethanol (tetracycline). Minimal inhibitory concentrations (MICs, defined here as the lowest concentration of antibiotic that completely inhibits visual bacterial growth after 24 hours) were determined from overnight cultures diluted to a titre of 10^5 cells/ ml in Mueller Hinton II broth containing 50 µm IPTG (to cause expression of TEM by inactivation of the *lacI*-repressor). Aliquots of 150 µl of this diluted culture were added to twofold serial dilutions of cefotaxime in 150 µl Mueller Hinton II in 96-well microtitre plates. In order to

increase accuracy, we used starting concentrations of 0.250 and 0.187 μg Ctx/ ml and diluted with twofold steps from there. The culture used as inoculum was also suitably diluted and plated on LB + Tet in order to check the true inoculum. MIC is defined as the lowest concentration of antibiotic that completely prevents visible growth. Once inoculated, cultures were grown for 24 hours at 37°C, after which MIC was determined by visual inspection. Prior to the MIC assay both alleles tested (TEM-1 and the silent mutant N175*) were placed into an isogenic vector background in order to exclude possible variation at the level of the plasmid. Both plasmids were then transformed into isogenic *E. coli* DH5αE cells by electroporation in order to exclude phenotypic variation in the host.

Annotation of mutations: Silent mutations in the TEM-allele are described by the respective nucleotide in the TEM-1 sequence, the number of the nucleotide according to Sutcliffe (1978) and the nucleotide in the respective mutant sequence. Amino acid numbering is according to Ambler (1991). In the text, some silent mutations are described by the amino acid involved. The mutation is then indicated by a single letter indicating the amino acid according to the IUPAC single letter code, the number of the respective amino acid according to Ambler (1991) and an asterix (indicating that the amino acid has not changed).

5.3 Results & discussion

Several mechanisms that can cause functional effects of silent mutations have been demonstrated in the past two decades (PARMLEY and HURST 2007). Two mechanisms that apply both to prokaryotes and eukaryotes have been particularly well documented in recent years; effects on mRNA secondary structure (NACKLEY *et al.* 2006) and on enzyme folding (KIMCHI-SARFATY *et al.* 2007). Below, we will first statistically analyse the dataset of experimentally obtained silent mutations in TEM-1 in order to address the question whether the frequency of occurrence of some of the silent mutations found is higher than expected under a uniform distribution. Subsequently, we directly measure the resistance of a mutant TEM-allele that differs from TEM-1 by a single silent mutation. Finally, we analyse the datasets of silent mutations in clinical and experimental isolates from the perspective of the two main mechanisms for functional effects of silent mutations described so far; mRNA secondary structure and enzyme folding.

Statistical arguments for the adaptive significance of silent mutations

A total of 85 silent mutations were found in the dataset of TEM-alleles isolated after directed evolution for increased cefotaxime resistance (Table 5.1). Within this dataset, three mutations (T141*, A150* and N175*) were found three times and one (H153*) was found four times. Since there are 616 silent mutations possible in the TEM-1 coding sequence (Table 5.1), it was suspected that the chances of finding the same silent mutation three or four times would be extremely low when 'drawing' 85 times (the number of silent mutations found in the total dataset) with replacement.

Table 5.1: Characterization of the 85 silent mutations found in the experimental dataset.

	Number of possible silent mutations in TEM-1	**Number of times found**				
		1	*2*	*3*	*4*	*Total*
Transitions						
A → G, T → C	165	26	0	1	0	29
G → A, C → T	126	15	2	2	1	29
Transversions						
A → T, T → A	96	8	2	0	0	12
A → C, T → G	93	2	0	0	0	2
G → C, C → G	63	1	0	0	0	1
G → T, C → A	73	10	1	0	0	12
Total	616	62	5	3	1	85

We investigated whether mutations found with high frequency in the experimental dataset occur more often than would be expected under randomness and independency model assumptions. We first focus on the most frequently found silent mutation (H153*, found in four independently evolved alleles) which is caused by a GC → AT transition. By focusing at the specific transition, we eliminate the mutational spectrum of the error-prone polymerase as a factor that can affect the frequency of occurrence of certain mutations. In total, 29 such transitions were found in the experimental dataset (the 'balls' in the statistical model below), while there are 126 different positions in the TEM-1 coding sequence where such a silent transition can be located (the 'bins' in our model).

An appropriate mathematical-statistical model is the following. Given are $N = 126$ bins and $n = 29$ balls.

The null-hypothesis H_0: 'balls are thrown independently and each bin has a constant probability of 1/126 of being hit' is tested against the alternative hypothesis

H_1: 'bins having been hit in the process already have a higher probability to be hit again'.

As a test statistic, we choose M with M the maximum number of hits taken over all bins. Under H_1, high outcomes of M have a higher probability than under H_0, so the rejection region is to be chosen right-hand sided. The experimental dataset and the corresponding Expected numbers under the null-hypothesis based on Binomial calculations, are as follows.

Number of hits	0	1	2	3	4	≥5
Number of bins	106	15	2	2	1	0
Expected number of bins under H_0	100.0035	23.2008	2.5985	0.1871	0.0097	0.0004

Outcome of the test statistic therefore is $M = 4$. In accordance with the right-hand sided rejection region, the significance of the test is P_0 ($M \geq 4$), where P_0 is the probability under the null-hypothesis. If the significance is sufficiently small, the null-hypothesis can be rejected. Let H_{ij} be the event that bin i is hit in attempt j (i = 1, 2, .., 126; j = 1, 2, .., 29). Let A_i be the event that bin i is hit more than or equal to 4 times. From elementary probability theory and symmetry arguments we have

$$P_0(M \geq 4) = P_0\left(\bigcup_{i=1}^{126} A_i\right) \leq \sum_{i=1}^{126} P_0(A_i)$$

$$= 126 * P_0(A_1) = 126 * \binom{29}{4} * P_0\left(H_{11} \cap H_{12} \cap H_{13} \cap H_{14}\right)$$

$$= 126 * \binom{29}{4} * (1/126)^4 \approx 0.0118$$

which supplies sufficient evidence to reject the null-hypothesis in favour of the alternative hypothesis. The conclusion therefore is that this silent mutation occurs too frequently to be explained by random chance if we assume a uniform distribution of chances.

Second dataset

The story is different for the remaining substitution that occurs three times (A150*), since this mutations is caused by another transition (AT → GC). The experimental dataset is in this case as follows.

Number of hits	0	1	2	3	≥4
Number of bins	138	26	0	1	0
Expected number of bins under H_0	138.3315	24.4611	2.0881	0.1146	0.0047

Using the same formulas as described above for $N = 165$, $n = 29$ and $M = 3$, we find P_0 ($M \geq 3$) ≈ 0.134. We conclude that this mutation does not occur at a frequency that is high enough to reject H_0. Overall, the only silent mutation that occurs more frequently than expected under a Binomial distribution is H153*. It thus seems that this silent mutation either confers some selective advantage or is the result of the presence of a mutational hotspot.

MIC measurement of a TEM-1 variant containing a single silent mutation

Next to the statistical arguments presented above, our idea that some of the silent mutations might in fact increase antibiotic resistance was strengthened by an observation in the experiments; one of the silent mutations that was found three times (N175*), was identified as the sole mutation present in a TEM allele that had been isolated after a single round of error-prone PCR and selection at a cefotaxime concentration higher than TEM-1's MIC. Based on this observation, we measured the MIC of four different clones of both TEM-1 and the silent mutant N175* with threefold replication, after having replaced both alleles in the naïve vector background and transformation into isogenic bacteria (Table 5.2). The MIC of the N175* mutant turned out to be significantly higher than that of TEM-1 (Mann-Whitney $U = 3.50$, $P < 0.001$), indicating a selective benefit of this silent mutation.

Next, we analyse the silent mutations found in clinical isolates and those found with high frequency in experiments in order to test the possible contribution of two of the main mechanisms that can cause functional effects, i.e. mRNA folding and protein folding.

Analysis of mRNA secondary structure

There is strong evidence that silent mutations can affect protein translation efficiency by

Table 5.2: MICs of overnight cultures of four different clones of TEM-1 and a silent variant of TEM-1 differing by a single silent mutation (N175*).

Allele	Clone	MIC[a]
TEM-1	1	0.031
	2	0.047
	3	0.031
	4	0.047
N175*	1	0.063
	2	0.094
	3	0.063
	4	0.094

[a] Median value across three replicates.

altering the secondary structure of the mRNA (NACKLEY *et al.* 2006). For instance, it has been demonstrated that in the specific case of TEM-1 the replacement of a stretch of 24 codons by their synonymous counterparts most frequently used in *E. coli* results in an increased mRNA half-life (DEANA *et al.* 1996). We have predicted the mRNA secondary structures of all known clinically isolated silent variants of TEM-1 (TEM-1b, c, d, e, f , g and an additional variant here named TEM-1-II) and of the four silent variants of TEM-1 that contain one of the silent mutations found with high frequency in the experimental dataset (T141*, A150*, H153* and N175*), using the Mfold (MATHEWS *et al.* 1999; ZUKER 2003) and Vienna (HOFACKER *et al.* 1994) programs (Table 5.3). Just a few of the clinically isolated TEM-variants have a Gibbs free energy (ΔG) that is consistently lower than that of TEM-1 (TEM-1d, TEM-1f and TEM-1-II), indicating that they may form a more stable mRNA secondary structure. There are no experimental isolates that have a consistently lower ΔG than TEM-1 (although A150* has a much lower ΔG than TEM-1 according to Mfold). The inconsistency of the outcome of the two methods used makes it hard to draw clear conclusions from these data. We therefore can neither support nor reject the hypothesis that adaptive effects of these silent mutations are the result of changes in mRNA stability. Measurements on mRNA half-life (DEANA *et al.* 1996) could provide more reliable data, but are beyond the scope of this study.

Putative effects on enzyme folding

Location of silent mutations in clinically and experimentally isolated TEM-alleles

It has been noted that rare codons are found more frequently than expected in parts of the sequence that encodes turns, loops and linker regions of proteins (KOMAR *et al.* 1999). It has been suggested that, since such regions are more slowly translated, they may serve as

interpunctuations separating different folding events (PURVIS *et al.* 1987). Therefore, if the silent mutations found in clinical and experimental isolates have functional effects, they may be found more often than expected in these regions.

Table 5.3: Silent variation in clinically isolated silent variants of TEM-1 and changes in Gibbs free energy (ΔG) of the respective mRNAs relative to TEM-1. The ΔΔG of the four silent mutations found with high frequency in experimental isolates is shown below.

Allele	**Occurrence**	**Silent mutation**							**ΔΔG (relative to TEM-1)**	
		c226t	*a346g*	*c436t*	*g604t*	*t682c*	*c913t*	*g925a*	*Mfold*	*Vienna*
TEM-1	14								0	0
TEM-1b	29	X		X	X				1.12	-1.43
TEM-1c	6			X					-0.11	0.19
TEM-1d	1			X			X		-3.85	-0.21
TEM-1e	1	X	X	X	X				1.12	-1.43
TEM-1f	61		X	X		X		X	-0.45	-1.27
TEM-1g	10			X	X				2.6	0.47
TEM-1-II	3	X		X					-1.59	-.143
T141*	3								-2.28	1.39
A150*	3								-4.35	-0.01
H153*	4								0	0
N175*	3								0	0

To test this, we first analysed the distribution of silent mutations in clinical TEM-isolates. Using the references in the Lahey-database and GenBank, we collected 175 unique TEM nucleotide sequences (belonging to 150 unique amino acid sequences). The location of the silent mutations found in these sequences is shown in Figure 5.1 and the details of all mutations depicted in this figure can be found in Table 5.4. Before analysing the distribution of these silent mutations over the various structural regions, we first corrected the data for the number of silent mutations possible per nucleotide and for the spectrum of spontaneous mutations in *E. coli* (BARLOW and HALL 2002). We found no evidence that silent mutations are more or less common then expected in the signal sequence ($\chi^2 = 2.00$, df = 1, P = 0.16), α-helices ($\chi^2 = 1.12$, df = 1, P = 0.29), ß-sheets ($\chi^2 = 1.14$, df = 1, P = 0.28), or the remaining turns, loops and hinges ($\chi^2 = 0.07$, df = 1, P = 0.80). Furthermore, the distribution over these four different regions of TEM-1 combined does not differ significantly from what would be expected under a uniform distribution ($\chi^2 = 3.78$, df = 3, P = 0.29). Concluding, we find no evidence that the location of the silent mutations found in the dataset of clinical isolates is clustered in specific structural regions within TEM-1.

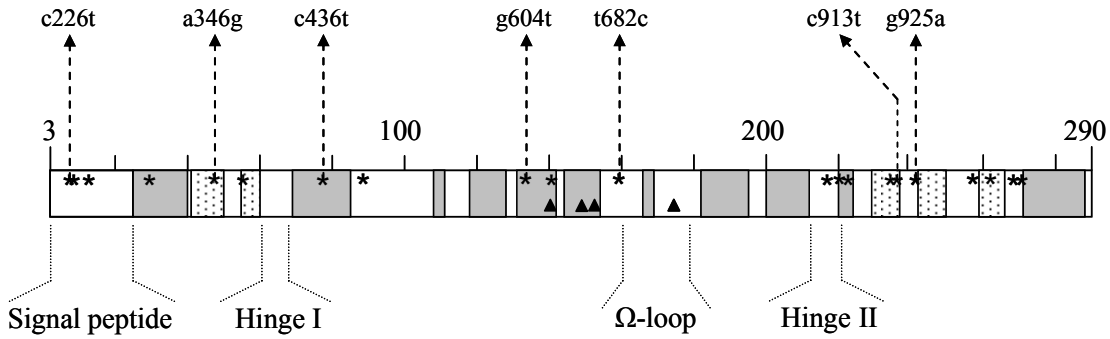


Figure 5.1: Location of silent mutations in TEM-1, as found in clinical isolates (asterices) and directed evolution studies (black triangles). Scale is the amino acid numbering according to Ambler (1991). Structural features like α-helices (grey) and ß-sheets (dotted) have been indicated according to Jelsch *et al.* (1993) and some additional structural features of TEM-1 are shown in text below. Nucleotide substitutions occurring in variants of TEM-1 that only contain silent mutations have been indicated by dotted arrows, the substitutions indicated by the respective nucleotide in TEM-1, the numbering according to Sutcliffe (1978) and the nucleotide in the mutant sequence.

The location of the four silent mutations found more than twice in the experimental dataset is shown in Figure 5.1 (black triangles). All four mutations occur in or in the vicinity of the structurally important Ω-loop (Figure 5.2). Because of the surprisingly short distance between the four mutations, we approximated the chance of finding a maximum distance of 34 amino acids when drawing four times with replacement from 273 possible locations (the total number of amino acids in TEM-1 minus the ones that cannot harbour silent mutations). Drawing ~10,000 sets as described before indicated that this chance is extremely low (~0.007), suggesting that the mutations indeed cluster together in a specific region of the TEM-allele. It is known that the exact positioning of the Ω-loop in and around which these mutations occur (see Figure 5.2) is of crucial importance for TEM-1 activity. On the one hand repositioning of this loop due to nonsynonymous mutations is known to increase TEM-1's resistance to several novel antibiotics (KNOX 1995). On he other hand, the Ω-loop contains several highly conserved amino acids and is relatively sensitive to nonsynonymous mutations when selection for ampicillin resistance is maintained (PALZKILL and BOTSTEIN 1992). It may therefore well be that the silent mutations found in the experimental dataset influence protein folding and thereby the position of the Ω-loop, as will be explained below.

Codon usage in the 15-codon neighbourhood of the experimental silent mutations

It has been hypothesised that protein folding is affected by the *in vivo* translation rate (PURVIS *et al.* 1987). Since translation kinetics are influenced by both the synonymous codon usage

bias and the availability of cognate tRNAs of a specific organism, regions of rare synonymous codons might be translated more slowly (KOMAR and JAENICKE 1995). In turn, protein folding is thought to be a co-translational process and it has been suggested that alteration of the translation kinetics can influence protein folding (KOMAR *et al.* 1999). Recent empirical work showing that silent mutations can indeed affect protein folding supports this hypothesis and it was suggested that the silent mutations involved may contribute to the formation of clusters of rare codons that alter the timing of co-translational folding (KIMCHI-SARFATY *et al.* 2007).

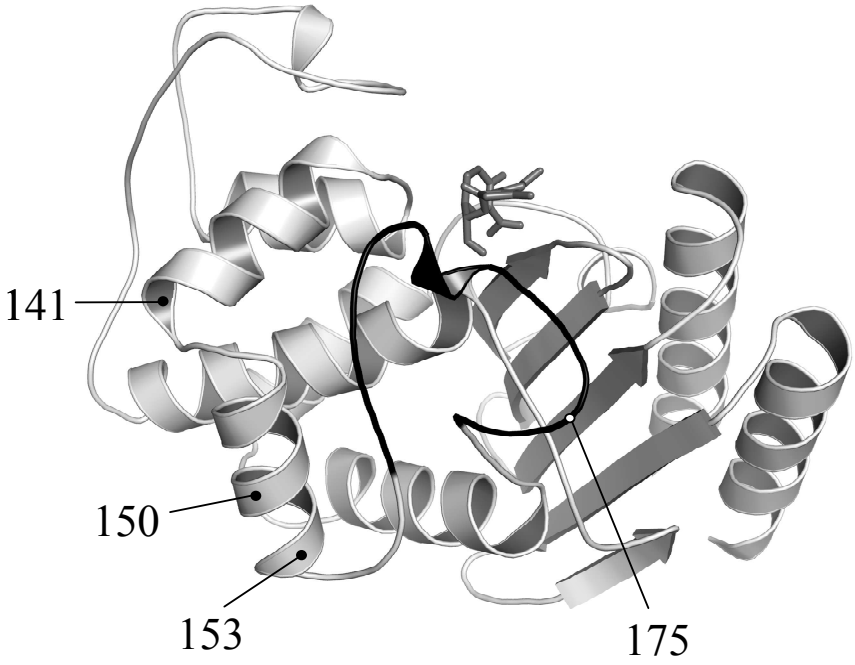


Figure 5.2: Location of the four silent mutations found with high frequency in experimental isolates in the 3d structure of TEM-1 (PDB-code: 1BTL). The Ω-loop is shown in black and a cefotaxime molecule (grey sticks) has been modelled in the binding pocket.

In *E. coli*, known to have a strong codon bias and a level of cognate tRNAs that is directly proportional to codon usage (IKEMURA 1981), rare codons are at the same time slowly translated codons (KOMAR *et al.* 1999). Thus, in *E. coli*, the main host of clinical TEM-alleles and the host of the TEM-alleles in our experimental dataset, mutations that alter a codon to a synonymous but rare codon might affect protein folding according to the model proposed by Kimchi-Sarfaty *et al.* (2007). We have analysed the codon usage in TEM-1 and in the codons altered by the silent mutations in the clinical and experimental datasets using two different codon preference criteria for *E. coli*. Ikemura (1981) used tRNA concentrations and the

efficiency of codon-anticodon interactions to describe a set of non-optimal codons. Because this set does not cover all amino acids, we also identified rare codons by looking at the codon usage in very highly expressed genes in *E. coli* (SHARP and LI 1986). The pattern of codon usage in very highly expressed genes can reveal which of the alternative synonymous codons for an amino acid are the most and the least efficient, since it has been shown that the strong codon usage bias in these genes is directly proportional to the level of cognate tRNAs (IKEMURA 1981).

The codon changes caused by synonymous mutations found in clinical isolates and by the four synonymous mutations found with high frequency in experimental isolates are shown in Table 5.4. The consequences of these changes regarding the two codon preference criteria described above are shown in the same table. Although most synonymous mutations in clinical isolates are towards less optimal or less frequent codons it is hard to draw conclusions from this observation, as it is unknown whether TEM-1 is the 'ancestral' sequence. There are seven variants of TEM-1 that only differ at synonymous sites and judging from the high frequency at which some of these mutations occur in the 175 clinical isolates analysed here, this actually seems unlikely. Our lack of information of the evolutionary relationships of the TEM alleles therefore complicates conclusions about the mechanism underlying a possible adaptive benefit of these silent mutations.

The four synonymous mutations found more than twice in experimental isolates all change their codon to the least optimal (with one exception where criteria do not apply) or least frequent one (Table 5.4). We have analysed the 15-codon neighbourhood of these four synonymous mutations (in analogy to Kimchi-Sarfaty *et al.*) using the codon preference criteria defined by Ikemura (1981) and the Relative Synonymous Codon Usage values (RSCU, i.e. the number of codons observed for a specific amino acid divided by the number of codons expected under a uniform distribution) of very highly expressed genes of *E. coli* (SHARP and LI 1987) (Table 5.5). All four silent mutations lower the RSCU value of their codon to the lowest possible value, while three of the four change the codon to the least preferred according to the Ikemura-criteria (the fourth codon change is not characterised by these criteria). In addition, all four mutations are surrounded by a large number of rare codons according to both methods of analysis, an image similar to that found by Kimchi-Sarfaty *et al.* (2007), which indicates that the mutations might be involved in the formation of clusters of rare codons.

Table 5.4: Silent mutations found in clinical isolates and silent mutations found with high frequency in directed evolution studies.

	Silent mutation[a]	Amino acid	Change	Occurrence	Change according to Ikemura[b]		Change according to Sharp & Li[c]	
					Before	*After*	*Before*	*After*
Clinical isolates	**c226t**	F8	ttc > ttt	42	Optimal	Least optimal	1.54*	0.46**
	t229a	R9	cgt > cga	1	Optimal	Least optimal	4.38*	0.02**
	t241a	I13	att > ata	1	Sub-optimal	Least optimal	0.47	0.01**
	g292t	L30	ctg > ctt	1	Optimal	Least optimal	5.33*	0.23
	a346g	E48	gaa > gag	90	Optimal	Least optimal	1.59*	0.41**
	c370t	I56	atc > att	3	Optimal	Sub-optimal	2.53*	0.47
	c436t	G78	ggc > ggt	149	Optimal	Optimal	1.65	2.28**
	g469a	E89	gag > gaa	2	Least optimal	Optimal	0.41**	1.59*
	g604t	A134	gcg > gct	50	Optimal	Optimal	0.80	1.88*
	g625a	T141	acg > aca	4	Least optimal	Least optimal	0.19	0.14**
	t682c	T160	act > acc	89	Optimal	Optimal	1.80	1.87*
	a856g	G218	gga > ggg	1	Least optimal	Sub-optimal	0.02**	0.04
	c863t	L221	ctg > ttg	1	Optimal	Least optimal	5.33*	0.11
	g871t	S223	tcg > tct	1	-	-	0.04**	2.57*
	t907a	S235	tct > tca	1	-	-	2.57*	0.20
	c913t	A237	gcc > gct	1	Optimal	Optimal	0.23**	1.88*
	g925a	G242	ggg > gga	74	Sub-optimal	Least optimal	0.04	0.02**
	c970t	S258	tcc > tct	1	-	-	1.91	2.57*
	c985t	I263	atc > att	1	Optimal	Sub-optimal	2.53*	0.47
	g1003a	Q269	cag > caa	1	Optimal	Sub-optimal	1.78*	0.22**
	t1009g	T271	act > acg	1	Optimal	Least optimal	1.80	0.19
Experimental isolates	g625a	T141	acg > aca	3	Least optimal	Least optimal	0.19	0.14**
	t652c	A150	gct > gcc	3	Optimal	Least optimal	1.88*	0.23**
	t661c	H153	cac > cat	4	-	-	1.55*	0.45**
	c727t	N175	aac > aat	3	Optimal	Least optimal	1.90*	0.10**

[a] Mutations in bold are present in one of the seven silent variants of TEM-1. All mutations shown are relative to the TEM-1 sequence (SUTCLIFFE 1978).
[b] Codon preference principles for *E. coli* as defined by Ikemura (1981); dash indicates that the criteria are not applicable
[c] Relative Synonymous Codon Usage (RSCU; observed frequency of codon/ expected frequency of codon when all codons for a particular amino acid are used equally) of highly expressed proteins in *E. coli* according to Sharp & Li (1987); dark shade indicates that the change is towards a lower RSCU value, * highest RSCU ** lowest RSCU

Table 5.5: RSCU values of the 15-codon neighborhood around the four silent mutations found with high frequency in the experimental data. For each codon the maximal ('max') and actual ('act') RSCU are shown, in analogy to the supplemental data of Kimchi-Sarfaty *et al.* (2007). The four silent mutations are boxed and the codon before (left) and after (right) the silent mutation is shown.

AA[a]	A	A	N	L	L	L	T	**T141**		I	G	G	P	K	E	L	
Codon[b]	GCG	GCC	AAC	TTA	CTT	CTG	ACA	**ACG**	**ACA**	ATC	GGA	GGA	CCG	AAG	GAG	CTA	
Max	1.88	1.88	1.90	5.33	5.33	5.33	1.87	**1.87**		2.53	2.28	2.28	3.29	1.60	1.59	5.33	
Act[c]	0.80	0.23	1.90	0.11	0.23	5.33	0.14	**0.19**	**0.14**	2.53	0.02	0.02	3.29	0.40	0.41	0.04	
AA[a]	G	G	P	K	E	L	T	**A150**		F	L	H	N	M	G	D	H
Codon[b]	GGA	GGA	CCG	AAG	GAG	CTA	ACC	**GCT**	**GCC**	TTT	TTG	CAC*	AAC	ATG	GGG	GAT*	CAT*
Max	2.28	2.28	3.29	1.60	1.59	5.33	1.87	**1.88**		1.54	5.33	1.55	1.90	X[d]	2.28	1.40	1.55
Act[c]	0.02	0.02	3.29	0.40	0.41	0.04	1.87	**1.88**	**0.23**	0.46	0.11	1.55	1.90	X[d]	0.04	0.61	0.45
AA[a]	K	E	L	T	A	F	L	**H153**		N	M	G	D	H	V	T	R
Codon[b]	AAG	GAG	CTA	ACC	GCT	TTT	TTG	**CAC***	**CAT***	AAC	ATG	GGG	GAT*	CAT*	GTA	ACT	CGC
Max	1.60	1.59	5.33	1.87	1.88	1.54	5.33	**1.55**		1.90	X[d]	2.28	1.40	1.55	2.24	1.87	4.38
Act[c]	0.40	0.41	0.04	1.87	1.88	0.46	0.11	**1.55**	**0.45**	1.90	X[d]	0.04	0.61	0.45	1.11	1.80	1.56
AA[a]	E	L	N	E	A	I	P	**N175**		D	E	R	D	T	T	M	P
Codon[b]	GAG	CTG	AAT	GAA	GCC	ATA	CCA	**AAC**	**AAT**	GAC*	GAG	CGT	GAC*	ACC	ACG	ATG	CCT
Max	1.59	5.33	1.90	1.59	1.88	2.53	3.29	**1.90**		1.40	1.59	4.48	1.40	1.87	1.87	X[d]	3.29
Act[c]	0.41	5.33	0.10	1.59	0.23	0.01	0.44	**1.90**	**0.10**	1.40	0.41	4.48	1.40	1.87	0.19	X[d]	0.23

[a] Amino acid (according to the IUPAC single letter code)
[b] Shade indicates that the codon used is not the most preferred according to Ikemura (1981). Dark shade indicates least preferred codons. Asterix indicates that no preference data are available for the respective codon.
[c] Shade indicates that the codon used is not the most common one according to the RSCU values of Sharp and Li (1987). Dark shade indicates lowest RSCU value.
[d] No RSCU value, since amino acid is encoded by a single codon only.

Since the Ω-loop is of high functional importance for TEM-1 and since the silent mutation found three times in this loop in experiments (N175*) was found to increase MIC levels, we have focussed on the codon usage in this loop and its border-regions in the same way as described above (Table 5.6). It is clear that the codon switch caused by the silent mutation at amino acid position 175 contributes to the formation of a rather large cluster of rare codons located around the C-terminal end of the Ω-loop. When this results in a local decrease in the speed of translation it may influence the folding of the Ω-loop, and in that way affect the MIC for cefotaxime, as was found for the mutant TEM-allele containing only this silent mutation.

Silent variation in signal sequence of natural isolates

Secretory signal sequences of *E. coli* are rich in non-optimal codons and it was hypothesised that this could facilitate targeting of secretory proteins by allowing interaction of the nascent signal peptide with the cell membrane prior to translocation (BURNS and BEACHAM 1985; POWER *et al.* 2004). It has also been shown that the codon composition of the signal sequence of TEM-1 (amino acids 3 to 25) can alter enzyme expression: when 13 non-optimal codons (according to the criteria of Ikemura 1981) in this region were replaced by optimal ones, a fourfold reduction in resistance to ampicillin was measured that was independent of plasmid copy number or mRNA levels (ZALUCKI *et al.* 2008). When this ill-performing mutant was subsequently allowed to revert its first six optimised codons using degenerate primers, the nine isolated colonies with increased resistance levels all contained reversions to non-optimal codons. Stimulated by these findings, we have analyzed the codon changes caused by silent mutations in the signal peptide region of clinical isolates of TEM-1. Again, we used two different methodologies to estimate codon preference in *E. coli* for comparison (IKEMURA 1981; SHARP and LI 1987). We found that the three silent mutations identified in the signal peptide of clinical TEM-isolates change the respective codons to the least optimal ones according to both datasets (Table 5.7 and 5.8).

Interestingly, it seems that non-optimal codons in the N-terminal half of the signal sequence are especially important to maintain ß-lactamase expression (ZALUCKI *et al.* 2008). All three silent mutations found in clinical isolates are in this region and they all change a (sub)optimal variants that only contain silent mutations (TEM-1b, TEM-1e and TEM-1-II) and is ubiquitous among clinically isolated TEM-alleles. We therefore speculate that in *E. coli* such a mutation could have a subtle but yet unnoticed effect on ß-lactamase expression.

Table 5.6: Codon usage in the Ω-loop. The change in codon caused by the silent mutation at amino acid position 175 has been boxed.

Structure[a]					Ω-loop	Ω-loop	Ω-loop	Ω-loop	Ω-loop	Ω-loop	Ω-loop	Ω-loop – H7	Ω-loop – H7	Ω-loop – H7	Ω-loop	Ω-loop	Ω-loop	Ω-loop	Ω-loop		Ω-loop	Ω-loop	Ω-loop	Ω-loop			
Amino acid number	157	158	159	160	161	162	163	164	165	166	167	168	169	170	171	172	173	174	175 - before	175 - after	176	177	178	179	180	181	182
Amino acid[b]	D	H	V	T	R	L	D	R	W	E	P	E	L	N	E	A	I	P	N	N	D	E	R	D	T	T	M
Ikemura[c]	*	*					*		X[e]												*			*			X[e]
Sharp and Li[d]									X[e]																		X[e]

[a] Structure according to Jelsch *et al.* (1993); note that a small α-helix is present between 168-170
[b] According to the IUPAC single letter code
[c] Codon preference according to the principles defined by Ikemura (1981); shade indicates codon is not the most preferred codon, dark shade indicates least preferred codons. Asterix indicates that no preference criteria are available.
[d] Codon preference according to the RSCU values of Sharp and Li (1987); shade indicates that the codon used is not the most common one, dark shade indicates the least common codons.
[e] Amino acid encoded by a single codon.

Table 5.7: Distribution of optimal (white) and non-optimal codons (shaded) in the signal sequence of TEM-1, according to the codon preference table of Ikemura (1981) (darkest shade indicating least preferred codons) and the codon distribution in very highly expressed proteins of *E. coli* as described by Sharp and Li (1987) (darkest shade indicating codons with lowest RSCU values). The codon-changes caused by the three silent mutations that can be found in clinical isolates (in amino acids 8, 9 and 13) are given next to the respective wild-type codon and in each case both have been boxed. Codons where criteria don't apply are indicated with an asterix.

Position	3	4	5	6	7	8		9		10	11	12	13		14	15	16	17	18	19	20	21	22	23	24	25
AA	M	S	I	Q	H	F		R		V	A	L	I		P	F	F	A	A	F	C	L	P	V	F	A
Codon	ATG	AGT	ATT	CAA	CAT	TTC	TTT	CGT	CGA	GTC	GCC	CTT	ATT	ATA	CCC	TTT	TTT	GCG	GCA	TTT	TGC	CTT	CCT	GTT	TTT	GCT
Ikemura	*	*			*																*					
Sharp	*																									

Table 5.8: Details of silent mutations found in the signal peptide region of TEM-1 in clinical and experimental isolates.

	Amino acid	**Silent mutation**	**# isolates**	**Change[a]**		**Change[b]**	
				From:	*To:*	*From:*	*To:*
Clinical isolates	F8	c226t	44	Optimal	Least-optimal	1.54*	0.46**
	R9	t229a	1	Optimal	Least-optimal	4.38*	0.02**
	I13	t241a	1	Sub-optimal	Least-optimal	0.47	0.01**
Experimental isolates	I5	t217a	1	Sub-optimal	Least-optimal	0.47	0.01**
	Q6	g220a	1	Optimal	Least-optimal	1.78*	0.22**
	R9	t229a	2	Optimal	Least-optimal	4.38*	0.02**
	I13	t241c	1	Sub-optimal	Optimal	0.47	2.53*
	A17	g253a	1	Optimal	Optimal	0.80	1.10

[a] According to codon usage preference as defined by Ikemura (1981).
[b] Relative Synonymous Codon Usage (RSCU) values based on the codon usage in very highly expressed proteins of *E. coli* according to Sharp and Li (1987); * = highest RSCU-value and ** = lowest RSCU-values.

Furthermore, the silent mutation in the amino acid at position nine (which changes the most preferred and most common codon 'CGT' for arginine to the least preferred and least common codon 'CGA') was not only found in a clinical isolate, but in two experimental isolates as well (Table 5.3). Although it is not possible do distinguish this event from mere chance (see our statistical arguments given above), both the location and the nature of this silent mutation are at least suspicious.

Concluding, we provide both direct and indirect empirical evidence for functional effects of silent mutations in a common antibiotic resistance gene. We show that one of the silent mutations found with high frequency in an experimental dataset increases resistance to a new antibiotic relative to the wild type enzyme. Another silent mutation was found with such a high frequency that the chance that it resulted from genetic drift was estimated to be extremely low under the assumption of a Binomial distribution of chances. Furthermore, we have reasons to believe that silent variation present in the signal sequence of clinical isolates of TEM-alleles is not selectively neutral. Because of this evidence, we think it is important that both nonsynonymous and synonymous mutations are reported when sequence data on antibiotic resistance genes are reported in order to obtain a better overview of putative functional effects of silent mutations. Since antibiotic resistance genes often strongly affect

the fitness of their host and are subject to lateral transfer (OCHMAN *et al.* 2000), which can alter the concentration of tRNAs in their environment, they might provide an excellent system for the detection and study of functional effects of synonymous mutations.

References

AMBLER, R. P., A. F. W. COULSON, J. M. FRERE, J. M. GHUYSEN, B. JORIS *et al.* 1991. A Standard Numbering Scheme for the Class-a Beta-Lactamases. Biochem. J. **276:** 269-270.

ANFINSEN, C. B. 1973. Principles That Govern Folding of Protein Chains. Science **181:** 223-230.

BARLOW, M., and B. G. HALL. 2002. Predicting evolutionary potential: in vitro evolution accurately reproduces natural evolution of the tem beta-lactamase. Genetics **160:** 823-832.

BARLOW, M., and B. G. HALL. 2003. Experimental prediction of the natural evolution of antibiotic resistance. Genetics **163:** 1237-1241.

BURNS, D. M., and I. R. BEACHAM. 1985. Rare Codons in Escherichia-Coli and S-Typhimurium Signal Sequences. FEBS Lett. **189:** 318-323.

DATTA, N., and P. KONTOMICHALOU. 1965. Penicillinase Synthesis Controlled by Infectious R Factors in Enterobacteriaceae. Nature **208:** 239-241.

DEANA, A., R. EHRLICH and C. REISS. 1996. Synonymous codon selection controls in vivo turnover and amount of mRNA in Escherichia coli bla and ompA genes. J. Bacteriol. **178:** 2718-2720.

HOFACKER, I. L., W. FONTANA, P. F. STADLER, L. S. BONHOEFFER, M. TACKER *et al.* 1994. Fast Folding and Comparison of Rna Secondary Structures. Monatshefte Fur Chemie **125:** 167-188.

IKEMURA, T. 1981. Correlation between the Abundance of Escherichia-Coli Transfer-Rnas and the Occurrence of the Respective Codons in Its Protein Genes. J. Mol. Biol. **146:** 1-21.

JELSCH, C., L. MOUREY, J. M. MASSON and J. P. SAMAMA. 1993. Crystal-Structure of Escherichia-Coli Tem1 Beta-Lactamase at 1.8-Angstrom Resolution. Proteins-Structure Function and Genetics **16:** 364-383.

KIMCHI-SARFATY, C., J. M. OH, I. W. KIM, Z. E. SAUNA, A. M. CALCAGNO *et al.* 2007. A "silent" polymorphism in the MDR1 gene changes substrate specificity. Science **315:** 525-528.

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.

KOMAR, A. A., and R. JAENICKE. 1995. Kinetics of Translation of Gamma-B Crystallin and Its Circularly Permutated Variant in an in-Vitro Cell-Free System - Possible Relations to Codon Distribution and Protein-Folding. FEBS Lett. **376:** 195-198.

KOMAR, A. A., T. LESNIK and C. REISS. 1999. Synonymous codon substitutions affect ribosome traffic and protein folding during in vitro translation. FEBS Lett. **462:** 387-391.

MATAGNE, A., J. LAMOTTE-BRASSEUR and J. M. FRERE. 1998. Catalytic properties of class A beta-lactamases: efficiency and diversity (vol 330, pg 581, 1998). Biochem. J. **331:** 975-975.

MATHEWS, D. H., J. SABINA, M. ZUKER and D. H. TURNER. 1999. Expanded sequence dependence of thermodynamic parameters improves prediction of RNA secondary structure. J. Mol. Biol. **288:** 911-940.

NACKLEY, A. G., S. A. SHABALINA, I. E. TCHIVILEVA, K. SATTERFIELD, O. KORCHYNSKYI *et al.* 2006. Human catechol-O-methyltransferase haplotypes modulate protein expression by altering mRNA secondary structure. Science **314:** 1930-1933.

OCHMAN, H., J. G. LAWRENCE and E. A. GROISMAN. 2000. Lateral gene transfer and the nature of bacterial innovation. Nature **405:** 299-304.

PALZKILL, T., and D. BOTSTEIN. 1992. Probing Beta-Lactamase Structure and Function Using Random Replacement Mutagenesis. Proteins-Structure Function and Genetics **14:** 29-44.

PARMLEY, J. L., and L. D. HURST. 2007. How do synonymous mutations affect fitness? BioEssays **29:** 515-519.

POWER, P. M., R. A. JONES, I. R. BEACHAM, C. BUCHOLTZ and M. P. JENNINGS. 2004. Whole genome analysis reveals a high incidence of non-optimal codons in secretory signal sequences of Escherichia coli. Biochem. Biophys. Res. Commun. **322:** 1038-1044.

PURVIS, I. J., A. J. E. BETTANY, T. C. SANTIAGO, J. R. COGGINS, K. DUNCAN *et al.* 1987. The Efficiency of Folding of Some Proteins Is Increased by Controlled Rates of Translation Invivo - a Hypothesis. J. Mol. Biol. **193:** 413-417.

SALVERDA, M., and J. A. G. M. DE VISSER. unpublished.

SHARP, P. M., and W. H. LI. 1986. Codon Usage in Regulatory Genes in Escherichia-Coli Does Not Reflect Selection for Rare Codons. Nucleic Acids Res. **14:** 7737-7749.

SHARP, P. M., and W. H. LI. 1987. The Codon Adaptation Index - a Measure of Directional Synonymous Codon Usage Bias, and Its Potential Applications. Nucleic Acids Res. **15:** 1281-1295.

SUTCLIFFE, J. G. 1978. Nucleotide-Sequence of Ampicillin Resistance Gene of Escherichia-Coli Plasmid Pbr322. Proc. Natl. Acad. Sci. U. S. A. **75:** 3737-3741.

VAKULENKO, S. B., B. GERYK, L. P. KOTRA, S. MOBASHERY and S. A. LERNER. 1998. Selection and characterization of beta-lactam-beta-lactamase inactivator-resistant mutants following PCR mutagenesis of the TEM-1 beta-lactamase gene. Antimicrob. Agents Chemother. **42:** 1542-1548.

ZALUCKI, Y. M., K. L. GITTINS and M. P. JENNINGS. 2008. Secretory signal sequence non-optimal codons are required for expression and export of beta-lactamase. Biochem. Biophys. Res. Commun. **366:** 135-141.

ZUKER, M. 2003. Mfold web server for nucleic acid folding and hybridization prediction. Nucleic Acids Res. **31:** 3406-3415.

-Chapter 6-

General discussion

Merijn Salverda

The enzyme studied in this thesis, TEM-1 ß-lactamase, confers resistance to ß-lactam antibiotics and is one of the best studied antibiotic resistance enzymes around. The remarkable phenotypic plasticity of TEM-1, that gradually became apparent as a result of the introduction of more modern ß-lactam antibiotics from the beginning of the 1980s onwards, makes it an ideal enzyme for the study of molecular evolution. Consequently, TEM-1 has been used as a model system for the study of protein structure-function relationships, directed evolution protocols and the prediction of antibiotic resistance. Elaborating on the detailed knowledge of both the natural and laboratory evolution of TEM-1, the molecular evolution of this notorious antibiotic resistance gene was studied in this thesis. This has resulted in a more detailed knowledge about the mechanisms shaping the evolution of TEM-1, such as recombination and the nature and interaction of adaptive mutations.

The use of TEM-1 as a model system for the study of epistasis and the fitness landscape

In ***Chapter 2***, the fitness landscape of TEM-1 was studied using directed evolution. The evolutionary potential to adapt to increasing concentrations of cefotaxime, an antibiotic to which TEM-1 confers only marginal resistance, was studied using *in vitro* evolution. Deviations from the predominant mutational pathway observed in an initial experiment led to the identification of two mutations (causing substitutions R164S and A237T) that had putative sign epistatic interactions with the initial mutation of the predominant mutational pathway to high resistance (G238S). Directed evolution starting with both single mutants indeed indicates that they give rise to one or possibly more alternative mutational pathways. Evidence that such interactions do not only play a role at the level of the first mutation in such pathways was also found.

It has been shown that, while sign epistasis can create ridges and valleys in the pathways that together constitute a fitness landscape (WEINREICH *et al.* 2005), a special form of sign epistasis named ‘reciprocal sign epistasis’ (where the effect of a mutation at one locus

depends on what mutation is present at another locus and vice versa) is a necessary condition for the occurrence of multiple peaks isolated by neighbouring genotypes of lower fitness (POELWIJK 2008). Such reciprocal sign epistasis is found between substitutions R164S and G238S: the double mutant carrying both substitutions has a MIC that is considerably lower than that of both single mutants and barely higher than that of TEM-1. In this respect, it would be interesting to see what solutions adaptive evolution 'finds' when the R164S/ G238S double mutant is subjected to *in vitro* evolution. It is already known that adaptive evolution in the background of this double mutant is possible by addition of substitution E104K or M182T (ZACCOLO and GHERARDI 1999), but it is unknown where this evolutionary trajectory goes and whether at some point back mutation S164R or S238G is the only way 'up'. The same questions apply to the evolutionary fate of the A237T/ G238S double mutant.

The comparison of natural and laboratory evolution

Considering laboratory evolution *in vivo* (experimental evolution) and *in vitro* (directed evolution), TEM-1 ß-lactamase is by far the best studied antibiotic resistance enzyme around. In ***Chapter 3***, the amino acid substitutions identified in 17 laboratory evolution studies on TEM-1 mediated antibiotic resistance (the majority focussing at the evolutionary potential for increased resistance to a variety of antibiotics that TEM-1 itself is susceptible to) were compared to those found in a dataset containing all known clinically isolated TEM-variants identified so far. This proved that laboratory evolution is highly successful and accurate in identifying the variation found in clinical TEM-alleles. Additionally, side-by-side comparison of laboratory studies resulted in the identification of amino acid substitutions that seem to constitute a selective benefit but that have nevertheless not (yet) been reported in clinical isolates. Time will tell whether these substitutions are part of sequence space not yet tapped by natural selection or whether they are laboratory artefacts that are not likely to appear in a clinical setting.

The big question is of course, even when the evolution of antibiotic resistance genes is largely predictable in the laboratory, to what extent predictive laboratory evolution can be exploited in the design and use of novel antibiotics. For one thing, *a priori* knowledge about the evolutionary potential of antibiotic resistance genes and detailed characterization of the pleiotropic effects of the different mutations involved may result in knowledge about which drug combinations to apply when a predicted mutant is identified in nature.

The effect of recombination and other variables influencing evolutionary dynamics *in vivo*

In ***Chapter 4***, several lines of evidence are presented indicating that recombination has had an important effect on the sequence evolution of TEM-alleles. In order to address the question whether recombination has adaptive significance from a theoretical point of view, both mutation rate and population size also have to be taken into account, because recombination can only be effective in populations that are sufficiently polymorphic.

There are reasons to believe that the *in vivo* mutation rate of TEM-alleles is higher than that of most bacterial strains. The frequency of hypermutable strains of *E. coli* is higher among clinical strains carrying TEM-alleles than for clinical strains not carrying TEM-alleles (BAQUERO *et al.* 2005). Since a association between hypermutable strains and acquisition of antibiotic resistance has been demonstrated (OLIVER *et al.* 2000), these results suggest that TEM-alleles might be subject to increased mutation rates *in vivo*. This could be connected to the finding that ß-lactam mediated inhibition of cell wall synthesis can induce an error prone DNA polymerase in *E. coli*, resulting in an increased mutation rate and a mutator phenotype (Perez-Capilla *et al.* 2005; Tompkins *et al.* 2003). The chemicals used to fight bacteria might thus enhance the rate of antibiotic resistance acquisition.

The mutation rate in *E. coli* per base pair per replication is ~5.4 x 10^{-10} (DRAKE *et al.* 1998). This means that per round of replication about one in five million *E. coli* cells harbours a TEM-allele with a point mutation (because TEM-alleles are usually plasmidic, the mutation rate is increased by a factor 10 assuming an average copy number of 10), which is increased to one in five hundred thousand if we assume the alleles reside in weak mutator strains (Baquero *et al.* 2005; Baquero *et al.* 2004). The exact population size of TEM-carrying bacteria in humans is hard to assess. TEM-alleles are most commonly found in bacteria isolated from urinary tract infections. Such infections are diagnosed when more than 100,000 bacterial cells are present per millilitre of urine (WRIGHT *et al.* 1986). This indicates that population sizes in such infections are at least in the order of 10^7 cells and probably much higher.

For recombination to have any effect at all there needs to be genetic polymorphism in a population. The 'natural' environment of TEM-alleles seems to promote this; within the highly compartmentalized human body, a high diversity of antibiotic concentration gradients exists (BAQUERO and NEGRI 1997), which have been shown to promote genetic

polymorphism among TEM-alleles (BAQUERO *et al.* 1997). In agreement with this, multiple TEM-alleles are sometimes identified in a single patient (BRADFORD *et al.* 1994; ESSACK *et al.* 2001; KIM *et al.* 2004; LAVIGNE *et al.* 2004). Inter- or intraplasmidic recombination can act upon such genetic polymorphism and increase the genetic variation upon which natural selection can act.

An interesting experimental approach to address the importance of recombination would be to study the adaptive evolution of TEM-1 mediated resistance to a novel antibiotic in normal *E. coli* cells and in *E. coli* suppressed in their recombination machinery (*rec*$^-$ mutants). Alternatively, *in vivo* evolution of normal *E. coli* cells carrying a TEM-allele with a single beneficial mutation like G238S could be studied. Cells carrying another single mutant with a silent marker and with a positive effect on resistance in the background of G238S (e.g. M182T, created by the introduction of a double point mutation) can be continuously added to the G238S culture that is evolving *in vivo*. Whether the G238S culture evolves by point mutation or by recombination can now be distinguished (note that transferable plasmids are required).

Evidence for selectable yet silent mutations

In ***Chapter 5***, evidence is presented indicating that a silent mutation can confer an adaptive benefit for cefotaxime resistance. Putative mechanisms behind this are discussed and silent mutations present in clinical isolates are also analyzed, suggesting that not all of these mutations need to be selectively neutral (as is often assumed).

Under strong antibiotic pressure, the functioning of antibiotic resistance enzymes is directly correlated to the fitness of the bacteria they reside in. Thus, the *in vitro* evolution of antibiotic resistance genes may provide a valuable tool for identifying and analyzing silent mutations with adaptive effects. It has been noted that compared to the codon usage of *E. coli*, TEM-1 contains a lot of rare codons. It would be interesting to see whether prolonged *in vitro* evolution and selection for improvement of ampicillin resistance would result in any translational selection. Frequent lateral transfer between different bacterial species might prevent such selection in nature.

Some other directions for future research

Selection *in vivo*

One of the problems of using *in vitro* evolution protocols is the difficulty of simulating a more or less natural mutational process where mutations appear and go to fixation one at a time. Although the mutation rate in an error-prone PCR can be accurately controlled, it is impossible to achieve a mutation rate that is so low that double mutants are excluded. An alternative is to keep the mutational process completely natural, by evolution of the bacteria carrying a plasmid with the antibiotic resistance gene to be studied *in vivo*. Such studies have been undertaken and they show that simple daily transfers of bacteria growing in increasing amounts of antibiotic yields mutant ß-lactamases within weeks (BLAZQUEZ *et al.* 2000; HOLLOWAY *et al.* 2007). The evolutionary trajectories described in ***Chapter 2***, as well as those described by others (WEINREICH *et al.* 2006) await testing in such a semi-natural *in vivo* system.

Another advantage of *in vivo* evolution is the possibility to follow the evolution of antibiotic resistance not only at the level of the plasmidic antibiotic resistance gene, but also at the level of evolution of the bacterial chromosome. When the re-transformation step is left out, bacterial genes are allowed to evolve along with the gene on the plasmid. Several other bacterial genes are known to evolve in response to antibiotic pressure, e.g. those coding for the outer membrane proteins and transpeptidases (FRERE *et al.* 1991; PERILLI *et al.* 2007). While the simultaneous monitoring of the molecular changes at multiple genes may be experimentally challenging, it also offers the opportunity to study the interactions between mutations at these different genes. A similar and easier approach can be achieved by omitting the replacement of the gene in a naïve vector background, which would allow evolution of the plasmid as a whole. Although it may seem intuitive that mutations at the regulatory level (like promoter mutations or mutations in genes regulating the promoter) result in additive fitness effects, increase in protein production can increase the aggregation propensity and thus the fitness effect of such regulatory mutations may rely on other mutations that can suppress this propensity (WEINREICH *et al.* 2006).

Fitness-assays

Since the fitness of a TEM-variant is directly correlated with the maximum concentration of a

specific ß-lactam antibiotic that bacteria carrying this variant can tolerate, MIC-assays have traditionally been used for assessing the fitness of TEM-alleles. A disadvantage of this method is the so-called 'inoculum effect'; killing of bacteria by lysis result in the release of additional ß-lactamases in the test medium and additional hydrolysis of the drug. This effect has typically been found with ß-lactam antibiotics against ß-lactamase producing bacterial strains (CRAIG *et al.* 2004). In practice, even when using exactly the same inoculum, slight variations in MIC can be recorded the next day (WEINREICH *et al.* 2006). This can be a problem when TEM-alleles that confer subtle differences in resistance level are compared. An alternative solution is the use of bacterial strains with neutral phenotypic markers. The fitness of different TEM-alleles can be assessed by direct competition between different strains carrying different alleles (NEGRI *et al.* 2000). Easier and more reliable methods relying on fluorescence based marker systems are available and could be of great use for more accurate fitness measurements of TEM-alleles.

Inaccessible sequence space

A better understanding of the molecular basis of adaptation is essential for the development of evolutionary biology as a predictive science for the study of protein evolution or emergent pathogens (COUNAGO *et al.* 2006). Using the approach of directed evolution, both areas were studied in this thesis. It seems that the adaptive solutions found by clinical TEM-isolates derive from similar areas in sequence space as the solutions uncovered by directed evolution studies, even though some of these studies have employed extremely high rates of recombination (STEMMER 1994) and mutation (ZACCOLO and GHERARDI 1999). Interestingly, some mutant TEM-variants partially constructed by computational design and optimized using directed evolution seem to derived from different areas of sequence space, probably because they contain a number of substitutions caused by double or triple nucleotide changes. It would be interesting to analyze the fitness (or its 'proxy' MIC) of these mutants mutation by mutation to see whether there truly is variation that is 'out of reach' for adaptive evolution.

References

Baquero, F., and M. C. Negri, 1997 Selective compartments for resistant microorganisms in antibiotic gradients. Bioessays **19:** 731-736.

Baquero, F., M. C. Negri, M. I. Morosini and J. Blazquez, 1997 Antibiotic selective process: Concentration-specific amplification of low-level resistant populations. Antibiotic Resistance: Origins, Evolution, Selection and Spread **207:** 93-105.

Baquero, M. R., J. C. Galan, M. D. Turrientes, R. Canton, T. M. Coque *et al.*, 2005 Increased mutation frequencies in Escherichia coli isolates harboring extended-spectrum beta-lactamases. Antimicrobial Agents and Chemotherapy **49:** 4754-4756.

Baquero, M. R., A. I. Nilsson, M. D. C. Turrientes, D. Sandvang, J. C. Galan *et al.*, 2004 Polymorphic mutation frequencies in Escherichia coli: Emergence of weak mutators in clinical isolates. Journal of Bacteriology **186:** 5538-5542.

Blazquez, J., M. I. Morosini, M. C. Negri and F. Baquero, 2000 Selection of naturally occurring extended-spectrum TEM beta-lactamase variants by fluctuating beta-lactam pressure. Antimicrobial Agents and Chemotherapy **44:** 2182-2184.

Bradford, P. A., C. E. Cherubin, V. Idemyor, B. A. Rasmussen and K. Bush, 1994 Multiply Resistant Klebsiella-Pneumoniae Strains from 2 Chicago Hospitals - Identification of the Extended-Spectrum Tem-12 and Tem-10 Ceftazidime-Hydrolyzing Beta-Lactamases in a Single Isolate. Antimicrobial Agents and Chemotherapy **38:** 761-766.

Counago, R., S. Chen and Y. Shamoo, 2006 In vivo molecular evolution reveals biophysical origins of organismal fitness. Molecular Cell **22:** 441-449.

Craig, W. A., S. M. Bhavnani and P. G. Ambrose, 2004 The inoculum effect: Fact or artifact? Diagnostic Microbiology and Infectious Disease **50:** 229-230.

Drake, J. W., B. Charlesworth, D. Charlesworth and J. F. Crow, 1998 Rates of spontaneous mutation. Genetics **148:** 1667-1686.

Essack, S. Y., L. M. C. Hall, D. G. Pillay, M. L. McFadyen and D. M. Livermore, 2001 Complexity and diversity of Klebsiella pneumoniae strains with extended-spectrum beta-lactamases isolated in 1994 and 1996 at a teaching hospital in Durban, South Africa. Antimicrobial Agents and Chemotherapy **45:** 88-95.

Frere, J. M., B. Joris, B. Granier, A. Matagne, F. Jacob *et al.*, 1991 Diversity of the Mechanisms of Resistance to Beta-Lactam Antibiotics. Research in Microbiology **142:** 705-710.

Holloway, A. K., T. Palzkill and J. J. Bull, 2007 Experimental evolution of gene duplicates in a bacterial plasmid model. Journal of Molecular Evolution **64:** 215-222.

Kim, S., J. Kim, Y. H. Kang, Y. Park and B. Lee, 2004 Occurrence of extended-spectrum genus Shigella in the beta-lactamases in members of the Republic of Korea. Journal of Clinical Microbiology **42:** 5264-5269.

Lavigne, J. P., N. Bouziges, C. Chanal, A. Mahamat, S. Michaux-Charachon *et al.*, 2004 Molecular epidemiology of Enterobacteriaceae isolates producing extended-spectrum beta-lactamases in a French hospital. Journal of Clinical Microbiology **42:** 3805-3808.

Negri, M. C., M. Lipsitch, J. Blazquez, B. R. Levin and F. Baquero, 2000 Concentration-dependent selection of small phenotypic differences in TEM beta-lactamase-mediated antibiotic resistance. Antimicrobial Agents and Chemotherapy **44:** 2485-2491.

Oliver, A., R. Canton, P. Campo, F. Baquero and J. Blazquez, 2000 High frequency of hypermutable Pseudomonas aeruginosa in cystic fibrosis lung infection. Science **288:** 1251-1253.

Perez-Capilla, T., M. R. Baquero, J. M. Gomez-Gomez, A. Ionel, S. Martin *et al.*, 2005 SOS-independent induction of dinB transcription by beta-lactam-mediated inhibition of cell wall synthesis in Escherichia coli. Journal of Bacteriology **187:** 1515-1518.

Perilli, M., B. Segatore, M. Tavio, D. Setacci, G. Celenza *et al.*, 2007 In vitro selection and characterization of mutants in TEM-1-producing Escherichia coli by ceftazidime and ceftibuten. Journal of Chemotherapy **19:** 123-126.

Poelwijk, F. J., 2008 Fitness landscapes of gene regulation in variable environments, pp. 184. Ph.D. thesis, FOM Institute for Atomic and Molecular Physics, Amsterdam.

Stemmer, W. P., 1994 Rapid evolution of a protein in vitro by DNA shuffling. Nature **370:** 389-391.

Tompkins, J. D., J. L. Nelson, J. C. Hazel, S. L. Leugers, J. D. Stumpf *et al.*, 2003 Error-prone polymerase, DNA polymerase IV, is responsible for transient hypermutation during adaptive mutation in Escherichia coli. Journal of Bacteriology **185:** 3469-3472.

Weinreich, D. M., N. F. Delaney, M. A. DePristo and D. L. Hartl, 2006 Darwinian evolution can follow only very few mutational paths to fitter proteins. Science **312:** 111-114.

Weinreich, D. M., R. A. Watson and L. Chao, 2005 Perspective: Sign epistasis and genetic constraint on evolutionary trajectories. Evolution **59:** 1165-1174.

Wright, D. N., B. Saxon and J. M. Matsen, 1986 Use of the Bac-T-Screen to Predict Bacteriuria from Urine Specimens Held at Room-Temperature. Journal of Clinical Microbiology **24:** 214-217.

Zaccolo, M., and E. Gherardi, 1999 The effect of high-frequency random mutagenesis on in vitro protein evolution: A study on TEM-1 beta-lactamase. Journal of Molecular Biology **285:** 775-783.

Samenvatting

Merijn Salverda

Antibiotica en hun ontdekking

Antibiotica zijn stoffen die gebruikt worden voor de bestrijding van ziekteverwekkende micro-organismen. Het eerste wereldwijd gebruikte antibioticum, penicilline, wordt geproduceerd door bepaalde schimmels. Penicilline speelt een rol in de strijd om voedingsstoffen die er woedt in de microbiële wereld. Een schimmel die penicilline produceert zorgt er namelijk voor dat er in zijn directe omgeving geen bacteriën kunnen groeien. Op die manier worden concurrerende schimmels en andere micro-organismen op een afstand gehouden.

Penicilline werd in 1928 per toeval ontdekt door de Britse microbioloog Alexander Fleming. Na terugkeer van zijn zomervakantie ontdekte hij dat veel van zijn achtergelaten bacterieplaten met daarop de bacterie *Staphylococcus aureus* besmet waren door een schimmel. Dit was niet zo verwonderlijk, aangezien Fleming bekend stond als een sloddervos en er door de aanwezigheid van een schimmellab een verdieping onder het lab van Fleming heel wat schimmelsporen door de lucht gezweefd moeten hebben. In eerste instantie zette Fleming zijn platen op een stapel in een bak met een ontsmettingsmiddel om de boel te desinfecteren. Toen hij die dag een plaat wilde laten zien aan een voormalig collega, greep hij een beschimmelde plaat die bovenop de stapel lag en zodoende ontsnapt was aan het ontsmettingsmiddel. Pas toen viel hem op dat er rondom de schimmel een gebied te zien was waarin geen bacteriën groeiden. Fleming was gefascineerd door deze waarneming. Met veel moeite slaagde hij er uiteindelijk in de stof te isoleren die door deze schimmel, genaamd *Penicillium notatum*, geproduceerd werd. Het zou nog tot aan het begin van de Tweede Wereldoorlog duren voordat penicilline efficiënt geproduceerd en gebruikt kon worden.

Antibioticumresistentie

Nog voordat penicilline haar klinische debuut maakte, werd door enkele Britse wetenschappers in een laboratorium ontdekt dat bacteriën al bij kortstondige blootstelling aan penicilline tekenen van resistentie vertonen. Helaas bleek deze waarneming ook ‘in het veld’ te gelden. Binnen enkele jaren na de introductie van penicilline doken de eerste berichten over resistente bacteriestammen op. Tegen het eind van de jaren veertig had penicilline veel van

zijn aanvankelijke kracht verloren en was men druk doende met de zoektocht naar betere antibiotica. Deze zoektocht duurt tot vandaag de dag voort, aangezien tot op heden steevast binnen enkele jaren na de introductie van een nieuw antibioticum bacteriële resistentie wordt ontdekt.

Met het probleem van antibioticumresistentie werd ook langzaamaan duidelijk dat bacteriën een ontzagwekkend aanpassingsvermogen hebben. Niet alleen bleken bacteriën in staat resistentie te ontwikkelen tegen elk nieuw antibioticum dat op de markt werd gebracht, ze bleken dit ook nog eens op verschillende manieren te doen. Zo kunnen bacteriën hun membranen minder doorlaatbaar maken voor antibiotica, actief antibioticum uit een cel pompen, de doelwitten waar antibiotica op aangrijpen aanpassen en zelfs antibiotica afbreken met behulp van enzymen.

Gelukkig heeft ook wat antibioticumresistentie betreft elk nadeel zijn voordeel. Zo is de nauwgezette studie van de interactie tussen antibiotica en enzymen die deze antibiotica afbreken bijzonder waardevol gebleken voor het beter begrijpen van de interactie tussen enzym en substraat, en voor de ontwikkeling van de enzymtechnologie in het algemeen. Ook worden antibioticumresistentie genen in de biotechnologie gebruikt om DNA te merken dat in bacteriecellen ingebracht wordt; cellen die erin slagen niet eigen DNA op te nemen, nemen ook het antibioticumresistentie gen op en zijn op basis van hun nieuw verworven resistentie gemakkelijk te scheiden van onveranderde cellen.

ß-lactam antibiotica en de resistentie hiertegen

In 1961 werd ampicilline op de markt gebracht. Ampicilline is een semi-synthetisch antibioticum dat gebaseerd is op penicilline en dat behoort tot de groep van de β-lactam antibiotica. Deze antibiotica verstoren de opbouw van de bacteriële celwand, waardoor de bacteriecellen uiteindelijk uit elkaar klappen. Net zoals bij alle voorgaande antibiotica volgde de eerste melding van ampicilline resistentie binnen enkele jaren en wel in 1963, toen in een Atheens ziekenhuis uit een patiënt ampicilline resistente bacteriën werden geïsoleerd. Later bleek dat deze resistentie veroorzaakt werd door de productie van een enzym dat ampicilline af kan breken. Dit enzym kennen we nu onder de naam TEM-1 β-lactamase . Het eerste deel van deze naam is afgeleid van de achternaam van de eerdergenoemde patiënt, Temoneira, terwijl het tweede deel aangeeft dat het hier gaat om een enzym dat ß-lactams kan afbreken.

Aan het begin van de jaren tachtig werd een serie nieuwe β-lactam antibiotica

geïntroduceerd. Ook in dit geval werden er binnen enkele jaren resistente bacteriën gevonden. Met behulp van experimenten in laboratoria was rond die tijd ook al vastgesteld dat kleine veranderingen in het DNA coderend voor TEM-1 β-lactamase de activiteit van dit enzym kunnen beïnvloeden. Zulke veranderingen kunnen spontaan ontstaan doordat het kopieermechanisme waarmee DNA vermenigvuldigd wordt niet foutloos is. Wanneer er bij het kopiëren van het DNA een mutatie optreedt (d.w.z., een foutje wordt gemaakt), kan dat ertoe leiden dat er in de keten aminozuren waaruit een enzym bestaat een bepaald aminozuur vervangen wordt door een ander aminozuur. Enzymen zijn chemische fabriekjes die één bepaalde stof (het substraat) afbreken of bewerken. Enzym en substraat passen op elkaar volgens een soort sleutel-slot principe en een kleine veranderingen in de structuur van een enzym kan er al toe leiden dat de substraat-voorkeur van een enzym verandert.

Al gauw bleek dat de problemen met resistentie tegen de nieuwe β-lactam antibiotica door precies zo'n mechanisme veroorzaakt werden. Mutaties in het TEM-1 gen zorgden voor structurele veranderingen in het ß-lactmase enzym en in zeldzame gevallen zorgde dit ervoor dat het enzym op zo'n manier veranderde dat het de nieuwe antibiotica beter kon binden. Ondertussen zijn er meer dan 150 van deze gemuteerde varianten van TEM-1 beschreven. De meeste van deze varianten kunnen de antibiotica die van penicilline en ampicilline zijn afgeleid veel beter afbreken dan TEM-1.

Proeven en resultaten beschreven in dit proefschrift

In dit proefschrift is onderzocht hoe TEM-1 zich aanpast aan een nieuw antibioticum. Die vraag is niet alleen interessant voor de antibioticumresistentie problematiek, maar ook voor de evolutiebiologie. Door stap voor stap, of in dit geval mutatie voor mutatie, te bekijken hoe een enzym beter wordt in het afbreken van een antibioticum kan namelijk belangrijke informatie vergaard worden over de manier waarop mutaties elkaar beïnvloeden. Het effect van mutaties kan simpelweg optelbaar zijn, maar ze kunnen elkaars effect ook versterken of afzwakken. Door de dynamiek van mutatie-interactie te onderzoeken kunnen we meer te weten komen over het zogenaamde fitnesslandschap van een enzym. Zo'n fitnesslandschap kun je voor je zien als een plaatje van een berglandschap, waarbij op het platte vlak (d.w.z., langs de x- en de y-as) alle mogelijke genotypen (genetische varianten) van een organisme zijn uitgezet, terwijl in de hoogte (langs de z-as) de 'fitness' van ieder genotype aangegeven is. Deze 'fitness' is een ingewikkelde term, die uitdrukt hoe goed een organisme, of in dit geval een

genotype, in staat is zich te reproduceren ten opzichte van andere organismen van dezelfde soort die in dezelfde omgeving leven. In de dalen vind je dus de genotypen die slecht aangepast zijn aan hun eigen omgeving terwijl de pieken genotypen voorstellen die juist zeer goed aan hun omgeving zijn aangepast. Een genotype dat zich in een dal bevindt kan via mutaties (die ieder op zich voordelig moeten zijn) uit dit dal klimmen en zich 'bergopwaarts' begeven. Telkens wanneer er zich een mutatie voordoet die een positief effect heeft op de fitness, krijgt het nieuwe genotype dat ontstaat een reproductief voordeel ten opzichte van het onveranderde, oude genotype. Het nieuwe genotype zal in een populatie dus het oude verdringen totdat er alleen nog nieuwe genotypen over zijn.

Één van de grote vragen die er bestaan over fitnesslandschap is of je zo'n landschap moet zien als een enkele, eenzame top, zeg de Mont Ventoux, of als een verzameling van pieken en dalen, zeg de Alpen. Welke van deze twee scenario's geldt, heeft directe invloed op een breed scala aan evolutionaire problemen als van de evolutie van seks, soortvorming en de vraag in welke mate evolutie beperkt wordt en herhaalbaar is.

In **Hoofdstuk 2** hebben we zo nauwkeurig mogelijk in kaart gebracht hoe TEM-1 zich stap voor stap aanpast aan een nieuw antibioticum (cefotaxime) waar het aanvankelijk nauwelijks resistentie tegen verleent. Dit is gedaan door natuurlijke evolutie te vervangen door evolutie in het laboratorium. Natuurlijke evolutie is afhankelijk van het ontstaan van toevallige foutjes in het DNA. Door het TEM-1 gen te vermenigvuldigen in een kopieermachine voor DNA en hierbij als het ware 'slechte inkt' te gebruiken, kunnen we het evolutieproces in een laboratorium nabootsen. Dit resulteert in miljoenen kopieën van het TEM-1 gen waarin allerhande 'foutjes' zitten. Door deze mutante kopieën in te bouwen in bacteriën en deze bacteriën vervolgens bloot te stellen aan steeds hogere concentraties van het antibioticum cefotaxime blijven door natuurlijke selectie juist bacteriën met die zeldzame TEM-1 varianten over die toevalligerwijs beter zijn geworden in het afbreken van dit 'nieuwe' antibioticum. Door deze nieuwe varianten te isoleren en het kopieerproces vervolgens te herhalen, kan de efficiëntie waarmee TEM-1 cefotaxime afbreekt door slechts drie of vier mutaties meer dan honderdduizend-voudig verhoogd worden. De proeven in hoofdstuk twee tonen aan dat toevalligheid in dit proces een grote rol speelt. Sommige mutaties werken elkaar namelijk tegen; wanneer ze los van elkaar voorkomen, verhogen ze ieder apart de resistentie tegen cefotaxime, maar wanneer ze gecombineerd worden, doen ze elkaars effect plotseling teniet.

Zulke effecten kunnen zorgen voor een fitnesslandschap met meerdere pieken, en kunnen beperkingen opleggen aan het adaptieve vermogen.

In **Hoofdstuk 3** wordt de natuurlijke evolutie van TEM-1 vergeleken met de laboratorium evolutie van TEM-1 om te onderzoeken in hoeverre natuurlijke evolutie voorspeld kan worden met behulp van laboratorium evolutie. Uit deze vergelijking blijkt dat bijna alle natuurlijke variatie die er tot nu toe binnen TEM-genen gevonden en beschreven is ook gevonden wordt in laboratorium evolutie. Op basis van de verzamelde gegevens worden bovendien enkele voorspellingen gedaan over mutaties die tot op heden niet in natuurlijke TEM-isolaten gevonden zijn, maar die in de toekomst op zouden kunnen duiken.

Hoewel bacteriën normaal gesproken hun erfelijk materiaal niet recombineren, tonen we in **Hoofdstuk 4** aan dat er waar het TEM-genen betreft verschillende bewijzen te vinden zijn die aangeven dat deze genen wel degelijk kunnen recombineren. TEM-genen worden met name gevonden op plasmides. Dat zijn, korte, cirkelvormige stukken DNA die makkelijk uitgewisseld kunnen worden tussen bacteriën. Recombinatie tussen twee TEM-genen binnen één plasmide kan ervoor zorgen dat er een enkel nieuw gen ontstaat. Dit proces zou kunnen resulteren in nieuwe genen die nieuwe combinaties van mutaties met zich meedragen. Er bestaat echter bewijs dat mutante TEM-genen minder goed zijn dan TEM-1 in het afbreken van penicilline en ampicilline. Gezien de relatief hoge frequentie van ongemuteerde TEM-1 genen in bacteriële populaties, lijkt het erop dat recombinatie door de invloed van penicilline en ampicilline, die beiden nog steeds relatief veel gebruikt worden, er juist voor zorgt dat TEM-1 in bacteriële populaties blijft overheersen.

Niet alle veranderingen in de DNA-code veroorzaken veranderingen in de aminozuren waaruit een enzym is opgebouwd. Dit is het gevolg van de dubbelzinnigheid van de genetische code. Een aminozuur wordt in het DNA gecodeerd door drie opeenvolgende basen (een triplet). Het DNA kent 4 verschillende basen en zo zijn er dus 4x4x4 = 64 combinaties van basen mogelijk die een aminozuur coderen. Er zijn echter maar twintig aminozuren. Als gevolg hiervan worden de meeste aminozuren door meer dan één triplet van basen gecodeerd. Een mutatie (de verandering van één van de drie basen in zo'n triplet) hoeft dus niet automatisch een verandering in een aminozuur te betekenen. Mutaties die wel het DNA maar

niet het aminozuur veranderen, noemen we synonieme mutaties. Tot een jaar of tien geleden werd er algemeen aangenomen dat dergelijke mutaties een enzym niet kunnen veranderen, aangezien de mutatie niet leidt tot een verandering in de keten van aminozuren waaruit het enzym is opgebouwd. Recentelijk zijn er echter overtuigende bewijzen gevonden dat dit niet altijd het geval is. In **hoofdstuk 5** vinden we aanwijzingen dat tenminste één synonieme mutatie in TEM-1 niet neutraal is. Een analyse van alle synonieme mutaties die er in natuurlijke TEM-isolaten gevonden zijn toont aan dat er mogelijk meer van dergelijke niet-neutrale, synonieme mutaties in TEM-genen te vinden zijn.

Dankwoord/ Acknowledgements

Het is dus toch zover gekomen... en hoewel er op dit moment nog geen proefschrift naast me ligt, ben ik dan toch eindelijk aan één van de belangrijkste onderdelen toegekomen; het dankwoord. Ik moet u vooraf waarschuwen dat mijn dank groot is en mijn dankwoord derhalve lang.

In de eerste plaats wil ik mijn co-promotoren Arjan de Visser en John van der Oost bedanken. Arjan, we hebben van een halfvol glas en een halfleeg glas één vol glas weten te maken, wat achteraf een Cruijffiaanse wijsheid mag lijken, maar wat onderweg niet altijd voor de hand heeft gelegen. Ik ben je dankbaar voor de vrijheid die ik de afgelopen jaren gehad heb, je enthousiasme en je altijd snelle en nuttige commentaar op mijn geschrijf. John, hoewel dit proefschrift uiteindelijk geen 'enzyme engineering' kant opgegaan is, zijn je pragmatische aanpak en doelgerichtheid een inspiratie voor me geweest. Hopelijk kunnen we in de toekomst nog vaker dingen samen doen.

Ten tweede wil ik mijn promotor Rolf Hoekstra bedanken. Rolf, ik heb de mogelijkheid om te allen tijde bij je binnen te lopen, je openheid en je eerlijkheid in de afgelopen jaren zeer gewaardeerd.

Fons Debets wil ik bedanken voor zijn deskundige hulp bij het navigeren in woelige AIO-wateren. Fons, ik zal onze gesprekken over de zin en onzin van de wetenschap, je optimistische kijk en het 'goedesmorgens' in de gang gaan missen. Gelukkig weet ik dat er een ervaren appelvinken-determineerder in de groep achterblijft.

This thesis would not have been here in its current form without the help of Miriam Barlow. Miriam, you have been my online helpdesk for practical and moral support ever since the start of my project. Thanks for the warm welcome in Merced last year and for all your help with shaping this thesis. I hope we can continue our cooperation in the future. I'll try and look for a new antibiotic resistance gene to work on!

Aan het labwerk dat vooraf gegaan is aan de keurige grafiekjes en tabellen die in dit proefschrift staan is heel wat gezucht, gesteun, gezweet en gevloek vooraf gegaan. Die tijd was ik nooit doorgekomen zonder de deskundige en nuchtere hulp van Bertha Koopmanschap. Bertha, ik heb enorm veel respect voor de onopvallende manier waarop je ervoor zorgt dat alles op het moleculair lab dagelijks op rolletjes loopt. Dankzij jouw zorgvuldige en geduldige manier van werken voelden zelfs de soms dagenlange MIC-assays die we samen deden als een eitje. En los van die praktische kant ben je in de afgelopen jaren een enorme steunpilaar voor me geweest. Ik kan alleen maar hopen dat ik in de toekomst weer ergens zo'n geweldige laborante tref!

Edu Holub stond, met zijn enorme ervaring en oog voor detail, altijd klaar om een breed scala aan problemen op te lossen. Edu, bedankt daarvoor en ik hoop in de toekomst nog met regelmaat een bakkie met je te doen op het Forumgebouw!

Via Edu kwam ik ook in contact met Hildo Offenberg, die mij op moleculair gebied veel vooruit geholpen heeft. Mijn student Thomas van der Velden worstelde zich door enkele 'site-directed mutagenesis' protocollen voordat we eindelijk beet hadden. Dank voor je geduld! Stan Brouns en Ans Geerling van de vakgroep Microbiologie wil ik bedanken voor hun vele tips en hulp. Op het transitorium zorgde Marijke den Hartog ervoor dat het sequencen van mijn gen altijd vlotjes verliep. En natuurlijk is een lab niks zonder iemand die ervoor zorgt dat alles netjes bijgehouden wordt. Anita, bedankt voor de vele wasjes die je voor me gedraaid

hebt en ook voor de leuke gesprekken. Jan van Kreel wil ik bedanken voor de altijd snelle reparaties, groot en klein. Bram van Putten van Biometris hielp me met enkele statistische problemen en bleek nog niets van zijn aanstekelijke enthousiasme van tien jaar geleden (toen ik als student cursussen bij hem volgde) verloren te hebben. Vanuit PE & RC was Claudius van de Vijver voor mij een onmisbaar aanspreekpunt, veel dank daarvoor.

Joanna Mroczkowska carried out some high-tech molecular work for me in the U.S. Thanks for that and for all the tourist advice! Furthermore, I would like to thank Daniel Weinreich for his always enthusiastic correspondence and for the inspiring days at Brown University.

Het Botanisch Centrum is, zoals elke koffiepauze maar weer blijkt, een uitermate gezellig gebouw. Met het risico in dit rijtje per ongeluk iemand te vergeten (excuses alvast), wil ik hier al diegenen die voor mij bijgedragen hebben aan die dagelijkse gezelligheid en die elders in dit dankwoord niet voorkomen bedanken; Diaan, Mariëlle, Corrie, Henk, Dick, Sander, Joost, Ana, Wessel, Artak, Sangita, Adriaan, Thijs, Maurice, Judith, Jochem, Leo, Ronnie, Wilco, Henny en Annelies, bedankt!

En dan onze groep zelf. Te beginnen met het ‘oude’ secretariaat, waar Aafke van der Kooi en Corrie Eekelder niet alleen veel achter-de-schermen werk verzorgden, maar ook een bijzonder gezellige sfeer creëerden. Aafke, ik wil jou in het bijzonder bedanken voor de vele gesprekken die we hebben gehad, serieuze en minder serieuze, en voor je luisterend oor. In de jaren dat ik als AIO werkzaam ben geweest, is er een grote verscheidenheid aan AIO’s en postdocs de revue gepasseerd. Van de ‘oude’ garde wil ik Henk, Siemen, Judith en Pieter bedanken voor de gezelligheid. Marc Maas was in mijn eerste jaren naast mijn kamergenoot ook mijn orakel en een geweldige collega. Merci beaucoup maestro Maas!

Anne van Diepeningen volgde Marc op als kamergenoot. Anne, je was een wijze en lieve kamergenote waar ik veel van heb geleerd. Ik hoop dat je mij m’n vlucht naar het torenkamertje hebt kunnen vergeven, maar ik ben nu eenmaal bang voor het klappen van de zweep.

Stefan Bosmans begon enkele maanden na mij als AIO, zodat je ons bijna AIO-broertjes kunt noemen. Stefan, bedankt voor de gezelligheid, de vele biertjes en het vele geouwehoer. Ik weet zeker dat er in de nabije toekomst een groeiend aantal jonge leraren met een bijzonder enthousiasme voor de evolutiebiologie aan de Nederlandse middelbare scholen gaat verschijnen.

Michelle Habets en Daniel Rozen begonnen net iets eerder dan ik en aangezien ze er al snel achter kwamen dat ze elkaar nog leuker vonden dan experimentele evolutie wil ik ze hier ook samen bedanken. Michelle, je was een fantastische collega en bent een geweldige vriendin met wie ik èn kan lachen èn kan praten. Iech hoop dat iech nog dèks carnaval mag koume vieren in Maastricht! Danny, besides a skilled experimental evolutionist and excellent scientist, you were most of all a superb colleague. I hope I can visit the Rozen family in the U.K. somewhere soon.

And then Károly Pál, my pal in and outside of the lab. Károly, köszönöm, hogy jó barátom és kollégám voltál az elmúlt években. Mindig ott voltál mellettem jóban rosszban. A közös vacsorák, péntek délutáni fotó bulik, arborétumi séták, Rajna esték, bulik, kirándulások emlékét örökre szívembe zártam. Mikor legközelebb Magyarországra megyek téged és apukádat kerecsensólyom módjára találllak majd meg… te vagy a kedvenc majmom!

Dominika Wloch-Salamon and her family were always a safe haven and a center of nice parties and other gatherings. Dominika, thanks for everything and especially for kicking my butt at exactly the right time. Dziękuję and I hope I can make it to Krákow soon!

De ‘nieuwe’ garde, Bart, Anna en Tânia bedankt voor de gezellige tijd en succes de komende jaren. Bart, speciale dank voor het feit dat je je behendigheid met computerprogrammatuur graag met onwetenden deelt. ‘Bart’s random trekker’ blijft een klassiek stukje Excel-kunst. Ook rest van de groep, Klaas, Duur, Marijke, Wietske en de Daniels bedankt voor de fijne jaren.

Over afleiding buiten het werk om heb ik de afgelopen jaren niks te klagen gehad. Vanaf halverwege de middelbare school tot en met mijn eerste AIO-jaren heb ik deel uitgemaakt van de band ‘Another Messiah’. Martino, Chris en Robbie, bedankt voor die mooie tijd en voor jullie legendarische bezoeken aan Wageningen. Het afgelopen jaar was ik actief in MaMaToNi, een veelbelovende maar helaas te vroeg ter ziele gegane band voor feesten, partijen en de betere funk. Niels, Maria en Marc, ik heb ervan genoten – ik hoop dat het gaat lukken met het online jammen.

Ook op sportief vlak heb ik me bijzonder vermaakt de afgelopen jaren. Bij dezen wil ik dan ook mijn teamgenoten bij Pluto, de zaalvoetballers en de leden van volleybalteam ‘Puntwaard’ bedanken voor alle gezweet. Echt afzien werd mij bijgebracht door een stel onverschrokken bikkels met wortels aan de Churchillweg. Jongens, vuur en wind zullen mij geen angst meer inboezemen na de al dat water en die midges van afgelopen zomer. P.S.: het onderwerp ‘the bad steps’ laat ik hier even rusten, hoewel Koen en Karel wel beter weten natuurlijk.

Door de jaren heen heb ik in en om Wageningen, en ook ver daarbuiten, heel wat afgevogeld. Rondom het Botanisch Centrum waren Linus en Mark altijd enthousiaste ‘medestanders’, en het moet gezegd worden; de oplettendheid onder het koffiedrinken, het Arboretum-struinen en de halfjaarlijkse vogelexcursies hebben tot een aardige lijst geleid (zie Appendix II). Net buiten Wageningen heb ik vele uren versleten op enkele kaalgetrapte vierkante meters rondom ‘paaltje 20’ in de Blauwe Kamer. Dat was nooit zo leuk geweest zonder al die andere enthousiastelingen aldaar, bedankt! Tom, Bas, Raoul, Herman, Klaas en Reinoud (AVG_W!) en Marijn, Frank, Ralf en Wouter; bedankt voor de mooie trips, waarvan er hopelijk nog vele mogen volgen.

Een dikke tien jaar geleden begon ik aan mijn studie Biologie in Wageningen. Daaraan heb ik een aantal hele goede vrienden overgehouden. Harmen en Sander, bedankt voor de filmpjes, koffie, biertjes, eigen brouwsels, culinaire avondjes en meer. Ralf en Carlijn; idem, minus de eigen brouwsels en plus de gezellige tv-avondjes.

Brammetje, het dankwoord! Geen chantagemogelijkheden meer voor mij na het typen van deze woorden. Vele, vele weekenduren hebben we samen versleten, arboretum-wandelingen met koffie tot besluit. Ik hoop in de toekomst nog vaak met je te mogen stappen en filosoferen en mocht je de Staatsloterij niet winnen dan hoop ik later toch aan de slag te kunnen als schoonmaker in je praktijk in Salamanca…

Twee vrienden -en hun eega’s- zijn mij bijzonder dierbaar. Ten eerste Herman van Oosten en Agata Siedlecka. Van het Oosten, zonder jou zou Wageningen voor mij Wageningen niet zijn. Legendarische avonden zouden hun glans verliezen, en kleine, donkermantelige meeuwtjes zouden mij, tezamen met zo vele gevederde vriendjes, onopgemerkt passeren. Maar boven dit alles ben je een geweldige vriend. Agata en Herman, I have enjoyed the many, many visits to the Hoogstraat and still do (though hopefully soon they will be to some bigger place). You have both been an enormous support and great friends. Dzięki za wszystko!

Ten tweede Krijn en Silvie Paaijmans. Krijnemans, wat kan ik mij nog meer wensen dan een vriend als jij. Je hebt op zoveel verschillende manieren een bijdrage geleverd aan dit

proefschrift dat je naam bijna op de kaft zou mogen. Zonder jou, onze AIO-lunches, Boulevard-avondjes en de vele, vele telefoontjes als je weer eens in den vreemde zat, had ik het niet volgehouden allemaal. Ge zijt dn goeiste! Krijn en Silvie, ondanks dat ik jullie nu allebei moet missen vind ik het geweldig dat jullie het zo naar de zin hebben in de V.S. met huis, bos en Boogie. Bedankt voor alles!

Tot slot wil ik mijn familie bedanken voor al hun steun in de afgelopen jaren. Irene, als broertje is het maar wat fijn om ook een zusje te hebben waar je zo fijn mee kunt praten, eten en koffieleuten. Joost, Marjolijn, Kasper en Boris, het is altijd weer gezellig om op de Edisonstraat langs te komen. Ik ben er enorm trots op dat ik de oom ben van twee zulke leuke neven! Anne Pier en Laura, beiden als ervaren proefschriftschrijvers, bedankt voor jullie steun. De vele overzeese telefoontjes en mailtjes heb ik zeer gewaardeerd. En tot slot mijn ouders. Netty en Marten, zonder jullie had ik hier nu niet gestaan, letterlijk en figuurlijk. Jullie steun heeft in de afgelopen jaren vele vormen gekend; bezoekjes, telefoontjes, Noordwesthoek-briefjes, knipsels, ingevroren promotiehapjes en zelfs een tafeltje-dekje service. Bedankt dat jullie er altijd voor me zijn.

Merijn.

Curriculum vitae

On July 18th 1979, I, Merijn Salverda, saw my first light of day in Deurne, the Netherlands. After obtaining my athenaeum-diploma at the Peelland College in Deurne in 1997, I started my Biology studies at Wageningen University, the Netherlands. I specialized in animal biology and soon became fascinated by evolution and genetics. Lured by the fascinating stories and tempting pictures of the Mojave dessert, I carried out my first Msc project at the University of California Riverside where, together with Peter Boons and Roland den Hollander and under the inspiring supervision of Ties Huigens, I studied sex ratio distorting elements in field population of *Trichogramma*, tiny parasitoid wasps. Back in Wageningen, I continued on my second project, studying the cytogenetics of the PSR chromosome in *Trichogramma kaykai* at the Laboratory of Entomology under the supervision of Richard Stouthamer. Meanwhile, my interest in molecular aspects of evolution grew, and I eventually ended up at the Laboratory of Genetics, Wageningen University, where I studied the *Ant1* transposon of *Aspergillus niger* under the supervision of Fons Debets. Via him, I came into contact with Menno Schilthuizen, then at the University of Malaysia Sabah. Under his supervision and together with Herman van Oosten, I spent five months in Malaysian Borneo, studying the population genetics of the tiny but incredibly elegant *Opisthostoma* land snails. After finishing my Msc in 2002, I worked for six months as a field assistant in a meadow bird ecology project led by Jort Verhulst from the Nature Conservation and Plant Ecology group, Wageningen University. In September 2003, I started my PhD at the Laboratory of Genetics, Wageningen University, studying the *in vitro* evolution of an antibiotic resistance gene under the supervision of Arjan de Visser, John van der Oost and Rolf Hoekstra. The results of this project are presented in this thesis.

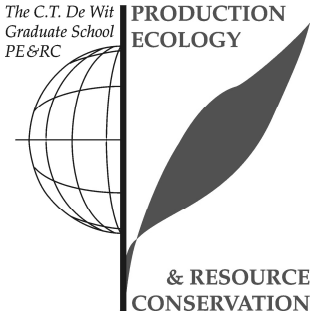
Publications

van Vugt, J.F.A., M. Salverda, J.H. de Jong, and R. Stouthamer. 2003. The paternal sex ratio chromosome in the parasitic wasp *Trichogramma kaykai* condenses the paternal chromosomes into a dense chromatin mass. Genome **46:** 580-587.

Schilthuizen, M., A. van Til, M. Salverda, S. Thor-Seng Liew, S.S. James, B. bin Elahan, and J.J. Vermeulen. 2006. Micro-geographical evolution of snail shell shape and predator behavior. Evolution **60 (9):** 1851-1858.

Barlow, M., J. Fatollahi, and M. Salverda. 2008. Evidence for recombination among plasmidic TEM-alleles. *Submitted to:* Journal of Antimicrobial Chemotherapy.

PE&RC PhD Education Certificate



With the educational activities listed below the PhD candidate
has complied with the educational requirements set by the C.T. de Wit Graduate School for Production Ecology and Resource Conservation (PE&RC) which comprises of a minimum total of 32 ECTS (= 22 weeks of activities)

Review of Literature (5.6 ECTS)

- 'Testing sex theories using the *in vitro* evolution of antibiotic resistance'

Laboratory Training and Working Visits (4.3 ECTS)

- Recombination in plasmidic TEM-alleles ; University of California Merced, Merced, U.S.A. (2007)
- Evolutionary pathways in TEM-1; Brown University, Providence, U.S.A. (2007)
- Empirical fitness landscapes; AMOLF institute, Amsterdam, the Netherlands (2008)

Post-Graduate Courses (4.0 ECTS)

- Population Genetics (2004)
- Enzyme engineering (2004)
- Bioinformatics (2007)

Competence Strengthening / Skills Courses (2.1 ECTS)

- Career perspectives (2006)
- PhD competence assessment (2006)

Discussion Groups / Local Seminars and Other Scientific Meetings (7.0 ECTS)

- Experimental evolution discussion group (2003-2008)
- Laboratory of Genetics discussion group (2003-2008)

PE&RC Annual Meetings, Seminars and the PE&RC Weekend (2.1 ECTS)

- PE & RC annual meeting (2003-2006)
- PE & RC introduction weekend (2005)

Symposia, Workshops and Conferences (7.0 ECTS)

- The evolutionary consequences of life without sex (2003)
- 10th PhD meeting in evolutionary biology (2004; oral presentation)
- Experimental evolution, fundamental and applied (2005)
- 10th congress of the European Society for Evolutionary Biology, Krákow, Poland (2005; poster)
- Joint annual meeting of the Society for the Study of Evolution, Stony Brook, U.S.A. (2006)

Courses in which the PhD has worked as a teacher

- Molecular and evolutionary ecology (2007); 30 days

Supervision of MSc students

- Testing controllable ways of DNA-shuffling; 30 days; 1 student

Appendix I

Supplementary table 3.1: Frequency of occurrence of amino acid substitutions in clinical and laboratory isolates of TEM-alleles.

Amino acid	Substitution	Total clinical isolates	Total laboratory isolates	Stemmer 1994	Zaccolo and Gherardi 1999	Blazquez *et al.* 2000	Long-McGie *et al.* 2000	Orencia *et al.* 2001	Barlow and Hall 2002	Barlow and Hall 2003	Camps *et al.* 2003	Fujii *et al.* 2004	Holloway *et al.* 2007	Kopsidas *et al.* 2007	Salverda *et al.* unpublished	Vakulenko *et al.* 1998	Vakulenko and Golemi 2002	Osuna *et al.* 2002	Bershtein *et al.* 2008	Hecky and Müller 2005
4	S4N		1																1	
5	I5T		1		1															
6	Q6K	4																		
	Q6R		1						1											
7	H7P		1																1	
	H7Q		1																1	
	H7R		1																1	
	H7Y		1																1	
10	V10I		6							2					4					
13	I13T		1																1	
	I13V		1																1	
15	F15C		1																1	
	F15L		1																1	
	F15V		1																1	
16	F16C		1																1	
	F16L	1																		
18	A18V		2	1											1					
19	F19C		2												1				1	
	F19I		1												1					
	F19L		2												1				1	
	F19V		1																1	
21	L21F	21																		
	L21I	1	1												1					
	L21P		3		2														1	
22	F22S		1											1						
23	V23A		2		2															
24	F24C		1												1					
27	P27L		1												1					
	P27S		1							1										
28	E28D	1																		
	E28K	1																		
31	V31A		3															1	1	1
32	K32R		1																1	
33	V33I		2												2					
34	K34E	1																		
35	D35A		1																1	
	D35P	1																		
36	A36T		1																1	
37	E37D		1																	1
38	D38G		1												1					
	D38N	1	1												1					
	D38V		1												1					
	D38Y		1							1										
39	Q39K	32	1												1					
	Q39R		2												1				1	
40	L40W		1									1								
41	G41S		1												1					
42	A42G		6	1					1						1	1	1			1
	A42V	1																		

Amino acid	Substitution	Total clinical isolates	Total laboratory isolates	Stemmer 1994	Zaccolo and Gherardi 1999	Blazquez *et al.* 2000	Long-McGie *et al.* 2000	Orencia *et al.* 2001	Barlow and Hall 2002	Barlow and Hall 2003	Camps *et al.* 2003	Fujii *et al.* 2004	Holloway *et al.* 2007	Kopsidas *et al.* 2007	Salverda *et al.* unpublished	Vakulenko *et al.* 1998	Vakulenko and Golemi 2002	Osuna *et al.* 2002	Bershtein *et al.* 2008	Hecky and Müller 2005
47	I47V		4												1		1	1	1	
48	E48V		1												1					
49	L49M	1	3												3					
	L49V		1												1					
50	D50E		1																1	
	L51F		2												1			1		
	L51P	1																		
52	N52D		3																2	1
52	N52H		1																1	
	N52S		3																2	1
	N52T		2																2	
	N52Y		1													1				
55	K55E		1																1	
	K55R		1															1		
56	I56T		1																	1
	I56V		1																1	
58	E58G		1															1		
	E58K		1												1					
59	S59G		1																	1
60	F60W		1																1	
62	P62S		1																1	
63	E63A		1																1	
	E63D		1																1	
	E63G		1																1	
	E63K		2																1	1
64	E64K	1																		
69	M69I	4																		
	M69L	12	1													1				
	M69V	9																		
79	A79V		1															1		
80	V80E	1																		
	V80I		1							1										
82	S82E		1																	1
	S82F		1																	1
88	Q88E		1																	1
	Q88H		1												1					
	Q88L		1												1					
90	G90S		1															1		
92	G92D	2	1												1					
	G92S		1	1																
96	H96D		1							1										
	H96Y		2																1	1
97	Y97F		1												1					
98	S98T		1															1		
	S98Y		1												1					
99	Q99R		1																1	
100	N100G		1																1	
	N100K		1												1					
	N100S	2																		
102	L102M		1												1					
	L102V	1																		
104	E104A		2															2		
	E104G		2																1	1
	E104K	42	36	1	1		2	1	7		2		4	1	16		1			
	E104V		1															1		
108	V108M		1															1		

Amino acid	Substitution	Total clinical isolates	Total laboratory isolates	Stemmer 1994	Zaccolo and Gherardi 1999	Blazquez *et al.* 2000	Long-McGie *et al.* 2000	Orencia *et al.* 2001	Barlow and Hall 2002	Barlow and Hall 2003	Camps *et al.* 2003	Fujii *et al.* 2004	Holloway *et al.* 2007	Kopsidas *et al.* 2007	Salverda *et al.* unpublished	Vakulenko *et al.* 1998	Vakulenko and Golemi 2002	Osuna *et al.* 2002	Bershtein *et al.* 2008	Hecky and Müller 2005
111	K111E		1																1	
	K111M		2												2					
	K111Q		2																2	
	K111R		3													1			2	
	K111T		2															1	1	
114	T114A		1																1	
	T114M		1												1					
	T114P	1																		
115	D115G	1																		
120	R120E		1																1	
120	R120G		5												1			1	2	1
	R120S		2																2	
124	S124G		2												1				1	
	S124N	1	1												1					
	S124R		2												1				1	
127	I127V	2																		
129	M129L		1												1					
130	S130G	3	1													1				
140	T140R		1												1					
145	P145A	1																		
	P145L		1												1					
	P145Q		1												1					
147	E147A		1																1	
	E147G		2																1	1
	E147K		2							1									1	
149	T149A		1												1					
150	A150T		1																	1
153	H153D		1													1				
	H153L		1												1					
	H153Q		1												1					
	H153R	3	6											1	1			2	1	1
	H153Y		1												1					
154	N154D		2															2		
	N154S		1																1	
	N154Y		1															1		
155	M155I		4							1					3					
157	D157E	1																		
158	H158N	1																		
	H158Y		1																1	
159	V159A		1																	1
	V159T																			
163	D163G	1																		
	D163H	1																		
	D163V		1												1					
164	R164C	4	2									1			1					
	R164G		2							1		1								
	R164H	18	15			5			1	6	1	1			1					
	R164N		1						1											
	R164S	26	21		1	11			2	1	1				5					
	R164Y		1												1					
165	W165C	1																		
	W165G		1												1					
	W165R	3																		
167	P167S		1														1			
	P167T		1												1					

Amino acid	Substitution	Total clinical isolates	Total laboratory isolates	Stemmer 1994	Zaccolo and Gherardi 1999	Blazquez *et al.* 2000	Long-McGie *et al.* 2000	Orencia *et al.* 2001	Barlow and Hall 2002	Barlow and Hall 2003	Camps *et al.* 2003	Fujii *et al.* 2004	Holloway *et al.* 2007	Kopsidas *et al.* 2007	Salverda *et al.* unpublished	Vakulenko *et al.* 1998	Vakulenko and Golemi 2002	Osuna *et al.* 2002	Bershtein *et al.* 2008	Hecky and Müller 2005
168	E168A		1																	1
	E168G		1															1		
	E168K		1						1											
	E168V		1																	1
169	L169P		2												1		1			
	L169R		2			2														
172	A172V		1												1					
173	I173L		1																1	
	I173R		1															1		
	I173T		2						1						1					
	I173V	1	12						3	6		1			2					
174	P174S		1							1										
175	N175D		5							1					3				1	
175	N175I	1																		
	N175Y		1												1					
177	E177K		1																	1
178	R178C		1									1								
	R178S		1							1										
179	D179E	1																		
	D179G		4			1						1			2					
	D179Y		1			1														
180	T180S		1												1					
182	M182K		1																1	
	M182L		1																1	
	M182R		1																1	
	M182T	21	26	1	1			1	2	2				1	14			1	2	1
	M182V		2																2	
184	A184T		1												1					
	A184V	(3)	4												3		1			
188	T188I		2												2					
	T188R		1												1					
191	R191H		2												2					
192	K192E		2												1					1
	K192N		1												1					
	K192R		1													1				
195	T195A		1													1				
	T195F		1																	1
	T195S		1																	1
196	G196D	1																		
	G196S	1	1												1					
197	E197K		1												1					
198	L198P		1																	1
199	L199P		1		1															
200	T200S		1												1					
201	L201I		1												1					
	L201P		4												2	1			1	
	L201Q		1																1	
	L201R		1																1	
202	A202V		1		1															
204	R204Q	1																		
205	Q205R		1												1					
206	Q206H		1																	1
207	L207S		1		1															
208	I208K		1												1					
	I208L		1																1	
	I208M		5												2			1	1	1
	I208T		1																1	
	I208V		2		1														1	

Amino acid	Substitution	Total clinical isolates	Total laboratory isolates	Stemmer 1994	Zaccolo and Gherardi 1999	Blazquez *et al.* 2000	Long-McGie *et al.* 2000	Orencia *et al.* 2001	Barlow and Hall 2002	Barlow and Hall 2003	Camps *et al.* 2003	Fujii *et al.* 2004	Holloway *et al.* 2007	Kopsidas *et al.* 2007	Salverda *et al.* unpublished	Vakulenko *et al.* 1998	Vakulenko and Golemi 2002	Osuna *et al.* 2002	Bershtein *et al.* 2008	Hecky and Müller 2005
209	D209G		1																1	
212	E212K		1															1		
213	A213P		1												1					
215	K215E		1		1															
	K215R		1																1	
216	V216I		2												2					
217	A217S		1												1					
218	G218E	1																		
219	P219S		1															1		
221	L221M	1																		
224	A224V	1	7							1					4			1		1
227	A227V		1									1								
235	S235T		2												2					
237	A237G	1																		
	A237S		2						2											
	A237T	8	6												6					
238	G238A		3										3							
	G238D	1																		
	G238N	1																		
	G238S	32	57	1	5		2	1	6				4	1	35	1	1			
240	E240G		4		2												1			1
	E240K	25	13		1				2						10					
	E240R	1																		
	E240V	1	1												1					
241	R241H		4	1	1										2					
243	S243G		1																1	
244	R244C	4	1													1				
	R244G	1																		
	R244H	3																		
	R244L	1																		
	R244S	7	1													1				
	R244T																			
247	I247V		3												1				1	1
248	A248T		1												1					
249	A249V		3												3					
254	D254G		3		1				1										1	
	D254H		1						1											
	D254N		2		2															
256	K256N		1												1					
	K256Q		1		1															
	K256R		1																	1
262	V262I	1																		
265	T265A		2		2															
	T265M	20	12				1		2						7				1	1
266	T266A		1						1											
267	G267A		1												1					
	G267R		2						1		1									
	G267V		1												1					
268	S268G	2	7										4	1	1		1			
	S268N		1							1										
	S268T		1												1					
271	T271A		1																1	
	T271I		2												2					
272	T272A		1													1				
275	R275A	2																		
	R275L	3	1															1		
	R275Q	4	1															1		

Amino acid	Substitution	Total clinical isolates	Total laboratory isolates	Stemmer 1994	Zaccolo and Gherardi 1999	Blazquez *et al.* 2000	Long-McGie *et al.* 2000	Orencia *et al.* 2001	Barlow and Hall 2002	Barlow and Hall 2003	Camps *et al.* 2003	Fujii *et al.* 2004	Holloway *et al.* 2007	Kopsidas *et al.* 2007	Salverda *et al.* unpublished	Vakulenko *et al.* 1998	Vakulenko and Golemi 2002	Osuna *et al.* 2002	Bershtein *et al.* 2008	Hecky and Müller 2005
276	N276D	12	3													1	1	1		
	N276R	1																		
	N276S	1	1															1		
277	R277K		1																	1
278	Q278E		1												1					
	Q278L		1															1		
	Q278R		2												1				1	
280	A280T		1															1		
	A280V	2																		
281	E281-		1												1					
284	A284G	1																		
285	S285T		1																1	
287	I287V		2							1									1	
288	K288E		2														1		1	
288	K288Q		1																1	
	K288R		2															1	1	
	K288T		1																1	
289	H289L																			
	H289Q		1												1					
	H289Y		1							1										
291	ter291K		1												1					

Appendix II

Supplement to acknowledgements. List of bird species recorded in and from the 'Dreijen Arboretum' and Botanical Centre, Wageningen, the Netherlands, in the period 2002-2008 (observations by L. van der Plas and M. Salverda). Explanation of status; 1 = breeding or strong indications of territory, 2 = regular but probably non-breeding visitor, 3 = fly over, 4 = rare visitor, 5 = rare fly over, 6 = seen during the bird watching excursions of the Botanical Centre (which include a walk to the 'Nevengeul').

.

Nr.	Dutch name	Status	English name	Latin name
1	Fuut	6	Great crested grebe	*Podiceps cristatus*
2	Aalscholver	2	Great cormorant	*Phalacrocorax carbo*
3	Kleine zilverreiger	6	Little egret	*Egretta garzetta*
4	Grote zilverreiger	5	Great egret	*Casmerodius albus*
5	Blauwe reiger	2	Grey heron	*Ardea cinerea*
6	Ooievaar	3	White stork	*Ciconia ciconia*
7	Lepelaar	5	Spoonbill	*Platalea leucorodia*
8	Knobbelzwaan	3	Mute swan	*Cygnus olor*
9	Kolgans	3	White-fronted goose	*Anser albifrons*
10	Grauwe gans	3	Greylag goose	*Anser anser*
11	Wilde eend	2	Mallard	*Anas platyrhynchos*
12	Kuifeend	6	Tufted duck	*Aythya fuligula*
13	Rode wouw	5	Red kite	*Milvus milvus*
14	Bruine kiekendief	5	Marsh harrier	*Circus aeruginosus*
15	Blauwe kiekendief	5	Hen harrier	*Circus cyaneus*
16	Buizerd	3	Buzzard	*Buteo buteo*
17	Wespendief	5	Honey buzzard	*Pernis apivorus*
18	Sperwer	2	Sparrowhawk	*Accipiter nisus*
19	Havik	5	Goshawk	*Accipiter gentilis*
20	Torenvalk	2	Kestrel	*Falco tinnunculus*
21	Boomvalk	3	Hobby	*Falco subbuteo*
22	Waterhoen	1	Moorhen	*Gallinula chloropus*
23	Meerkoet	6	Eurasian coot	*Fulica atra*
24	Scholekster	3	Oystercatcher	*Haematopus ostralegus*
25	Kievit	3	Lapwing	*Vanellus vanellus*
26	Tureluur	6	Redshank	*Tringa totanus*
27	Grutto	6	Black-tailed godwit	*Limosa limosa*
28	Wulp	6	Curlew	*Numenius arquata*
29	Houtsnip	4	Eurasian woodcock	*Scolopax rusticola*
30	Kokmeeuw	3	Black-headed gull	*Larus ridibundus*
31	Stormmeeuw	3	Mew gull	*Larus canus*
32	Zwartkopmeeuw	5	Mediterranean gull	*Larus melanocephalus*
33	Zilvermeeuw	3	Herring gull	*Larus argentatus*
34	Kleine mantelmeeuw	3	Lesser black-backed gull	*Larus graellsii*
35	Visdief	4	Common tern	*Sterna hirundo*
36	Holenduif	1	Stock dove	*Columba oenas*
37	Houtduif	1	Woodpigeon	*Columba palumbus*
38	Turkse tortel	2	Collared dove	*Streptorelia decaocto*
39	Koekoek	2	Cuckoo	*Cuculus canorus*
40	Velduil	5	Short-eared owl	*Asio flammeus*
41	Steenuil	4	Little owl	*Athene noctua*
42	Gierzwaluw	3	Swift	*Apus apus*
43	Alpengierzwaluw	5	Alpine swift	*Apus melba*
44	IJsvogel	2	Kingfisher	*Alcedo atthis*
45	Zwarte specht	5	Black woodpecker	*Dryocopus martius*
46	Groene specht	2	Green woodpecker	*Picus viridis*
47	Grote bonte specht	1	Great spotted woodpecker	*Dendrocopos major*
48	Kleine bonte specht	4	Lesser spotted woodpecker	*Dendrocopos minor*
49	Veldleeuwerik	3	Skylark	*Alauda arvensis*
50	Oeverzwaluw	3	Sand martin	*Riparia riparia*
51	Boerenzwaluw	3	Barn swallow	*Hirundo rustica*

52	Huiszwaluw	3	House martin	*Delichon urbica*
53	Graspieper	3	Meadow pipit	*Anthus pratensis*
54	Witte kwikstaart	2	White wagtail	*Motacilla alba*
55	Gele kwikstaart	6	Yellow wagtail	*Motacilla flava*
56	Grote gele kwikstaart	5	Grey wagtail	*Motacilla cinerea*
57	Winterkoning	1	Wren	*Troglodytes troglodytes*
58	Pestvogel	4	Bohemian waxwing	*Bombycilla garrulus*
59	Heggenmus	1	Dunnock	*Prunella modularis*
60	Roodborst	1	Robin	*Erithacus rubecula*
61	Blauwborst	6	Bluethroat	*Luscinia svecica*
62	Zwarte roodstaart	1	Black redstart	*Phoenicurus ochruros*
63	Roodborsttapuit	6	Stonechat	*Saxicola rubicola*
64	Zanglijster	2	Song thrush	*Turdus philomelos*
65	Koperwiek	2	Redwing	*Turdus iliacus*
66	Grote lijster	1	Mistle thrush	*Turdus viscivorus*
67	Kramsvogel	2	Fieldfare	*Turdus pilaris*
68	Merel	1	Blackbird	*Turdus merula*
69	Beflijster	4	Ring ouzel	*Turdus torquatus*
70	Tuinfluiter	6	Garden warbler	*Sylvia borin*
71	Zwartkop	1	Blackcap	*Sylvia atricapilla*
72	Braamsluiper	1	Lesser whitethroat	*Sylvia curruca*
73	Grasmus	6	Whitethroat	*Sylvia communis*
74	Sprinkhaanzanger	6	Grasshopper warbler	*Locustella naevia*
75	Fitis	2	Willow warbler	*Phylloscopus trochilus*
76	Tjiftjaf	1	Chiffchaff	*Phylloscopus collybita*
77	Goudhaan	1	Goldcrest	*Regulus regulus*
78	Vuurgoudhaan	2	Firecrest	*Regulus ignicapillus*
79	Grauwe vliegenvanger	4	Spotted flycatcher	*Muscicapa striata*
80	Bonte vliegenvanger	4	Pied flycatcher	*Ficedula hypoleuca*
81	Koolmees	1	Great tit	*Parus major*
82	Zwarte mees	2	Coal tit	*Parus ater*
83	Pimpelmees	1	Blue tit	*Parus caeruleus*
84	Kuifmees	1	Crested tit	*Parus cristatus*
85	Glanskop	2	Marsh tit	*Parus palustris*
86	Staartmees	1	Long-tailed tit	*Aegithalos caudatus*
87	Boomklever	1	Nuthatch	*Sitta europaea*
88	Boomkruiper	1	Short-toed treecreeper	*Certhia brachydactyla*
89	Ekster	1	Magpie	*Pica pica*
90	Gaai	1	Jay	*Garrulus glandarius*
91	Kauw	3	Jackdaw	*Corvus monedula*
92	Zwarte kraai	1	Carrion crow	*Corvus corone*
93	Spreeuw	2	Starling	*Sturnus vulgaris*
94	Huismus	2	House sparrow	*Passer domesticus*
95	Ringmus	4	Tree sparrow	*Passer montanus*
96	Vink	1	Chaffinch	*Fringilla coelebs*
97	Keep	4	Brambling	*Fringilla montifringilla*
98	Kleine barmsijs	4	Mealy redpoll	*Carduelis cabaret*
99	Putter	2	Goldfinch	*Carduelis carduelis*
100	Groenling	1	Greenfinch	*Chloris chloris*
101	Sijs	2	Siskin	*Carduelis spinus*
102	Europese kanarie	4	Serin	*Serinus serinus*
103	Goudvink	1	Bullfinch	*Pyrrhula pyrrhula*
104	Appelvink	1	Hawfinch	*Coccothraustes coccothraustes*
105	Kruisbek	4	Crossbill	*Loxia curvirostra*
106	Rietgors	4	Common reed bunting	*Emberiza schoeniclus*

Exotic species and escapes

107	Nijlgans	3	Egyptian goose	*Alopochen aegyptiacus*
108	Fazant	6	Ring-necked pheasant	*Phasianus colchicus*
109	Japanse pestvogel	4	Japanese waxwing	*Bombycilla japonica*
110	Kanarie	4	Canary	*Serinus canaria*

Funding
The work presented in this thesis was carried out at the Laboratory of Genetics, Wageningen University, the Netherlands, with financial support from the Netherlands Organization for Scientific Research (NWO), grant number 813.04.003.

Publisher
Ponsen & Looijen bv, Wageningen, the Netherlands