Essentiality of conserved amino acid residues in β-lactamase

Chikunova, A.

Chapter 1
Introduction

Protein evolution

"...it is not the most intellectual of the species that survives; it is not the strongest that survives; but the species that survives is the one that is able best to adapt and adjust to the changing environment in which it finds itself."

Leon C. Megginson, 1963

All organisms are optimized for their environment and are found in a state that shows little or no morphological change over time. This state is called stasis and represents the best phenotype in a given condition. However, if the environment changes, it forces organisms to change too. These changes were believed to occur gradually and continuously via a process called anagenesis, transforming the whole species into another[1]. This phyletic gradualism model was later transformed into a punctuated equilibria model, which was proposed by Niels Eldredge and Stephen Jay Gould in 1972[2]. In contrast to anagenesis, according to the punctuated equilibria model, species are changing via cladogenesis, dividing into two distinct species. Cladogenesis is a response to a rare and rapid environmental change and happens in a short period of time. However, our perception of "short" and "rapid" is different from the evolutionary time scale. Charles Darwin wrote that "although each species must have passed through numerous transitional stages, it is probable that the periods, during which each underwent modification, though many and long as measured by years, have been short in comparison with the periods during which each remained in an unchanged condition"[3], so the transition from one stasis to another still happens gradually, with no noticeable difference between generations.

Under changed conditions organisms need to be able to adapt, so their proteins need to evolve, and they have been doing it successfully for billions of years. Our understanding of protein evolution is far from complete. While we might have obtained a great portion of knowledge about molecular mechanisms involved, we still lack the full knowledge of fundamental aspects of it.

The evolutionary process in proteins is determined by three main components: mutation in the DNA code, which occurs randomly and is the mechanism of evolution, selection, which determines the viability of mutations, and drift, which allows for a fixation of neutral mutations. Each single mutation, in order to persist, needs to be "accepted" by a functional protein. This was first proposed by John Maynard Smith in 1970. According to his concept, for evolution by natural selection to occur, functional proteins must form a continuous network in sequence space with single mutation steps without passing through nonfunctional intermediates[4]. A selective model of evolution was for a long time the main model. It suggested that all mutations are either detrimental for a protein and, therefore are not carried on, or beneficial. In 1968, Motoo Kimura for the first time proposed the neutral model of evolution, arguing that most mutations arising during evolution are actually neutral[5,6], a theory that is well-accepted today. Although initially he based his theory on a very small dataset, later he was able to find proof for its relevance in the facts that conservative substitutions (changes to another amino acid of the same group) occur more often than non-conservative and that the rate of synonymous mutations (DNA changes in a gene that do not affect the amino acid sequence) are almost always higher than non-synonymous ones (DNA changes leading to changes in the amino acid sequence)[7].

Proteins populate fitness peaks, which are the local optima in the fitness landscape. Fitness is a quantitative representation of an evolutionary selection process, and as such fitness peaks can be considered phenotypes that are more successful in present conditions than other phenotypes. Changing environmental conditions affect the fitness of the organism and thus may change the fitness landscape of its proteins. A number of fitness criteria for proteins have been proposed, such as foldability, folding rate, stability, substrate affinity and, for enzymes, accessibility of active site residues[8–12]. If a protein adopts a new function, it needs to traverse the fitness landscape from one fitness peak to another. That means it must pass through lower fitness intermediates, an adaptive valley. Weak selection that purges only detrimental mutations aids in crossing from one fitness peak to another, because it allows for accumulation of neutral mutations. That was experimentally demonstrated with in-lab evolution performed on TEM-1 β-lactamase[13]. Crossing through this valley requires a close coaction of two important evolutionary properties: evolvability and evolutionary robustness.

Evolvability and evolutionary robustness

Evolvability and evolutionary robustness are fundamental features of biological systems on all levels of organization. Evolvability or adaptability describes the capability of a system to acquire a new trait. Evolutionary robustness or neutrality describes the phenotypical resistance to changes, in terms of proteins, the ability to accumulate mutations without losing the original function (Figure 1.1). Enzymes are often highly robust to mutations. It was shown for example, for bacteriophage T4 lysozyme that mutations in more than half of the sequence are well tolerated[14]. Evolvability on the other hand is not a basic property of enzymes, some families are known for their high evolvability and are able to acquire a new function relatively easy, such as β-lactamases[15,16], and some enzyme families are highly specific and invariable, such as dihydrofolate reductase[17]. Robustness and evolvability may appear to be contractionary traits at first, but evolvability is heavily dependent on robustness[18,19]. To obtain a new functional mutation, proteins must sample various non-functional mutations without losing the native phenotype, and, most importantly, in a robust system a phenotype is presented by an ensemble of genotypes (neutral network), which increases the number of pathways evolution can take to generate a new phenotype.

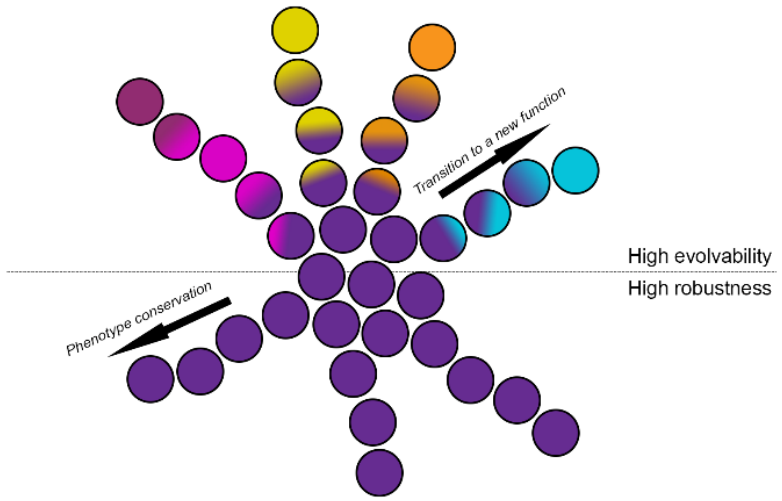

Figure 1.1. Schematic representation of evolvability and evolutionary robustness traits. Each circle represents a genotype, color represents phenotype. Mutational pathways are represented by successive genotypes radiating from the center.

Another interesting concept, which to date has not received much attention, is evolution of evolvability. It seems logical that nature would favor mutations that enhance evolvability, at least for species that have a high chance of meeting new selection pressures due to environmental changes. However, the underlying mechanisms are extremely difficult to define. So, is evolvability evolvable? Some studies oppose this concept[20,21], while others argue that it is plausible[22,23]. A study by Zheng *et al.* (2020) showed that selection does indeed enhance evolvability by increasing robustness[24]. In this study a population with yellow fluorescence protein was subjected to either strong or weak selection and then evolved toward green fluorescence. It was then demonstrated that the population under weak selection adapted the new phenotype at a higher frequency at first, however, eventually, the population under strong selection evolved more rapidly due to increased robustness and foldability in this population.

Role of stability in the evolution process

Structural stability (and foldability) of a protein is named as one of the main fitness criteria, however, generally, only marginal stability and foldability are required to obtain functional enzymes. Enzymes can therefore evolve within a "neutral" range of stability that maintains their fitness. Mutations that do not cause a large destabilizing effect are carried on. Such destabilizing mutations occur relatively often (about a third of all mutations[25]) and are purged by negative selection. Some mutations, although rare, can increase stability too much, compromising the dynamics of the enzyme, and these mutations also do not persist.

Protein stability promotes evolutionary robustness. Stabilizing mutations do not influence fitness on their own, but they buffer the destabilizing effect of other mutations. Stabilizing

mutations can cause local or global effects, compensating for destabilizing mutations of distant residues. Most functionally beneficial mutations in enzymes are destabilizing, as active site residues must satisfy a number of constraints, such as geometry, interaction ability, and hydrophobicity. These make active site residues highly optimized for activity and, consequently, poorly optimized for stability[26]. It was even proposed by Beadle and Shoichet that regions with relative instability or "regions where stability rules are broken" can be used as an indicator for active sites in enzymes with unknown function[27]. Stability and activity tradeoff has been a subject of debate for decades. It was first demonstrated by Wang *et al.* that mutations leading to new functions are highly destabilizing[28]. While some works refer to activity and stability as conflicting features[27,29], several studies showed that some highly stabilized enzymes can still be fully functional[8,30]. Studies of ancestral enzymes demonstrated that these proteins were much more stable than today's enzymes, and that the proteins lost the excess stability during evolution for the sake of new functions[31]. It was also shown that destabilized mutants with improved functions become closer to the consensus/ancestor sequence upon acquiring the stabilizing mutations. For example, many clinically isolated TEM β-lactamase mutants harbor the stabilizing mutation M182T, while for ancestral β-lactamase threonine is the canonical residue at position 182[31]. Tokuriki *et al.* (2008) indicated that although new-function mutations surely are destabilizing, the destabilization effect is not greater than that of a random destabilizing mutation[32]. Thus, for a new function to evolve, an enzyme must be able to cope with destabilizing mutations. In that way, general stability promotes evolvability. Multiple studies showed a link between increased stability and increased evolvability. This has been demonstrated computationally and also experimentally for cytochrome P450 by Bloom *et al.*[33]. On the other hand, more stability can mean less plasticity and a more rigid protein, which generally impairs evolvability, at least in the case of enzymes. Dellus-Gur *et al.* demonstrated using TEM-1 that it is critical that the enzyme scaffold is stable but that the active site needs sufficient flexibility[17].

Conserved residues

Protein evolution works via mutagenesis, and for that reason it is desirable to have a limited number of essential residues. A small number of essential residues is directly linked to evolutionary robustness, as it ensures that most mutations have a low chance of a detrimental effect on protein. Also evolvability benefits from a small number of crucial residues, because it leaves more room for new and possibly functional mutations to occur.

Mutation of an essential residue may affect the enzyme in a variety of ways (Figure 1.2), all leading to loss of enzyme function. These residues are usually highly conserved among orthologous proteins and, thus, conservation is used as a proxy for essentiality[34–37]. Considering all the factors that are in play to create a functional protein (Figure 1.2), the number of highly conserved residues in enzymes is surprisingly low.

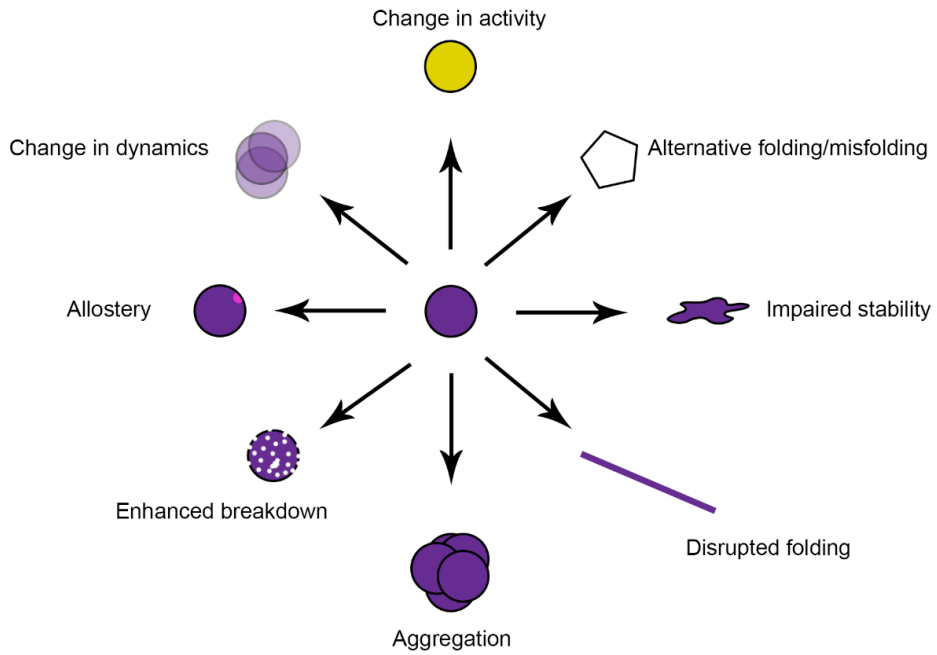

Figure 1.2. Schematic representation of possible effects of mutation of a conserved residue.

Residues involved in enzymatic catalysis, unsurprisingly, tend to be highly conserved. A link between evolution rate and the distance from the active site was demonstrated by Jack *et al.* An almost linear dependence was reported, showing decreasing conservation with increasing distance from the active site[38]. However, even far from catalytic sites of enzymes conserved residues can be found. Often, these distant conserved residues are discussed in the context of maintaining the three-dimensional structure, i.e. the fold[39–41]. The number of protein folds is insignificant compared to the number of reactions that can be catalyzed by enzymes. Proteins with similar fold can be involved in highly diverse reactions, for example phosphocarrier HPr protein and human carboxypeptidase A2 share the α/β plait fold[42]. Proteins with the same fold have the same secondary structure elements arranged in the same three-dimensional structure, yet the residues in these elements can vary greatly. Fold conservation is much greater than sequence conservation. Similar folds can be obtained with sequence identity as low as 10-15%[43]. This observation suggests that the folds have a small set of "staple" residues that arrange secondary structure into three dimensions. These residues are usually highly conserved. The presence of such small number of "staple" residues aids the evolutionary robustness of a protein, as it leaves more freedom for random mutations to occur inside the secondary structure elements, while the 3D arrangement stays intact. The concept of stapling residues is discussed in more detail in Chapter 2. Conserved residues are often found to

form clusters[44]. Such clusters can be found, for example, at protein-protein interfaces[45] or be involved in formation of the folding nucleus, which aids the formation of the fold[42]. In the study by Mirny and Shakhnovich the five most common folds were analyzed for detection of conserved residues that might illuminate the folding nuclei. For each of the studied folds they identified positions that are conserved in most of the protein families sharing this fold. The residues were forming clusters but the nature of interactions stabilizing these clusters were found to be different in different folds and even in different families of the same fold[42].

Here, we divide conserved residues of enzymes into three groups based on their position in the structure. The first-shell residues form the active site, responsible for substrate binding and catalysis. The second group includes the residues outside active site but in close proximity to the residues of the active site, the residues of this group are not directly involved in the enzymatic action (second shell). Catalysis requires accurate positioning of the atoms involved in chemical reactions, so residues of the second shell can be important for the precise sub-Ångström positioning of the residues of the active site. The third group or third shell includes the residues that are located very far from the active site, spread over the protein, and without direct interactions with active site residues. This group of residues is likely to be involved in maintaining the protein fold, creating a 3D-scaffold for the active site.

Subject of the study

"The Captain of all these men of death, that came against him to take him away, was the Consumption, for it was that that brought him down to the grave"

John Bunyan, 1680

Tuberculosis

Tuberculosis ("Consumption") is a deadly disease, a leader of all infectious diseases, which accompanied humans throughout history. Multiple studies confirm the presence of the disease in populations dating from more than 8000 years ago, as well as in many Egyptian mummies[46–50]. Although no hard evidence is present, it has been argued that human beings fell victim to tuberculosis even earlier, since the early human migrations[51]. While first human populations suffered from tuberculosis incidentally, urbanization increased population density enabling tuberculosis to be spread more easily. This development, combined with the lack of a treatment, led to high mortality due to tuberculosis. The incidence of the disease peaked in the 18th-19th centuries, becoming an epidemic in Europe. Even nowadays, until the rise of coronavirus infections in 2020, tuberculosis was the leading cause of death from a single infectious agent[52]. Acknowledgement of tuberculosis as a disease can be traced also in literature throughout history, in which it is described under various names already in ancient times by Cicero, Hippocrates[53] and in the Old Testament[54]. Later, in the 19th century, because of the wide spread of tuberculosis, it even became a fashion statement. The looks of patients with tuberculosis were romanticized for their pale faces with rosy cheeks and eyes that "sparkled as bright as diamonds"[55].

Throughout history physicians have tried to propose a cause of tuberculosis disease. The theories varied from it being hereditary to being spontaneous. In 1865, Jean-Antoine Villemin first reported that he was able to transmit tuberculosis from a human patient to a rabbit. And in 1882, Heinrich Hermann Robert Koch announced that he found a pathogen responsible for the disease, a discovery honored with the Nobel Prize in 1905. This pathogen was named *Mycobacterium tuberculosis* in 1883. *M. tuberculosis* is the main etiologic agent of tuberculosis. It has no environmental reservoir, with humans being its natural host and is believed to have co-evolved with us over millennia. For example, the ability of the pathogen to encapsulate itself inside a host and stay dormant for decades, which developed 20,000-30,000 years ago, is considered a survival strategy in response to limited contagiosity in small, widely dispersed populations of human beings[51]. On the other hand, the genome of humans adapted to the tuberculosis infection as well[56].

Giving the spread and lethality of tuberculosis, the human race always faced a need for prevention and remedy but attempts for treatment were for a long time ineffective. For many

years starting from ancient times fresh air, milk and sea voyages were recommended as the cure for tuberculosis and, while these recommendations are generally nice to follow and may have benefitted the general condition of the patient, they did little harm to a deadly infection. Starting from 496 and for the next several hundred years, a "royal touch" (a healing touch by a hand of a royal)[57] was commonly used in Europe to treat extrapulmonary tuberculosis condition, which, unsurprisingly, also did not work. High-altitude sanatoria became very popular in the 19th century, where the main treatment course included aero- and heliotherapy, better nutrition, and bed rest. At the same time, more harmful and unsafe methods such as bloodletting, chest massages with acetic acid, inhalation of antiseptics and even pneumonectomy and "artificial pneumothorax"[58] were implemented in a desperate search for a therapy. The big step towards the control of tuberculosis spreading was the introduction of the BCG (Bacille Calmette-Guerin) vaccine by Albert Calmette and Camille Guerin in 1921[59], which is still widely used nowadays for vaccination of newborn babies in high-risk countries, but it was only by the discovery of antibiotics later in 20th century that tuberculosis became curable. The first two drugs discovered in 1944 with a clear effect on *M. tuberculosis* were streptomycin, introduced by Selman Walksman and Albert Schatz[60] and para-aminosalycilic acid (PAS) synthesized by Jörgen Lehmann[61]. Although resistance to these drugs developed rather quickly, the possibility of curing tuberculosis boosted extensive research in pursuit of new compounds or to test existing compounds against *M. tuberculosis*.

Drug therapy of tuberculosis posed another issue. High mutability of *M. tuberculosis* requires the use of a combination therapy. Moreover, due to the thick cell wall of *M. tuberculosis*, its very slow generation time and its ability to become a dormant infection inside granulomas in lung tissue, the antibiotic treatment needs to be carried out for months. Modern chemotherapy of tuberculosis consists of co-administration of ethambutol, isoniazid, rifampicin, and pyrazinamide for six months. Such prolonged therapy of multiple drugs is heavily associated with poor medication compliance. Incomplete antibiotic treatments resulted in isolation of multidrug-resistant (MDR) strains in the 1990s[62], followed by isolation of extensively drug-resistant (XDR)[63] and totally drug-resistant strains in the 2000s[64–67]. Nowadays, more than half a million new MDR cases are reported each year with about 6% of them being XDR forms.

β-Lactam antibiotics

The rise of antibiotic resistance in *M. tuberculosis* puts pressure on humanity to seek new antibiotics. The process of new drug development and approval for clinical use is very complex and time-consuming. Furthermore, the perceived poor business case for new antibiotics has led to limited interest of large pharma companies to pursue their development[68]. Perhaps, simplicity is the ultimate sophistication and the need for a new antituberculosis strategy can be satisfied with the existing remedies. The most widely used group of antibiotics are β-lactam antibiotics. The first antibiotic discovered in 1929 by Alexander Fleming was a β-lactam[69], nowadays this

group of antimicrobials counts multiple subgroups with numerous compounds in each[70] (Figure 1.3). β-Lactam antibiotics are well-researched and there are many advantages of the use of this group of drugs, such as low toxicity, low cost, and a simple delivery method. However, β-lactam antibiotics are not generally used in antituberculosis therapy because *M. tuberculosis* produces a β-lactamase, which makes the bacteria resistant to this group of antibiotics[71]. The production of β-lactamases is the most common mechanism of resistance to β-lactam antibiotics[72], and it was overcome by the introduction of β-lactamase inhibitors. β-Lactamase inhibitors are compounds that bind to β-lactamases, rendering the enzyme inactive, thus allowing β-lactam antibiotics to act. The combination of β-lactamase inhibitors with β-lactam antibiotics was shown to be effective against *M. tuberculosis* in multiple studies[73–76] and has been recommended as a treatment of last resort by the WHO[77,78]. It is still possible for β-lactamases to gain resistance against inhibitors through sequence modifications. Many inhibitor-resistant variants of *Escherichia coli* β-lactamases were clinically isolated[79]. Despite evidence once suggesting that resistance against β-lactamase inhibitors is not likely to arise due to mutations in β-lactamase in *M. tuberculosis*[80], inhibitor inactivation due to synergistic double mutations in BlaC has since been observed[81], as well as a few single mutants less susceptible to inhibitors[82–85].

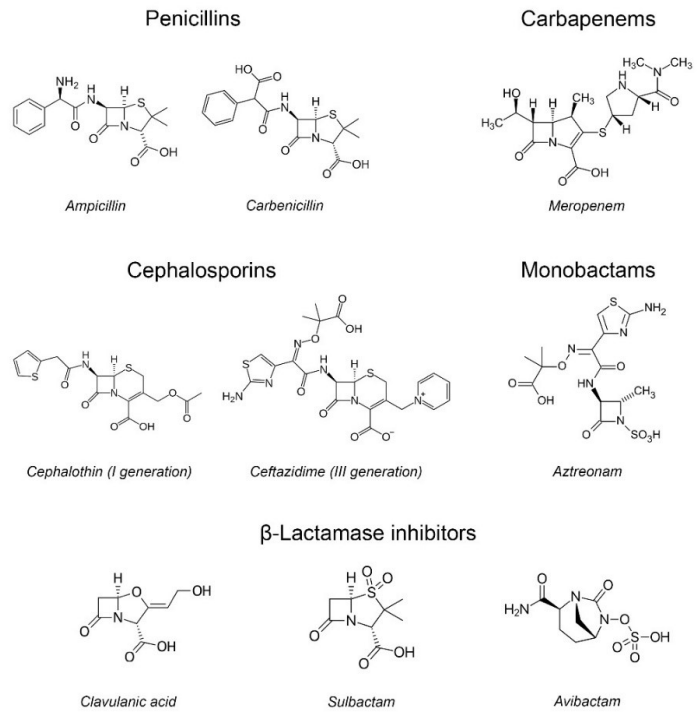

Figure 1.3. Structures of the main groups of β-lactam antibiotics with examples and structures of β-lactamase inhibitors. Clavulanic acid and sulbactam are β-lactam β-lactamase inhibitors and avibactam is a non-β-lactam β-lactamase inhibitor. All shown compounds except for aztreonam were used in this study.

Thus, the resistance against β-lactam antibiotics and β-lactamase inhibitors can routinely occur via mutations in β-lactamases. In such conditions the design of new drugs might be only a temporary solution. However, most of the mutations clinically found in resistant variants were in non-essential amino acids. So, what if we take advantage of the fact that mutations in certain essential residues cannot take place easily? Will it be possible to design an inhibitor to which resistance cannot evolve because it would require a change in an essential residue? This is an exciting idea that can be brought to life only with the comprehensive study of such essential residues and the evolutionary pathways an enzyme has taken and can take.

β-Lactamases, BlaC

β-Lactamases are the enzymes that are able to inactivate β-lactam antibiotics. They were first detected in 1940[86], interestingly already before the wide use of β-lactam antibiotics in the clinic. The introduction of β-lactam compounds as therapy intensified the process of β-lactamase evolution, and nowadays the β-lactamase family counts more than a thousand unique enzymes with various substrate profiles[87]. β-Lactamases are present in the periplasmic space of Gram-negative bacteria. In Gram-positive bacteria they are found to be bound to the cytoplasmic membrane or excreted[88].

All β-lactamases are generally classified by two main classification systems. β-Lactamases can be divided in four classes based on their primary sequence homology (Ambler classification)[89]. Classes A, C and D contain serine-proteases, which use a serine residue in the active site[90,91]. The Ser70 hydroxyl group is involved in the nucleophilic attack on the carbonyl carbon atom in the β-lactam ring[92]. Class B enzymes are metallo-β-lactamases and require one or two zinc ion(s) in the active site[93,94]. Despite having low sequence identity, different classes still have a similar fold, with two domains, an α-domain and an α/β-domain (Figure 1.4a and 1.4b) or 2 α/β-domains (class B). β-Lactamases also share structural similarity with penicillin-binding proteins (PBPs), transpeptidases that are responsible for the final stages of cell-wall synthesis in bacteria[95,96]. They are the main target of β-lactam antibiotics and are believed to be the ancestors of β-lactamases[97].

The Bush-Jacoby classification uses substrate profiles to distinguish three or four groups. Group 1 consists of cephalosporinases and matches class C β-lactamases. Group 2 is defined by its susceptibility to inhibition by clavulanic acid, and includes β-lactamases from molecular classes A and D. It is further divided in subgroups based on the preferred substrate. Group 3 contains class B metallo-β-lactamases with various substrate specificity (mostly carbapenemases) and group 4 includes penicillinases which are not prone to clavulanate inhibition. However, this group sometimes is not included in classification due to unknown enzyme structure[98–100].

The β-lactamase from *M. tuberculosis*, BlaC, is a broad-spectrum class A β-lactam degrading enzyme that is encoded by a chromosomal gene[101]. Although known for years, it was first purified and characterized only in 1998[71] and the first crystal structure of it was obtained in 2006[101]. BlaC is classified in Bush-Jacoby group 2 and belongs to the extended spectrum β-lactamases (ESBL)[102], as it was shown to degrade penicillins, cephalosporins and several carbapenems[101,103], although less effectively than other β-lactamases.

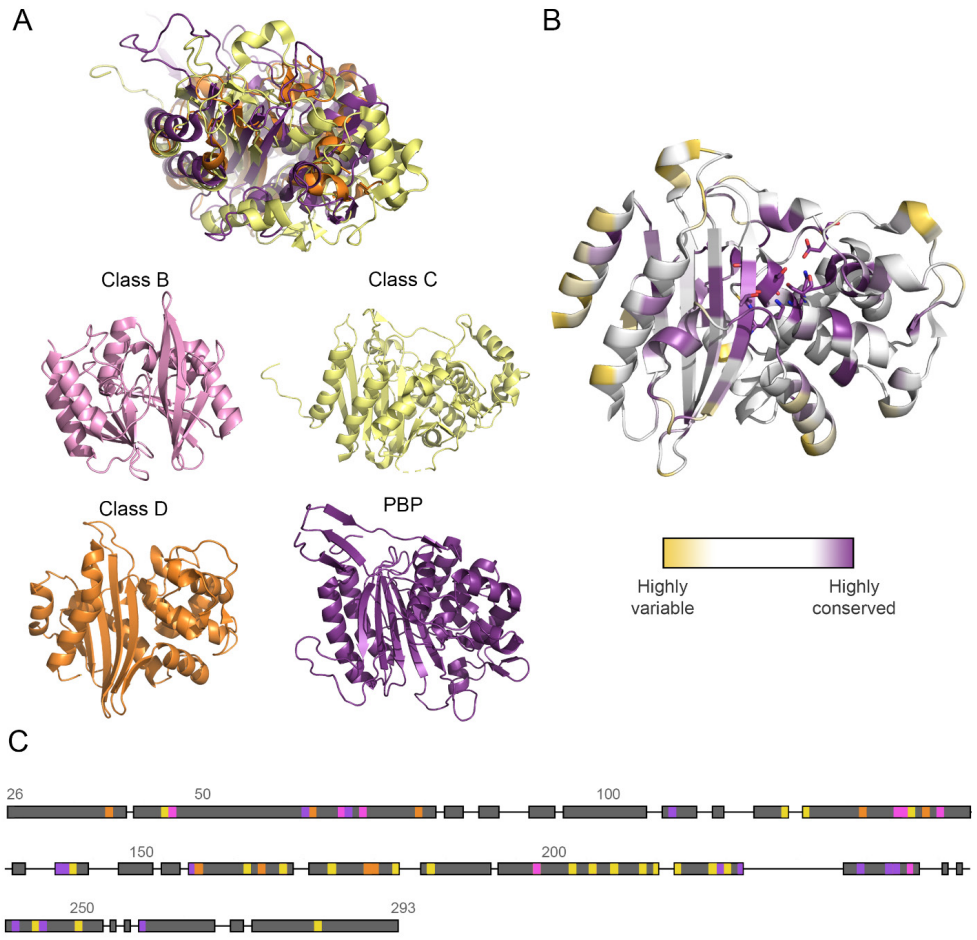

Figure 1.4. (A) Examples of classes B (4D1T[104]), C (5K1D[105]), D (1K57[106]) β-lactamases and PBP (2Z2L[107]). At the top the overlay of class C, D and PBP structures is shown; (B) BlaC (class A, 2GDN[101]) colored by conservation of amino acid residues. Active site residues are shown in sticks; (C) Consensus sequence of class A β-lactamases with Ambler numbering. Residues that are 100%, >99%, >95% and >90% conserved are shown in pink, orange, purple and yellow, respectively, thin lines show the gaps in consensus sequence.

Conserved residues of β-lactamases

Class A β-lactamases are the largest group of β-lactamases, and they have been extensively studied structurally and mechanistically. A sequence alignment of class A β-lactamases done by a ConSurf server[108,109] indicate that around 15% of the amino acids are identical in >90% of the sequences (Figure 1.4b and 1.4c, more on conserved residues can be found in Chapter 2 of this work). The conserved residues are concentrated in and around the active site but are also found throughout the enzyme structure.

Active site

The process of β-lactam antibiotic hydrolysis starts with substrate binding in the active site, which is followed by nucleophilic attack on the β-lactam carbonyl group (acylation step) (Figure 1.5a). The protonation of the β-lactam nitrogen results in formation of an acyl-enzyme intermediate, which is then hydrolyzed due to a second nucleophilic attack by an activated water molecule, to yield the resting state enzyme and inactivated antibiotic molecule (deacylation step). In 1992, it was shown by Strynadka and colleagues that the acyl-enzyme intermediate is formed by a substrate covalently bound to Ser70 in a deacylation-defective β-lactamase mutant. Thus, Ser70 was recognized as the attacking nucleophile in the acylation step[110]. The authors also proposed a role for Lys73 as the general base in this process. However, still today debate is ongoing over the involvement of Glu166, the active site water, and Ser130 in Ser70 activation[79,111]. In 1993, a study by Knox *et al.*[112] proposed a role of Glu166 in deacylation step but not in acylation step, but later other research groups showed the possibility of Glu166 together with a water molecule to act as the general base[113,114]. It cannot be excluded that the hydrogen transfer pathways differ among different β-lactamases.

Hence, even though the exact mechanisms of proton shuffling, and the initial activation remain controversial, the composition of an active site in class A β-lactamases is well established. The first crystal structure of PC1 β-lactamase derived from *Staphylococcus aureus* solved in 1987 (refined later in 1991 to 2 Å) by Osnat Herzberg allowed to identify the position of the active site residues[115,116]. The active site of class A β-lactamases consists of Ser70, Lys73, Ser130, Glu166 as the catalytic residues and several residues involved in substrate binding (Figure 1.5b). The conservation of the catalytic residues is extreme, residues 70, 73 and 130 are invariable, and residue 166 is almost invariable with only a few natural substitutions occurring (for example, in some *Nocardiopsis* species, which were isolated from soil in China[117], but no crystal structure or substrate information is available for those enzymes). Mutational studies on the catalytic residues allowed for capturing various adducts inside the active site, making the catalysis process clearer. The conservation of the substrate-interacting residues, however, is lower, even though it is still extremely high for some of them. This room for substitutions is likely important for adaptability of the active site for new substrates. For

example, the residue at position 234 is generally lysine and was shown to act as an anchor for the C3 carboxylate of penicillins[118]. Arginine 234 is found in most carbenicillinases from class A and is experimentally associated with carbenicillinase activity and decreased clavulanate sensitivity[83,120-123]. Another example of variability in the active site residues is residue 105, which canonically is an aromatic residue (tryptophan or phenylalanine), while in BlaC it is isoleucine. This residue hangs over the active site pocket and regulates the accessibility of the active site for substrates. Due to its small size in BlaC, this "gatekeeper" residue results in a wider active site entrance, which may be one of the reasons for the broader substrate range of BlaC[80].

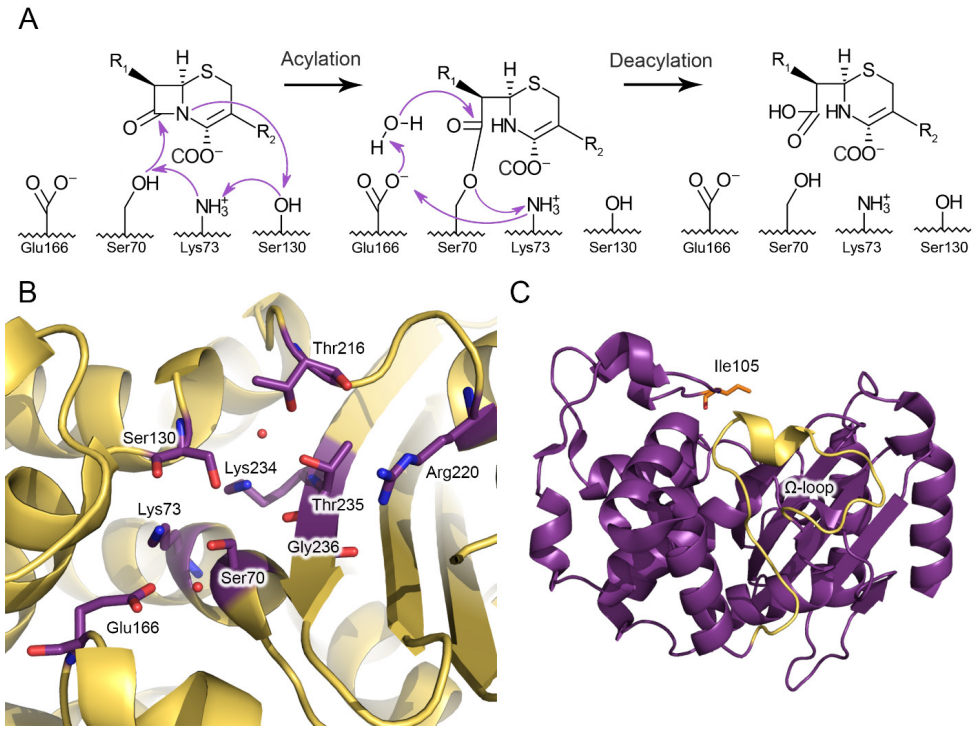

Figure 1.5. (A) Schematic representation of the proposed catalytic mechanism of class A β-lactamases; (B) Active site of BlaC (2GDN) with important residues shown in sticks and waters as red spheres; (C) BlaC structure with the gatekeeper residue 105 in orange sticks and the Ω-loop (in yellow).

Around active site

The group of residues around the active site also received some attention, as they directly influence the position or orientation of active site residues. Extensive research was done on the "gatekeeper" loop (residues 103-106), which includes "gatekeeper" residue 105 (Figure 1.5c). It was shown, for example, for residue Asn106 in CTX-M β-lactamase, that mutation leads to an increased thermostability of the protein, however, it dramatically decreases activity against cefotaxime, due to altered positions of active site resiudes[123]. The highly conserved residue Asn136 was demonstrated to have a crucial effect on activity upon mutation, as mutations lead to a displacement of a catalytical residue, Glu166[124]. Several studies showed the importance of the invariable residue Asp131 for folding and formation of a stable protein[125,126]. Another key element of β-lactamase, the Ω-loop (Figure 1.5c) was researched thoroughly as well. The Ω-loop contains the catalytic residue Glu166 and was demonstrated to be important for substrate recognition. One of the Ω-loop residues, Asn179, is highly conserved, and its mutation altered the substrate profile in TEM and KPC-2[127–130]. In many β-lactamases Asn179 makes an interaction with Arg164 and the disruption of this interaction is believed to contribute to a changed substrate profile.

Far from active site

Conserved residues far from the active site have received limited attention, probably because they do not contribute to a possible novel activity and therefore are not of great interest. It was shown that the conserved Trp229 might be involved in the modulation of an allosteric site[131,132]. Meneksedag *et al.* showed that β-lactamase inhibitor protein caused changes in the flexibility of TEM-1 regions away from the main binding site, thus pointing toward the binding to the alternative site. Such an alternative binding site contained Trp229 and the authors discussed the importance of Trp229 for the modulation of the allosteric communication between the main and alternative binding sites[132]. A similar conclusion was made in another study on TEM-1, where the importance of Trp229 for regulation of an allosteric site was discussed together with Pro226, another highly conserved residue[131].

Research question and thesis overview

The aim of this research was to better understand evolutionary processes in enzymes, specifically in β-lactamases, in terms of the patterns of residue conservation. Under the assumption that enzymes under heavy and varying selection pressures, such as is the case for β-lactamases, will minimize the number of essential residues to increase evolutionary robustness and evolvability, the question of which critical roles the remaining essential residues play, is highly relevant. Understanding these roles in relation to the location in the structure (first, second and third shell) can provide insight into the relationship between structure, dynamics, activity and evolution of β-lactamases and enzymes in general. While most research is focused on catalytic residues, the conserved residues of second and third shells also convey an interesting message. We have formulated possible roles of such conserved residues and attempted to test these hypotheses with extensive mutational studies to serve a basis for understanding molecular mechanisms involved in enzyme evolution.

Chapter 1 of this thesis provides relevant background information in the fields of protein evolution and β-lactamases, as well as an overview of questions discussed in this work. A large-scale mutational study, including *in vivo* and *in vitro* characterization is described in *chapter 2*, where the roles of all second-shell and third-shell conserved residues are discussed. It is shown that second-shell residues contribute to the arrangement of the active site residues, while residues from the third shell are mostly responsible for the folding process and stability. Interesting observations on the structural effects of second-shell mutations are described and various roles in the folding process are proposed. The general observations and patterns presented in this chapter are supported with in-depth discussions of specific mechanisms of function for several second- and third-shell residues in the next chapters. *Chapter 3* gives an insight into the exact roles of several second-shell residues. It is proposed that mutations in these residues create a mobile part displacing the substrate-binding residues. Moreover, this chapter demonstrates co-evolution of residues because comparison of the sequences reveals a surprising pattern in the position of a substrate binding residue. In *chapter 4* a thorough discussion of several third-shell residues is presented. Two residues very far from active site are shown to be involved in different aspects of the folding of β-lactamase. *Chapter 5* touches on some characteristics of residues that exhibit unexpected results upon mutation. While it is expected that conserved residues are essential, for some of them mutations are not only non-damaging but, in some cases, even advantageous. In *chapter 6*, crystal structures of a variety of BlaC mutants which were found in directed evolution studies are discussed. This chapter reveals possible changes that can occur in a protein upon gaining an improved phenotype. *Chapter 7* concludes this work with a general discussion and an outlook on further research.