Chapter 6

Summary, general discussion and perspectives
Summary

This thesis deals with the directed evolution of a cephalosporin acylase. Acylases play an important role in the environmentally friendly production of semi-synthetic β-lactam antibiotics, which are the largest group of antibiotics used worldwide. Unfortunately, resistance towards this important class of antibiotics is still rising, stressing the need for new variants of these and other antibiotics. Since the specificity of antibiotics is mainly due to their side chains, semi-synthetic antibiotics with new side chains and thereby new enzymes to produce these compounds are needed. New enzymes can be made by mimicking the evolution in the laboratory while directing it towards the desired activity. This directed evolution technique has resulted in many improved model enzymes in the last decade and promised to be useful for industrial biocatalysts as well. We decided to pursue directed evolution on a industrially applicable enzyme, the cephalosporin acylase of Pseudomonas SY-77. A general introduction on this enzyme and directed evolution is given in chapter 1 of this thesis.

The SY-77 glutaryl acylase is transcribed as a polypeptide which folds and cleaves itself into a heterodimer. This cleavage results in the formation of the N-terminal serine of the β-subunit that is the active site residue performing the nucleophilic attack of the peptide bond of the substrate. Although a lot of different enzymes belong to the so-called cephalosporin acylase family, their activity towards cephalosporin C (CPC) is very low ($k_{cat}$ of $< 0.15 \text{ s}^{-1}$) in comparison to the preferred β-lactam substrate glutaryl-7-aminocephalosporanic acid (glutaryl-7-ACA). The key intermediates in the production of semi-synthetic cephalosporins are 7-ACA and 7-aminodesacetoxy-cephalosporanic acid (7-ADCA). In the two-step process from CPC to 7-ACA, a D-amino acid oxidase removes the amino group from the side chain, followed by the hydrolysis of the glutaryl side chain by a cephalosporin acylase. In a new production route of 7-ADCA a genetically modified fungus produces adipyl-7-ADCA instead of CPC. Optimisation of a single step production process of 7-ADCA would require a cephalosporin acylase that hydrolyses this compound, adipyl-7-ADCA. Both examples clearly demonstrate that there is need for new cephalosporin acylases which are able to hydrolyse adipyl-7-ADCA or CPC at a high rate.

Directed evolution is a fairly new area of protein engineering. It follows the path of Mother Nature to find new and improved enzyme activities. In nature new enzymes and enzyme activities have evolved by mutation and recombination. In the laboratory, researchers try to direct enzymes towards new activities by mimicking this process, followed by selection for the desired activity. In this process two parts are crucial for the success of an experiment, the quality of the mutant library and the method to search for the best mutant. The perfect mutant library comprises all possible mutants in an equal amount. Several research groups have been working towards this goal in the last decade, resulting in numerous protocols for mutagenesis and recombination, each having their
own advantages and drawbacks\cite{55,64}. The second crucial part of a directed evolution experiment involves the method of finding the best mutant. This can be done by selection or screening. Selection is a process in which the desired activity is coupled to a selectable marker. Since only active enzymes will be sampled, a much larger pool of mutants can be searched. For activities which cannot be coupled to a selectable marker screening is the method of choice. In screening procedures all mutants have to be tested for the desired activity, which decreases the number of searchable mutants. Fortunately, high throughput screening assays have been established for many different enzymes, resulting in the possibility of screening larger libraries.

At the time the work for this thesis was started, the first successful examples of directed evolution were described\cite{125}. These examples, however, were mainly dealing with the exploration of the new techniques. Therefore, model systems were used handling enzymes with easy detectable activities, like improving antibiotic resistance or thermostability of an enzyme, or using substrates that are converted into coloured products. Although these models are very well suited to improve directed evolution techniques, most of them do not lead to useful biocatalysts. Therefore, the aim of this thesis is to use directed evolution in order to find variants of the industrially relevant glutaryl acylase of \textit{Pseudomonas SY-77} with improved hydrolysis towards specific β-lactam compounds.

\textbf{Chapter 2} describes a random mutagenesis approach aimed at the finding of amino acids, important for substrate specificity. As the α-subunit of the enzyme was already mutagenised before, we decided to focus on the β-subunit. This part of the gene consists of 1566 bps. Mutagenising this stretch of DNA will result in a library too large to select. Therefore, we divided the gene into five highly overlapping parts of about 550 bps and mutagenised every part separately by epPCR, adjusting the Mn$^{2+}$ concentration in such a way that every mutant comprises one or two mutations on average. The mutagenised parts were cloned back into the rest of the gene and transformed to the leucine auxotroph \textit{Escherichia coli} DH10B. The transformants were plated on minimal medium containing adipyl-leucine as sole leucine source. In this way, only bacteria expressing a mutant enzyme capable of hydrolysing adipyl-leucine to liberate leucine will be able to grow. After two weeks several colonies were picked from the selective plates and crude cell extracts were used to determine the hydrolysis of adipyl-7-ADCA, in order to see if transformants were not only active towards the selection substrate adipyl-leucine, but also towards the desired β-lactam compound. Hydrolysis of glutaryl-7-ACA was also determined and the ratio was calculated to correct for differences in enzyme concentration in the crude cell extracts. Sequencing of the selected mutants showed that most of them comprised a mutation at position 266 or 375. Five mutants had an improved ratio of hydrolysis of adipyl-7-ADCA over glutaryl-7-ACA, in order to see if transformants were not only active towards the selection substrate adipyl-leucine, but also towards the desired β-lactam compound. Hydrolysis of glutaryl-7-ACA was also determined and the ratio was calculated to correct for differences in enzyme concentration in the crude cell extracts. Sequencing of the selected mutants showed that most of them comprised a mutation at position 266 or 375. Five mutants had an improved ratio of hydrolysis of adipyl-7-ADCA over glutaryl-7-ACA. These enzymes were purified and catalytic parameters were determined. Mutant SY-77$^{N266H}$ showed an almost ten times improved catalytic efficiency towards adipyl-7-ADCA, while the improvement of the ratio of mutation F375L was merely due to a reduction of the
hydrolytic activity towards glutaryl-7-ACA. The structure of the highly homologous cephalosporin acylase of *Pseudomonas diminuta* KAC-1 revealed that both Asn266 and Phe375 are part of the active site of the enzyme.

In epPCR it is, however, unlikely that two bases within the same codon are mutated. This implies that on average one amino acid may be changed in seven other amino acids, being mainly conservative mutations. This means that the method is highly suited to point out residues that are important for the desired activity, but only a limited part of the sequence space can be explored. Chapter 3 and 4 describe a complete mutational analysis of positions 266 and 375, respectively, in order to explore the function of these residues in substrate specificity in depth. In each case, all 20 mutant proteins were purified and their catalytic parameters were determined. Most mutant enzymes could be fully functionally expressed in *E. coli* and they displayed a broad spectrum of affinities and activities towards glutaryl-7-ACA, adipyl-7-ADCA and CPC, confirming the importance of these residues for substrate specificity and pointing towards the flexibility of the enzyme at these positions. The mutants SY-77N266M and SY-77F375C appeared to show the most improved activity towards adipyl-7-ADCA, while SY-77N266Q showed a small increase towards cephalosporin C. It should be noted that each of these mutations could only be achieved by a change of two bases in the respective codons, clearly demonstrating how this saturation mutagenesis approach complements the epPCR.

Although up to a 15-fold increase of the catalytic efficiency for hydrolysis of adipyl-7-ADCA was shown in single mutants, it appeared likely that further improvement of the activity on this substrate and activity towards CPC is possible by creating multiple mutations. In chapter 5 both rational and random approaches are described to find the best combination of mutations in the most important substrate specificity residues. In a rational approach of combining the best mutant in the α-subunit with the most improved mutants found in the β-subunit, only the combination of mutations Y178H and F375L showed an improved hydrolysis activity towards adipyl-7-ADCA compared to both single mutants. Therefore, we approached the combination of two residues in a more random way. The 20 amino acids at position 266 were combined with all 20 amino acids at position 375. This library of β-subunit mutants was cloned into a wild type background and into a plasmid bearing the Y178H mutation in the α-subunit. The resulting libraries were transformed to a leucine deficient *E. coli* strain and selected by growing them on minimal medium with adipyl-leucine or amino adipyl-leucine as sole leucine source. Unfortunately, none of the three transformants picked from the amino adipyl selection plates showed enzyme production. Some transformants from the adipyl selection plates did show enzyme production and were tested for activity towards glutaryl-7-ACA and adipyl-7-ADCA. Four mutants showed improved activity towards the desired substrate, so they were purified and catalytic parameters were determined. Upon sequence analysis it was discovered that the best multiple mutant was a combination of the best single mutants from earlier research (SY-77Y178H and
Summary & general discussion

SY-77\textsuperscript{N266M}) resulting in an additive activity towards adipyl-7-ADCA. With respect to the acylase activity towards CPC, none of the four selected mutant combinations, N266Q+F375L, Y178H+N266H, Y178H+N266M and Y178H+N266Q+F375M, resulted in a higher conversion of this substrate relative to the single mutant SY-77\textsuperscript{Y178H}. In the search for an even better acylase we started with the generation of a library in which all important residues for substrate specificity discovered until now are totally randomised. Since it was obvious from former research that using oligonucleotides for saturation mutagenesis may result in a biased library towards the original codon, we decided to construct the mutated part of the gene synthetically. The four amino acids Tyr178, Tyr231, Arg 255 and Asn266 were totally randomised using NNS codons on both strands. Although the assembly of ten pairs of oligonucleotides is very easy in theory, the process had to be monitored very closely in order to achieve a library of 3*10\textsuperscript{6} mutants, which statistically comprise 94\% of all possible mutants. The cloning of this library, however, was not very successful resulting in only 0.13\% of the total possible mutant library of all five residues randomised. Plating this small library onto minimal medium with amino adipyl-leucine as sole leucine source did not result in the selection of any active mutant. Although it is clear that this experiment has to be repeated with a complete library, the results will be part of research that goes beyond this thesis.

General discussion

Directed evolution started as a poor man’s protein engineering tool to improve enzymes from which no crystal structures were available. Over the last decades, however, several impressive successes were achieved, turning it into a major protein engineering tool next to rational design. After years of competing, it is clear now that both methods do not exclude each other, but have compatible strengths. Therefore, the two methods of rational design and directed evolution should no longer be used separately, but are to be combined in order to search a larger sequence space\textsuperscript{126}.

In this thesis both random mutagenesis strategies and more rational methods have been used to develop cephalosporin acylases with improved activity towards adipyl-7-ADCA and CPC. As is clear from chapters 2, 3 and 4, the combination of a directed evolution experiment, comprising of a random mutagenesis strategy followed by selection to find the residues that are important for substrate specificity, with a more detailed saturation mutagenesis of the important residues, to deeply explore sequence space, is a good strategy. However, a reliable and fast screening assay as well as an efficient protein purification method are equally important to grant success. Mutants were found with improved catalytic efficiency up to 15 times towards adipyl-7-ADCA and up to a 3-fold improved hydrolysis rate of CPC. Although this does not seem to be a very large improvement, the mutant showing the highest activity towards
adipyl-7-ADCA displays a $k_{cat}$ that is only two times lower than the $k_{cat}$ of the wild type enzyme towards glutaryl-7-ACA. So, this mutant enzyme is almost industrially applicable. The enzyme is, however, already a very efficient biocatalyst, which means that an improved activity of five to ten times already is a significant achievement from a scientific point of view. Furthermore, every activity towards CPC is a big achievement, since several decades of screening for this activity did not result in any Cephalosporin C acylase yet [12,13].

The constructed library with all five important substrate specificities randomised was apparently not large enough to permit isolation of even more improved enzymes. It can be stated beyond doubt that several mutations are needed in order to change substrate specificity of an enzyme. As an example, in parallel to the research described in this thesis, Oh et al. generated several mutant libraries of the cephalosporin acylase of *P. diminuta* [24]. Guided by the crystal structure, important substrate specificity residue were mutagenised by site directed mutagenesis one at the time. The best mutant was used as a backbone for the total randomisation of a second amino acid and the best mutant out of this library was used for the third round. It was shown that the best single, double and triple mutant had an improved hydrolysis activity towards CPC of respectively 1.8, 6 and 7.9 times the wild type activity. This corroborates with the hypothesis that multiple mutants will improve the activity towards the desired substrate. The combination of new mutants with the best mutant of a former library may, however, not always result in an improvement of activity. The rational combination of two improved mutants in the $\alpha$- and $\beta$-subunit did not result in an improved catalytic efficiency of the multiple mutant in comparison to the single mutants (chapter 5). The mutants selected on adipyl-leucine displayed an improved activity towards adipyl-7-ADCA. However, in none of these multiple mutants the activity towards CPC was increased compared to single mutant SY-77 Y178H. This suggests that starting from a fixed mutant backbone may result in the entrapment at a local minimum in the enzyme activity landscape. It is therefore important in protein engineering strategies to keep some randomness upon combination of mutants [127].

Along with the strategy to construct the library, the selection or screening method is a very important part of a directed evolution experiment. The selection method used in this thesis is an elegant method in which the desired enzyme activity is used for the survival of the expression host, resulting in several improved mutants as described throughout the thesis. The advantage of the method is that the selective pressure is on enzyme activity rather than binding of the enzyme to the desired substrate as is described in many other directed evolution experiments [72]. Furthermore, it is a straightforward and technically simple method, since it is an *in vivo* selection, and applicable to numerous peptide cleaving enzymes. Although this method has been used before to change penicillin acylases into cephalosporin acylases [68,128], it only resulted in mutants with improved activity towards the selection substrates and not towards the similar $\beta$-lactam compounds. This confirms the first law of directed evolution “you get
what you select/screen for” [69]. However, in our case the mutants selected on adipyl-leucine did show activity towards adipyl-7-ADCA. Perhaps cephalosporin acylases bind differently to the core of the substrate than penicillin acylases. Alternatively, this selection strategy may only be applicable for small changes in enzyme activity and the change from an aromatic penicillin to an aliphatic cephalosporin side chain may be too big a step.

This last hypothesis may also explain why we did not find any active mutants on the selection plate containing amino adipyl-leucine. On the other hand, the selection procedure may also be limiting. For example, the single mutant SY-77Y178H, when plated separately, showed small colonies on amino adipyl-leucine selection plates within nine days, but was not selected in the YNF library (chapter 5). Since the randomness of the library was ensured, this strongly suggests that the selection method is not appropriate to find mutants that show only three times improved hydrolysis activity in vitro towards a high concentration of CPC. First of all, the visual identification of small white colonies on the white amino adipyl-leucine selection plates is difficult, but can be trained. Furthermore, the concentration of selection substrate and thereby the leucine concentration in the plate may be limiting. Mutants that are impaired in the expression of acylase have the advantage of producing less protein. So they need less substrate and may grow faster and be selected. In order to improve the selection procedure coloured plates may be used to improve distinguishing a colony from the plate. Another suggestion for improvement of the assay is using a higher concentration of selection substrate in order to minimise the selection pressure on the affinity of the mutants towards the substrate.

The second step in our selection procedure is the ratio of hydrolysis of adipyl-7-ADCA over that of glutaryl-7-ACA. As shown in chapter 2, an improved ratio may be the result of an improvement of activity towards adipyl-7-ADCA or of the decreased hydrolysis of glutaryl-7-ACA. Since the selection procedure is not as unambiguous as we would like it to be and we want to improve enzymes rather towards β-lactam compounds than towards the selection substrates, we do need a second screening step. Purifying all of the selected mutant enzymes is, however, not possible when large libraries are selected, since this process takes several days and only two mutants can be purified per day. Using a method to quantify the amount of active enzyme in the crude cell extract would be a reliable alternative. Unfortunately, all known inhibitors for acylases and peptidases do not affect cephalosporin acylases, which means that an inhibition reaction can not be used for the quantification of the amount of active enzyme. Furthermore, the polyclonal SY-77 antibody, which is used for detection of acylase, appeared to be not suitable for quantification. The absolute values of the crude cell extracts seem, however, to have a predictive value for the hydrolysis activity of the purified mutant. So, both ratio and values should be considered in this second step. Another medium to high throughput screening method would be the use of BIAcore to detect the amount of enzyme binding to the substrate coated on a chip. Although this is an elegant and quantitative method it is a fairly new method and may be difficult to establish.
The last step in the characterisation of the mutant enzymes is the determination of the kinetic parameters by using a range of different substrate concentrations in a discontinuous assay using fluorescamine. The validation of this strategy and the use of the appropriate controls resulted in reproducible kinetic parameters for both adipyl-7-ADCA en glutaryl-7-ACA. However, some aspects should be taken into account. Since the assay is based on the production of primary amines by cephalosporin acylases, it is also prone to background fluorescence of other primary amines. Although the chosen wavelength prevents the detection of primary amines on amino acids (and thus proteins), the presence of the amine group on the side chain of CPC does interfere. This results in a much higher background value and leaves very little space to see small improvements of activity towards the substrate. It is therefore crucial to perform at least two different experiments in triplicate or more for each mutant. As can be seen in chapters 3 and 5, this did generate reproducible hydrolysis rates of CPC.

Although the main theme of this thesis is about the mutagenesis of active site residues and most successful protein engineering strategies aim at these amino acids, there are papers showing that small modifications of an enzyme outside its active site also result in improved enzyme activity. These mutations are usually difficult to explain, but are mostly thought to stabilise conformational changes in the enzyme, which are the result of another transition state with the new substrate or of modified hydrophobic and hydrophilic interactions with the desired solvent. The epPCR of the five parts of the enzyme only resulted in one improved mutant outside the active site, although this residue might be important for the binding of the core of the substrate (chapter 2). Since one mutation, either inside or outside of the active site, might not be enough to improve enzyme activity and structurally close residues are not always closely together in the gene, it would be a good idea to shuffle all five epPCR parts in order to get multiple mutations throughout the whole gene. In this way, mutations that do not improve activity on their own may be discovered.

Another way of obtaining more diversity in an enzyme is by shuffling it with other homologous enzymes. Family shuffling of acylases seems not feasible, since they are mainly homologous in a structural way and not at the amino acid level. Some parts of the enzyme do have more homology to each other than other parts. This might just be enough for shuffling. On the other hand, comparing the differences between several acylases in combination with their substrate specificity might lead to the discovery of other important substrate specificity residues. This approach needs, however, several crystal structures or models of the enzymes to be compared and an extensive calculation capacity. The crystallisation of natural enzymes as well as (multiple) mutants with or without a co-crystallised substrate can increase our knowledge on structure-function relationship and should therefore be a main goal for the coming years.

The cephalosporin acylase of Pseudomonas SY-77 appears to be quite amenable to modification of enzymatic activity. A single mutation improved the hydrolytic activity
towards adipyl-7-ADCA already 15-fold, which could be easily selected with a straightforward selection procedure. This may be attributable to the fact that substrate specificity is mainly associated with the side chain and not with the core of the substrate resulting in a broad substrate range. Cephalosporin acylase also shows enantioselective amidase and esterase activity [33]. As cephalosporin acylases are used in the different industrial processes to obtain 7-ACA or 7-ADCA, conditions have already been developed and optimised for its production and quality assurance. Finding another enzyme activity in a mutant cephalosporin acylase only requires the (re)design of a (new) selection or screening strategy, giving it a plethora of industrially interesting and important functions to be explored. This makes the enzyme a perfect adaptable biocatalyst, not only for the production of semi-synthetic cephalosporins, but also in the resolution of racemic mixtures and potentially in the synthesis of peptides [10].

In conclusion, the evolution of a glutaryl acylase into an industrial cephalosporin acylase is most definitely feasible. The foundation has been laid, the house can be build.