Towards Understanding of Selectivity & Enantioconvergence of an Epoxide Hydrolase

ÅSA JANFALK CARLSSON
Abstract

Epoxide hydrolase I from Solanum tuberosum (StEH1) and isolated variants thereof has been studied for mapping structure-function relationships with the ultimate goal of being able to in silico predict modifications needed for a certain activity or selectivity. To solve this, directed evolution using CASTing and an ISM approach was applied to improve selectivity towards either of the enantiomeric product diols from (2,3-epoxypropyl)benzene (1).

A set of variants showing a range of activities and selectivities was isolated and characterized to show that both enantio- and regioselectivity was changed thus the enrichment in product purity was not solely due to kinetic resolution but also enantioconvergence. Chosen library residues do also influence selectivity and activity for other structurally similar epoxides styrene oxide (2), trans-2-methyl styrene oxide (3) and trans-stilbene oxide (5), despite these not being selected for.

The isolated hits were used to study varying selectivity and activity with different epoxides. The complex kinetic behaviour observed was combined with X-ray crystallization and QM/MM studies, powerful tools in trying to explain structure-function relationships. Crystal structures were solved for all isolated variants adding accuracy to the EVB calculations and the theoretical models did successfully reproduce experimental data for activities and selectivities in most cases for 2 and 5. Major findings from calculations were that regioselectivity is not always determined in the alkylation step and for smaller and more flexible epoxides additional binding modes are possible, complicating predictions and the reaction scheme further. Involved residues for the catalytic mechanism were confirmed and a highly conserved histidine was found to have major influence on activity thus suggesting an expansion of the catalytic triad to also include H104.

Docking of 1 into the active site of the solved crystal structures was performed in an attempt to rationalize regioselectivity from binding. This was indeed successful and an additional binding mode was identified, involving F33 and F189, both residues targeted for engineering.

For biocatalytic purpose the enzyme were was successfully immobilized on alumina oxide membranes to function in a two-step biocatalytic reaction with immobilized alcoholdehydrogenase A from Rhodococcus ruber, producing 2-hydroxyacetophenone from racemic 2.

Keywords: Epoxide hydrolase, Epoxide, Enantioselectivity, Regioselectivity, Enantioconvergence, Crystal structures, Biocatalysis, Immobilization, Transient kinetics, CASTing, Directed evolution

Åsa Janfalk Carlsson, Department of Chemistry - BMC, Box 576, Uppsala University, SE-75123 Uppsala, Sweden.

© Åsa Janfalk Carlsson 2016

ISSN 1651-6214
urn:nbn:se:uu:diva-286557 (http://urn.kb.se/resolve?urn=urn:nbn:se:uu:diva-286557)
Ju mer man tänker, ju mer inser man att det inte finns något enkelt svar

Nalle Puh
List of Papers

This thesis is based on the following papers, which are referred to in the text by their Roman numerals.

I \hspace{1em} Janfalk Carlsson, Å., Bauer, P., Ma, H. and Widersten M. Obtaining optical purity in product diols in enzyme catalyzed hydrolysis: contributions from changes in both enantio- and regioselectivity. Biochemistry 2012 vol. 51 7627-7637


IV Janfalk Carlsson, Å., Bauer, P., Nilsson, M., Dobritzsch, D. and Widersten, M. Laboratory evolved enzymes provide snapshots of the development of enantioconvergence in enzyme-catalyzed epoxide hydrolysis. Submitted

V Janfalk Carlsson, Å., Bauer, Paul., Kamerlin, S. C. L. and Widersten, M. Complex Kinetic Schemes are Required to Describe Epoxide Hydrolase Catalyzed Production of Diols. Manuscript

VI Billinger, E., † Janfalk Carlsson, Å., † Widersten, M. and Johansson, G. Integrated action of Solanum tuberosum epoxide hydrolase I and Rhodococcus ruber alcohol dehydrogenase A in a two-step biocatalytic reactor. Manuscript

† These authors contributed equally.
Paper I reprinted with permission from *Biochemistry*
http://pubs.acs.org/doi/abs/10.1021/bi3007725 and for addition/corrections
http://pubs.acs.org/doi/abs/10.1021/acs.biochem.6b00197.

Paper II reprinted with permission from *ACS Catalysis*

Paper III reprinted with permission from the *Royal Society of Chemistry*
http://pubs.rsc.org/en/content/articlelanding/2016/ob/c6ob00060f#!divAbstract
Contribution report

I. Performed most of the experimental work, contributed to the writing of the manuscript.

II. Constructed E35Q/H300N and performed kinetic studies for this variant, wrote the corresponding parts for the manuscript.

III. Performed laboratory work to obtain crystals for R-C1B1 and parts of the kinetics, contributed to the structuring of the manuscript and wrote the corresponding parts of the work.

IV. Performed laboratory work to obtain crystals for all variants presented, contributed to the writing of the manuscript.

V. Performed all experimental work, contributed to the writing of the manuscript.

VI. Performed part of the experimental work, contributed to the writing of the manuscript.
Abbreviations

ADH-A Alcohol dehydrogenase A from *Rhodococcus ruber*
CASTing Combinatorial active site testing
DNA Deoxyribonucleic acid
*E. coli* *Escherichia coli*
EH Epoxide hydrolase
ESI/MS Electrospray ionization mass spectrometry
HPLC High-performance liquid chromatography
ISM Iterative saturation mutagenesis
LDH L-lactate dehydrogenase
NADH Nicotinamide adenine dinucleotide
NOX NADH oxidase from *Lactobacillus sanfranciscencis*
PDB Protein data bank
QM/MM Quantum mechanics/molecular mechanics
StEH1 Epoxide hydrolase 1 from *Solanum tuberosum*
*
Indicates a chiral carbon
Introduction to my work

Welcome to join in on my journey towards understanding of the structure-function relationship of epoxide hydrolase I from Solanum tuberosum. This enzyme has previously been thoroughly studied and characterized, with a well-understood reaction mechanism and solved crystal structure, thus fulfilling many of the criteria wished for when setting out for a task like this. The aim for this study was to construct and characterize modified enzyme variants with improved enantio- or regioselectivity. Together with theoretical chemists and crystallographers we took quite a few steps towards understanding the underlying mechanisms. This will aid in future work of being able to foresee modifications needed for a certain selectivity or activity with also other epoxides. As a far-end goal we want to use this enzyme together with other well-characterized enzymes to perform multi-step biocatalytic reactions as an alternative to man-made catalysts used in industry today.

Green chemistry

To save our planet we need to treat it in a sustainable and greener way and to think more about the environment. This is true for chemical industry in particular where organic chemists work hard on trying to fulfill the twelve principles of green chemistry, a list created already in 1998 (Box I). The principles can be summed up as to: conserve energy and resources and avoid waste and hazardous material in industrial processes. Parameters to describe processes has arisen and amongst them E-factor (mass ratio of waste to desired product), atom economy (molecular weight of desired product divided by the sum of molecular weight of all substances produced), and step economy are commonly used, all serving to increase yield and reduce waste. Earlier, efficiency was only discussed in high yields and said nothing about the waste where today process optimization also involves the use of raw material and elimination of waste and toxic compounds. When it comes to waste, not only the actual amount but also the nature of it is important and various factors have been introduced (e.g. the unfriendliness factor) to be used in combination with the E-factor.
Box I. The twelve principles of green chemistry that serve as points of optimization for chemical processes.

- Prevent waste
- Atom economy
- Less hazardous syntheses
- Design safer chemicals
- Safer solvents and auxiliaries
- Design for energy efficiency
- Use of renewable feedstocks
- Reduce derivatives
- Catalysis
- Design for degradation
- Real-time analysis for pollution prevention
- Inherently safer chemistry accident prevention

Biocatalysis

Biocatalysis is nowadays broadly defined as the use of enzymes or whole cells as catalysts for industrial synthetic chemistry. If using whole cells, the cell membrane needs to be permeable to the substrate or the enzyme needs to be expressed on the surface. Whole cells are easy and cheap to produce and co-factor regeneration is solved by the endogenous metabolic network. This network, however, could also be a true drawback due to other pathways acting to disturb and regulate the target reaction. This is avoided if synthetic pathways are created, by combining free enzymes. Here, drawbacks are instead the need of high stability, for the enzyme not to be degraded outside the cell, and purification, adding a cost in both time and money.

Why enzymes are suitable catalysts

Biocatalytic reaction pathways are great alternatives for making industrial chemical processes greener* and more environmentally friendly. What biocatalysts often do, compared to traditional metallo- or organocatalysts, is reducing the total waste, eliminating the use of transition metals and some-

* Green biotechnology refers to the application in agricultural processes and for industrial processes the term white biotechnology is appropriate.
times-toxic compounds and improving yield and productivity. Enzymes, by their inherent properties, are meeting many of the twelve principles of green chemistry in being biodegradable and harm-free to both environment and humans, functioning well at ambient temperatures and pressures in water-based solutions. They are efficient and often highly chemo-regio- and enantioselective thus increasing product quality and simplifying downstream processing. Most enzymes function under similar conditions and can be combined to co-function in one pot without the need for group activation or protection/deprotection, leading to less waste and a more atom- and step-economic processes.

Previous drawbacks with enzymes as catalysts included high production cost and low stability, something that has been overcome by the enormous progress in DNA technology, where today it is possible to express almost any enzyme in large amounts using simple E. coli systems. The stability issue is routinely solved by various immobilization techniques or protein engineering.

Industrial Biocatalysis

What the chemical manufacturing industry desire is stable, selective and productive catalysts functioning well under the conditions presently used in the process. Thus, the talk about enzymes being environmentally friendly due to functioning in water is indeed true but to be implemented in existing processes they need to be active also in the organic solvents used. There is on-going research in trying to perform organic synthesis in water-based systems but focus is also on the nature of the organic solvents (to be less harmful according to FDA) or alternatives like ionic liquids. Most enzymes do not withstand high concentration of organic solvents and therefore stability needs to be improved. Additionally, if there is a need for co-factors, (compounds that can be expensive, especially in larger scale) either whole cells or co-factor regeneration should be considered for a cost-efficient process.

Enzymes will not necessarily replace all steps in present industrial processes but rather be used for key steps where high selectivity is of importance. Enzymes come in especially handy when working with chiral molecules, with the wish for enantiopure products, due to enzymes’ chiral nature and their possibility to show high selectivity. Most synthetic reactions in pharmaceutical industry include steps with chiral molecules and despite the fact that two enantiomers are highly similar (Figure 1), being mirror images of each other, their biological activities can be completely different in nature. When it comes to pharmaceuticals, usually only one of the enantiomers has the targeted effect while the other could be not only inefficient but also detrimental. One such classic example is thalidomide, a tranquilizer recommend-
ed for pregnant women where only one enantiomer gave the calming effect while the other gave severe birth defects. From an economical and green point of view, also in cases where the inefficient enantiomer is harmless, it is beneficial to produce only the bioactive one.

Figure 1. Two stereoisomers that are mirror images thus called enantiomers.

Enzymes are remarkable catalysts and many enzymes are used in chemical industry today, however, they could be implemented even more in industrial processes. There is still a knowledge gap between the two fields of biocatalysis and organic synthesis, reducing the transfer of enzymes into this industry, the situation is improving (one successful example being the recent contract between Codexis and Merck). In most cases, enzymes need to be engineered to fulfill the criteria of stability, activity, selectivity and substrate scope where companies like BASF, Codexis and Novozymes all have platforms with enzymes and high-throughput techniques to obtain tailored variants for a specific need.

The importance of selectivity has been presented. Even though enzymes in theory look like the ideal catalyst they often need to be engineered to carry the preferred activity, selectivity, stability or substrate scope wished for. Before going there, I will talk a bit about enzyme selectivity and catalysis.

Enzymes’ role in catalysis

Enzymes catalyze reactions by lowering the activation energy for the transition-state intermediate by the tight fit allowing for electrostatic and hydrophobic interactions between the enzyme and the substrate. Additionally, the structure of the active site allow for selection of substrates by steric hindrance and also for positioning of substrates to be close to the catalytic residues of the enzyme. This is especially important for reactions with more than one substrate, or reactions needing co-factors, where the enzyme keeps the components close in space, thereby increasing the probability for the reaction to occur. Furthermore, in the active site of an enzyme, reactions occur in a hydrophobic pocket allowing for more controlled reaction conditions. The low water content means that even hydrolysis can occur under controlled and
selective conditions. The low water content in combination with various interactions of closely positioned amino acids often lead to abnormal pK\textsubscript{a} values and the positioning of amino acids residues also lead to coupled protonation states from e.g. charge repulsion. These are the fundamentals of enzyme catalysis, but what is actually making enzymes as remarkable as they are is still not completely understood, and might only be for nature to know. Another factor that has arisen to explain enzyme catalysis is dynamics. Molecular dynamics show that enzymes are not rigid but in constant motion where some parts are more rigid, e.g. the core, while others are more flexible, e.g. the loops. Dynamic motions range on the scale from femtoseconds to minutes and their contributions to the catalytic powers of enzymes is an on-going debate.\textsuperscript{15–22}

Enzyme selectivity

First, it is important to clarify the difference between specificity and selectivity due to these two terms being sometimes used as synonyms. Selectivity is when one product is favored over the other and thus mostly one product is formed. Specificity, on the other hand, refers to complete selectivity meaning that only one product is formed. In other words, complete selectivity is the same as specificity (IUPAC).

Enzymes are often described as being both specific and highly chemo-, regio- and enantioselective (Figure 2), which they indeed can be, but in many cases they are promiscuous and function for a range of substrates.\textsuperscript{23} For biocatalytic purposes perfect selectivities are aimed for in order to obtain pure products and low amount of waste. Thus in most cases protein engineering needs to be applied.

![Figure 2. Schematic figures of (from left to right) chemo-, enantio- and regioselectivity.](image)

When working with chiral catalysis (also named enantioselective catalysis or asymmetric catalysis) and a racemic mixture of substrate both enantio- and regioselectivity is important. Formation of enantiopure product is either met by kinetic resolution (Figure 3, A) or by enanticonvergence (Figure 3, B). For kinetic resolution only one enantiomer is reacted while the other is left
unreacted thus leading to a limit of 50 % conversion. This limit is overcome by enzymes that show enantioconvergence, where there is a change in regioslectivity for the different substrate enantiomers. Enantioconvergence is sometimes seen naturally in enzymes but could otherwise be obtained from protein engineering or by using a combination of enzymes with complementary selectivity.

![Figure 3](image)

Figure 3. (A) Kinetic resolution of two enantiomers resulting in maximum 50 % enantiopure diol and the non-reacted epoxide enantiomer and (B) complete enantioconvergence resulting in 100 % enantiopure product from a racemic substrate.

To be able to obtain a successful biocatalyst, in most cases the catalytic properties and/or the substrate selectivity needs to be improved. Thus a stable enzyme susceptible to modifications is wished for.

Improving enzymes to be better catalysts

Stability

The wish to improve enzyme stability has already been touched upon when discussing tolerance towards organic solvents and stability was the first property to be targeted for both enzymes and whole cells, solved by immobilization. Immobilization of enzymes does not only improve stability but also adds the possibility of re-use of the catalyst without complicated downstream processing. A drawback with immobilization is a possible decrease in activity thus an optimal method needs to be found in each case. Different techniques allow for more or less controlled enzyme immobilization where a totally random approach will result in low efficiency due to binding also occurring to the active site. An alternative or additional method for improving stability is protein engineering where modifications known to be beneficial for stability (e.g. di-sulphide bridges and residues allowing for salt bridges within the enzyme) is introduced. If the structure of the protein is
not known, random mutagenesis can be applied targeting the whole sequence, thus a screen for variants with improved stability and retained activity is needed.

Activity, selectivity and substrate scope

Directed evolution is a process where a rapid version of nature’s own evolution is copied in a test tube and this method has emerged as an effective process for tuning selectivity and activity and to broaden the substrate scope for enzymes. At first, this procedure was performed completely randomly but upon more protein structures being solved and made available, more structure-based approaches were allowed and also more rational design. Protein engineering is nowadays rather straightforward and in most cases variants with improved selectivity and activity are indeed isolated.

More focused approaches often result in larger improvement in selectivity compared to random alternatives, thus motivating the use of methods focusing on residues lining the active site. However, also modifications far away from the active site, in so called second or third shell residues, have shown to be beneficial, most likely due to compensating to stability loss upon binding to a new substrate but possibly also by influencing enzyme dynamics. This brings a challenge for both experiments and computations, as the whole protein needs to be targeted for engineering purposes and during computational modeling. The capacity for calculations is increasing and in silico studies are used more and more, allowing also for larger systems to be included in calculations to describe kinetics. Now it is even possible to make de novo design of enzymes with decent and also completely new activities. In many cases, however, de novo design result in poor activities, suggested to be from lack of dynamics, but computational power in theoretical modeling is nevertheless aiding in directed evolution, making it more rational.

Challenges & bottlenecks

For the directed evolution, mutations can either be introduced one by one or in combination of two or more residues. If introduced one at the time the direct effect is seen while one miss out on possible cooperative effects. Also, a large number of modifications might be needed and since enzymes are evolved to function to perfection in its natural habitat, introducing too many modifications normally comes with a loss in activity. This could be met by first introducing stabilizing mutations for the enzyme to be more susceptible to further engineering and possible loss in binding energy. This could either be made by introducing random or site-directed mutations or by reconstructing an ancestral protein. These common ancestors are thought to be generalists, with a broader substrate scope, that has evolved into different
specialists. Thus despite the search for a highly selective catalyst, starting out with a promiscuous enzyme is in many cases preferred, allowing for reactions with a broader range of substrates and to stepwise create different specialists. The use of the one-by-one approach is supported by the fact that a property is not unique but can occur from a number of different modifications.\textsuperscript{49,50} When instead introducing two or more residues simultaneously synergistic effects are allowed for.

A common bottleneck in directed evolution is library construction and screening. Enzyme libraries consisting of more than one residue is often constructed, due to multiple modifications not always being additive but one by one could have no or detrimental effect but in concert they are highly beneficial (epistasis).\textsuperscript{36} A drawback with modifying too many residues at the same time is the cost in screening effort and a high-throughput method is crucial. Examples of fluorescent cell sorting, colony screens\textsuperscript{51,52} and ESI-MS for screening enantioselectivity\textsuperscript{53,54} have been reported as well as a solid-phase gene synthesis.\textsuperscript{55} The fast improvement in DNA technology has resulted in the possibility to order any gene of interest and to codon-optimize it for expression in an organism of choice. Thus the time from idea to isolation of protein is significantly reduced. Even whole pre-made libraries can be ordered, saving not only time for construction but also for screening and sequencing due the fact that geno- and phenotype is linked. Thus, there are ways to overcome the bottleneck of library construction and the need of extensive oversampling.

\textit{Nowadays, with increasing power of DNA technology, enzyme cost and availability is in most cases not an issue but instead the limit lies in engineering the enzyme to more efficiently catalyse the reaction of choice. To reduce this bottleneck there is a need for understanding structure-function relationships to be able to predict modifications needed for a certain property. The target system for this thesis is epoxide hydrolysis by epoxide hydrolases.}

\section*{Epoxides}

Epoxides are a group of ethers being reactive due to the strained angles of the three-membered epoxide ring and the partial positive charge on the oxirane carbons, making them strong electrophiles. Since there are two of these carbons and due to these substrates, in many cases, being chiral, both regio- and enantioselectivity is important to avoid multiple products.

Hydrolysis of epoxides can be either acid or base catalyzed. At neutral or basic pH, steric factors usually dominate and the reaction occurs via $\text{S_N}_2$ attack on the least substituted epoxide carbon. Under acidic conditions, the
Epoxide is protonated and reaction can occur via both $S_N1$ and $S_N2$ at the most substituted carbon due to electron donating effects stabilizing the positive charge of the oxirane carbon attacked. In other words, regioselectivity can be directed by simply changing the pH, adding complexity is that the nature of the substituents also influences reactivity of the carbons. Here, electron-donating groups allow for attack at the substituted carbon also at neutral and basic pH. An example of this is for styrene oxide where under basic conditions both carbons are attacked to the same extent. In addition, when starting from cheaper racemic substrates, also enantioselectivity is highly important and this is poor in non-enzymatic systems since enantiomers often appear identical to free nucleophiles. Instead, chiral catalysts like enzymes can be used for kinetic resolution or, in combination with changed regioselectivity, enantioconvergence.

The reason why epoxide hydrolases are of interest for industrial purposes and biocatalysis is due to the chiral nature of both the substrate epoxides and the product diols and their presence as common key building blocks in organic synthesis. In current processes including epoxide hydrolysis with organic catalysts the main issue is the need of high amount of expensive catalyst and a limited efficiency with sometimes varying regioselectivity and lack of enantioselectivity. It is here enzymes come in handy in fulfilling all these criteria.

**Epoxides studied here**

The environment in the active site of an enzyme allows for other reactions than those predicted to occur in solution. The epoxides studied here are presented in Figure 4. For epoxides 2 and 3 the attack is assumed to occur at the phenyl-substituted carbon but wild-type StEH1 did previously show not to obey this rule. However, for epoxide 1 the phenyl-ring is positioned one carbon away from the oxirane carbons thus steering of the nucleophilic attack would be even harder, but assumed to be at the carbon of least steric hindrance. Epoxide 5 is bulky and does not allow for different binding modes in the active site, also only one product is formed independent of where the nucleophilic attack occurs thus only enantioselectivity could be monitored experimentally.
A problem with phenyl substituted epoxide substrates, when it comes to enzymatic hydrolysis, is their low solubility in water and if solvents is added the enzyme have to be active and stable in such environment. To improve stability, immobilization could be used or an alternative two-phase system either by adding an organic phase or by simply adding more epoxide than can be dissolved thus forming droplets of substrate in the buffer solution, keeping the concentration of free epoxide low enough. Another problem could be spontaneous non-selective hydrolysis thus a sufficiently fast reaction is important to be able to control product outcome.

**Epoxide hydrolases**

Epoxide hydrolases are widespread in both host and function\textsuperscript{62,64,65} and is therefore covering hydrolysis of a range of epoxide substrates with different activities and selectivities. Most of the epoxide hydrolases belong to the same superfamily of $\alpha/\beta$-hydrolase fold, sharing a common structure and catalytic triad and therefore also mechanism. This leads to the fact that understanding the structure-function relationships for one enzyme could also be applicable to related ones. As mentioned, there are epoxide hydrolases active with different epoxides and with different enantioselectivities and they do function to give enantioconvergent conversions, either alone\textsuperscript{66,67} or in tandem in one-pot reactions\textsuperscript{68–70}.

This $\alpha/\beta$-hydrolase fold family\textsuperscript{71} consists not only of epoxide hydrolases but also lipases, esterases, peptidases, proteases, peroxidases and dehalogenases. These enzymes have diverged from a common ancestor with a preserved
catalytic triad but a varying binding site, leading to enzymes covering a wide range of substrates. The family is named after the typical structural fold they share which is built up by β-sheets surrounded by α-helices and where the epoxide hydrolases of this superfamily all have a lid domain covering the core domain, together forming the active site pocket. The catalytic triad consists of a nucleophile, a histidine base and an acid for charge relay. The histidine is completely conserved within this superfamily and for epoxide hydrolases also the aspartic nucleophile while for other members this can be a serine or a cysteine. In the active site there is a so-called oxyanion hole built up by backbone amides to stabilize the negative reaction intermediates formed. For epoxide hydrolases there are two conserved tyrosine residues, necessary for positioning of the epoxide for nucleophilic attack. The kinetic mechanism is well understood and shared within the α/β-hydrolase family and proceeds in two steps, first alkyl-enzyme formation followed by hydrolysis, as described in Figure 5 where numbering in brackets refers to the residues in StEH1. For most epoxide hydrolases the hydrolysis step has been shown to be rate limiting, allowing for build up of alkylenzyme suggesting that its presence could be detected in the pre-steady state kinetics.

Figure 5. The reaction mechanism for epoxide hydrolases of the α/β-hydrolase family is a two-step reaction where an alkylenzyme formation is followed by a hydrolytic
step. The two tyrosines (Y154 and Y235) of the lid domain aid in forming the Michaelis complex and positioning the epoxide ring for nucleophilic attack. In the first step where the nucleophile (D105), activated by H300 that is charge-relayed by D265, attacks either of the oxirane carbons to form a covalent negatively charged alkyl-enzyme intermediate where again the lid tyrosines play an important role in charge stabilization. In the second step, a hydrolytic water molecule, positioned by E357 and activated by H300, attacks the alkyl enzyme to form a tetrahedral intermediate stabilized by backbone amides (Figure from Paper I).

Numerous protein-engineering studies have been performed for this class of enzymes to increase both substrate scope and selectivity. There are many successful cases where enantioselectivity was increased, sometimes referred to as the selectivity between substrates and sometimes for the actual product outcome. Methods used have been both random processes like error-prone PCR and DNA-shuffling, often motivated by the findings that also mutations far from the active site could influence selectivity, but also more rational approaches and site-directed mutagenesis based on structural information. Examples of residues at the entrance of the active site and further out have been pinpointed but most influence in selectivity is obtained from modifications of residues surrounding the active site, especially the residue next to the nucleophilic aspartate and the lid tyrosines. Due to structural similarities of this class of enzymes this is valuable information for future engineering of other epoxide hydrolases. To help in translating an amino acid position in the enzyme of one organism to another there are available databases made up from structures for all α/β-hydrolases that are crystallized.

These engineering strategies have been applied for a range of different epoxide hydrolases, amongst these are different limonene epoxide hydrolases (LEH) and EH from Agrobacterium radiobacter (ArEH) and Aspergillus niger (ArEH), all having known crystal structures. From these, it can be seen that the limonene epoxide hydrolase is not an α/β-hydrolase enzyme thus the mechanism is different. Here, the hydrolysis is acid catalyzed where the epoxide is protonated before nucleophilic attack by a hydroxide ion, thus no covalent intermediate is formed. All of these epoxide hydrolases, as most of the known ones, are homo-dimers, and they have rather small active sites and thus show rather narrow substrate scope.

There have been numerous QM/MM studies for systems related to those shown in this thesis and it has also been possible to predict the increase in enantioselectivity for LEH from in silico studies when compared to laboratory evolved enzymes. Interestingly, the modifications were not at all the same, possibly due to the large number of amino-acid residues modified simultaneously for the computational approach allowing for introduction of compensating mutations to allow for larger changes of the active site.
Potential industrial applications for epoxide hydrolases

Epoxide hydrolases are interesting for industrial purposes due to their asymmetric synthesis of vicinal diols, as both them and the chiral epoxides are important building blocks in pharmaceutical synthesis. The fact that there are so many different epoxide hydrolases with varying activities and selectivities allow for combination of the two to perform an all in all enantioconvergent process. There are also examples of where the enzymes themselves show enantioconvergence, something that is not seen in (m)any other enzymes. Epoxide hydrolases, including StEH1, have been implemented in industrial processes of different scales, for production of a number of known drugs.

A drawback is these enzymes inability to function in organic solvents where it needs to be stabilized, possibly by immobilization, for implementation in present industrial chemical processes. There are attempts where epoxide hydrolases has been successfully immobilized to function also in organic solvents but a problem is the need for water in order to function in hydrolysis. Alternatives to organic solvents are ionic liquids (ILs) that constitute a major research field within green chemistry where a range of epoxide hydrolases show retained activity and selectivity despite the activity being completely abolished in some organic solvents.

Epoxide hydrolase from potato

The epoxide hydrolase from Solanum tuberosum (StEH1) is a promising candidate for biocatalytic purposes due to it being a rather small (36.1 kDa) and monomeric enzyme without any need for co-factors. Its native role is not completely clear but the preferred substrates are fatty-acid epoxides and the enzyme is mostly expressed in the leaves of the potato plant suggesting a role in formation of cutin. This enzyme have a rather broad substrate scope due to a relatively large active site making it a good candidate for engineering purposes towards increased activity for a range of epoxides. The wild-type structure was solved to high resolution allowing for site-directed mutagenesis to confirm residues involved in the catalytic mechanism and also for more rational approaches for enzyme engineering. Further, StEH1 is easy to express in large amounts and is stable over time thus making it easy to work with and this enzyme has previously shown to be susceptible to modifications in both enantio- and regioselectivity where it is also known to show the remarkable property of enantioconvergence with some epoxides.
Figure 6. To the left is StEH1 (PDB entry 2cjp) with the catalytic triad and the two tyrosines involved in positioning the epoxide. To the right, a close-up of the active site.

Now, we have an enzyme that looks good in theory. The next step is to create modified variants and if successful these will be further characterized in order to map structure-function relationships.
The workflow

*Figure 7.* An overview of the workflow used for this study, from library construction via screening and hit characterization. Everything is performed in an iterative manner where the best hit is further characterized and template the next round of modification and screening. Each part is described in the following paragraphs of this booklet.

**Enzyme engineering**

Enzymes can be engineered in different ways depending on what the aim is and what knowledge is available (Table 1). To perform *rational design* structural information is needed, if not the crystal structure so at least the amino-acid sequence for the possibility to create a homology model, and knowledge about function of the enzyme. This information is used to perform *site-directed mutagenesis* where a certain residue is modified to a set one. This is often the method of choice to introduce e.g. disulphide bridges or study the importance of a certain residue and verify its involvement in the catalytic mechanism, as mentioned earlier.
In the other end of the spectrum is directed evolution that is instead a random process in an attempt to simulate a rapid version of nature’s own evolution. Here, no structural information is needed and some of the most commonly used methods include error-prone PCR, DNA shuffling and circular permutation.

Table 1. Different methods applicable for protein engineering.

<table>
<thead>
<tr>
<th>Method</th>
<th>Strategy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Error-prone PCR</td>
<td>Random</td>
</tr>
<tr>
<td>DNA shuffling</td>
<td>Random</td>
</tr>
<tr>
<td>Circular permutation</td>
<td>Random/Target flexibility and dynamics</td>
</tr>
<tr>
<td>Saturation mutagenesis</td>
<td>Specific</td>
</tr>
</tbody>
</table>

Directed evolution can also be used in a semi-rational way by restricting the area to modify. One way of identifying important residues is CASTing (Combinatorial Active-Site Test) where small focused libraries surrounding the active site is chosen, each of them being subject to random mutagenesis. One benefit with this method is the allowance for synergistic effects otherwise not seen. This method demands a known 3D-structure of the enzyme, with preferably the substrate bound, to identify residues possibly interacting with the substrate or having influence of the structure of the active site as such.

Once the method for directed evolution has been decided upon and as well as the residues to include in each library these sites need to be examined in a controlled manner.

Iterative Saturation Mutagenesis (ISM)

A commonly used method is ISM where iterative cycles of saturation mutagenesis are performed. The resulting hit from each round is used to template mutagenesis of another library and this process is repeated until satisfaction. ISM allows for doing this systematically to make sure all combinations in all orders are visited. If used in complete, ISM allows for epistasis where the effect of one modification is dependent on the presence of another or simply that there is a non-additive effect in combination compared to individual modifications. In short the ISM method involves: i) identification and ranking of important residues, ii) selection and grouping of target residues, iii) choice of diversification method, screening effort and amino acid alphabet.
For ranking the importance of a residue the priority goes: for residues reported to enhance activity > influence substrate scope > act in determining stereoselectivity > enhance stability. The ranking is needed to keep the number of colonies to screen manageable. At this point it is important to consult available data regarding conserved residues.

To reduce library sizes, the residues included in each library should be kept low (typically not more than 3). To further reduce screening effort a degenerate codon set could be considered, a common alternative being NDT allowing for 12 out of 20 amino acids covering most of the different characteristics of the natural amino acids (Table 2). Despite this being considered a reasonable trade-off, the reduced library size comes with a sacrifice in possible missed potential hits. The use of NNK codon allows for all 20 natural amino acids but involves 32 possible codons, thus several amino acids are overrepresented leading to a bias in the method and also more mutants to screen. When applying NDT for libraries consisting of two amino acids only 430 clones need to be screened for a 95 % coverage as compared to 3068 with NNK.131

Table 2. The amino acids coded for using degenerate codon set NDT.

<table>
<thead>
<tr>
<th>Amino acids</th>
<th>NDT&lt;sup&gt;a&lt;/sup&gt;</th>
<th>Nature of sidechain</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>F L I V Y</td>
<td>Hydrophobic</td>
</tr>
<tr>
<td></td>
<td>H R D N S C</td>
<td>Charged or polar</td>
</tr>
<tr>
<td></td>
<td>G</td>
<td>Small</td>
</tr>
</tbody>
</table>

<sup>a</sup> N = all bases, D = A, G, T

To overcome the problem with biased codon-set and having to sacrifice the number of amino acids included, small intelligent libraries are constructed by using a combination of codon sets resulting in only one codon per amino acid.132,133 Here, all twenty natural amino acids are allowed for while reducing the screening effort to 1450 clones for the same coverage as discussed above. The decision about how restricted the libraries need to be is often a matter of the potential throughput of the screening method in combination with the knowledge about the surroundings of the site for modification.

Screening and selection

All efforts in trying to reduce library size are aimed towards having a manageable set of variants for screening and hit identification, something that is often the bottleneck in directed evolution. To allow for cooperativity, often more than one residue is modified at the time, thus the number of clones to
screen demands a method of relatively high throughput. The screen is designed dependent on the property searched for where stability, activity or enantioselectivity are common alternatives. Often libraries are transferred to 96-well microtiterplates where a screen for the lysate is applied. Examples of fast methods are colorimetric screens or the use of substrates where product formation can be followed by UV/Vis spectrophotometry. In these cases the screen is used to directly get activity and data can be gathered also for variants that do not show improved behavior. There is a lot of information to get from these variants where also modifications detrimental to the sought-after property can be pin-pointed. Often an initial, less complicated screen is used and the hits isolated from here are further analyzed to confirm potential hits. The colorimetric method used here show different shades of blue depending on how much unreacted epoxide is left.134,135 If the same lysate is used for screening of two enantiomers in parallel, enantioselectivity can be selected for. This method has been developed to also function for whole cells and colonies directly from the plate, something that would increase the throughput.52 The identified top enzyme variants will be sent for sequencing and then be analyzed in more detail to confirm the hit.

Now when a set of enzyme variants has been selected from a primary screen and these need to be characterized in further detail to describe their actual activity and selectivity.

Characterization

All isolated hits need to be characterized to assure and map their properties. Decisions have to be taken concerning parameters to examine based on what property is searched for. A covering characterization can be time-consuming and in some cases only the very best variant can be documented. Parameters of interest are normally describing kinetics and selectivity but also stability and of course the actual sequencing. It can include measurements with a whole spectrum of substrates under different pH and temperatures, solvents and buffer conditions.

Enzyme kinetics

The most important property for characterization of enzyme function is kinetics, describing the enzyme’s catalytic efficiency. There are two different time-scales to measure on, the longer resulting in steady-state and the shorter for pre-steady-state kinetics.
**Steady-state kinetics**

Many enzymes have complicated reaction mechanisms but most obey, more or less strictly, Michaelis-Menten kinetics, in its simplest form shown in Scheme 1.

\[
E + S \xrightleftharpoons[k_1^{-1}]{k_1} ES \xrightarrow{k_2} E + P
\]

*Scheme 1.* The simplest form of kinetic mechanism where substrate and enzyme form a Michaelis complex (ES) and subsequently broken down either back to free enzyme and substrate or to form product.

This model is described by Equation 1 where it is important to make sure that initial velocities are measured such that the change in substrate concentration can be considered constant and the product concentration neglected. In other words, the method involves two assumptions; \( k_2 \) is negligible and the concentration of Michaelis complex (ES) is constant.

\[
v_0 = \frac{v_{\text{max}}[S]}{K_M + [S]}
\]  

(1)

where the initial rates’ \( (v_0) \) dependence on substrate concentration \( ([S]) \) is shown, allowing for determination of the maximal rate \( (V_{\text{max}}) \) and the Michaelis constant \( (K_M) \).

Any positive cooperativity adds sigmoidality to the saturation curve being it from allosteric regulation or conformational changes. Complexity is also added by the existence of more than one enzyme bound intermediate (Scheme 2) where expression for \( k_{\text{cat}} \) and \( K_M \) gets more complicated.

\[
E + S \xrightleftharpoons[k_1^{-1}]{k_1} ES \xrightleftharpoons[k_2^{-1}]{k_2} EA \xrightarrow{k_3} E + P
\]

*Scheme 2.* The kinetic mechanism an additional enzyme intermediate (EA) is formed.

Steady-state measurements describes the “equilibrium” reaction giving the combined constants \( k_{\text{cat}} \) and \( K_M \). \( k_{\text{cat}} \) is the turnover number, the amount of product formed per enzyme molecule per time unit when all enzyme are bound to substrate and include the rate constants for all steps of the reaction after substrate binding. \( K_M \) is an apparent dissociation constant of all enzyme-bound substrate complexes where substrate turnover is not included.
$K_M$ is often described as the affinity for the substrate, something that is only true for the simple reaction in scheme 1 if $k_2$ is negligible in comparison to $k_{-1}$. The affinity is instead described by the dissociation constant $K_S (k_{-1}/k_1)$.

$K_M$ and $k_{cat}$ are complex expressions built up of a number of rate constants that can be determined by pre-steady state kinetic measurements. These rates are important for attempts to pinpoint rate-determining steps and explaining product outcome.

In steady-state kinetic measurements activity is often detected spectrophotometrically, demanding different absorbance spectra for substrate and product. If there is no such difference, endpoint measurements must instead be performed, taking samples at different time points and analyze the product formation e.g. by HPLC or mass spectrometry.

**Pre-steady-state kinetics**

To reveal the transient rate constants building up $K_M$ and $k_{cat}$ pre-steady-state kinetics needs to be analyzed where an instrument able to measure on a shorter time scale (typically ms) is used. An example is stopped-flow spectro/fluorophotometry where one needs to have a detectable signal upon e.g. binding of substrate in absorbance or fluorescence. In the case of StEH1, tryptophan fluorescence is quenched upon alkylenzyme formation allowing for detection of rates for formation and decay thereof. Equation 2 can be applied to reveal the different intrinsic rate constants. For multiple-turnover measurements, enzyme concentration must be well below substrate concentration to allow for assumption of this to be constant during the detection. What is measured is the difference in rates of decay plus formation of alkylenzyme.

$$k_{obs} = \frac{k_2[S]}{K_S+[S]} + k_{-2} + k_3 \quad (2)$$

The observed rates are substrate dependent and the intrinsic rate constants can be obtained. Single-turnover experiments are only possible for reactions with low $K_S$ due to the need for substrate concentrations well above this value to assure saturation and enzyme concentration to be well above this concentration to assume all substrate to be bound and consumed during one catalytic cycle.

**Selectivity & Specificity**

How could specificity be described and compared between different enzyme variants? $k_{cat}$ does not include substrate binding and is therefore not a suitable measure whereas $K_M$ does not include substrate turnover thus neither of these parameters are alone a good measure. Instead a combined quota is used the
so-called specific activity, $k_{\text{cat}}/K_M$. This parameter is used to compare the specificity of two substrates or two enantiomers of one substrate. In the latter case this gives the enantioselectivity thus how the enzyme distinguishes between the two enantiomers of a substrate, often presented by the enantiomeric ratio (E) described by $\left(\frac{k_{\text{cat}}/K_M}{E}\right)^S_{R} / \left(\frac{k_{\text{cat}}/K_M}{E}\right)^S_{S}$.\(^{136}\)

Another important parameter for describing enzymes is *regioselectivity*. In the case of epoxides this describes the enzyme’s ability to distinguish between the two oxirane carbons for nucleophilic attack (Figure 8). When varying regioselectivities are involved the E-value is not suitable for describing product outcome but enantiomeric excess of product or product ratio should be used.

\[\text{Figure 8. The product outcome from an enantiopure substrate upon different regioselectivity.}\]

Once the enzyme is characterized and well described there is a wish to try to understand structure-function relationships in order to be able to predict mutations needed for a certain catalytic property. In an attempt to solve this question experimental data can be combined with crystallization and theoretical modeling.

The power of structure determinations

The number of protein structures deposited in PDB has exploded during the last decades with X-ray crystallography still being the by far most commonly used method. Determination of the tertiary structure of proteins adds invaluable information for understanding of biological systems. X-ray crystallography provides atomic resolution and thus the possibility to study proteins in detail.

Crystallization can be time-consuming and for some proteins not achievable. Conditions for a successful crystallization cannot be foreseen but are determined by trial and error. When crystallizing a range of modified variants of the same enzyme a good starting point is to use similar conditions but this does not necessarily have to be successful. Methods used for crystallization are normally hanging or sitting drop vapor diffusion, where in both cases the difference in precipitant concentration in the drop and the container solution
will make water diffuse from the drop thus causing an increase in protein concentration that initiates precipitation.

When comparing X-ray crystallization data for small organic molecules and large proteins the accuracy is much lower for the latter, a consequence of proteins being dynamic and in constant motion. If a few conformers are overrepresented these could be resolved one by one from the same scattering data but if there are more (which is often the case) then they cannot be distinguished and thus an average structure resembling the most commonly occupied one will be obtained. This average conformer does not necessarily resemble the catalytically efficient one since crystallographic conditions differ from the native enzyme habitat in demanding high protein concentrations, cryo protectants, precipitants and in many cases absence of substrate. One should also be careful when tuning the pH used since this will contribute to shifts in the structure depending on different protonation states of possibly interacting residues. Crystal structures do nevertheless contribute a lot to understanding of protein structure, and can identify residues putatively contributing to catalysis or substrate binding, especially in cases where there is a possibility to co-crystallize the enzyme with a ligand or substrate. To trap a covalently bound substrate in the enzyme normally demands modifications of the enzyme to be catalytically impaired or the design of an inhibitor or transition-state analog that will allow binding but not product formation.

When dealing with proteins that do not crystallize, multidimensional nuclear magnetic resonance (NMR) is an alternative method. Here, enzyme dynamics and conformational flexibility is studied at near-physiological conditions (in solution and at lower enzyme concentration). With this method differences in structure upon binding of substrate can be monitored, however, on behalf of resolution. Other possible drawbacks are the size limit of 35 kDa for the studied protein, and that the protein needs to be stable at room temperature during the run.

A third alternative for protein structure determination is cryo-electron microscopy (cryo-EM). This method has made an enormous development in recent years and thus the number of structures determined by cryo-EM to date (about 1/100 of the number of crystal structures) will for sure increase rapidly, especially with achieved resolutions getting better where today 2.2 Å is possible. One drawback of cryo-EM is that the protein studied needs to have a certain size: thus so far it is mostly used to study large complexes and membrane proteins.

For the present work X-ray crystallization was the method of choice where the structural information obtained allows for site-directed mutagenesis and also a possibility to perform more rational approaches for directed evolution. The crystal structure for wild-type StEH1 was solved previously and pro-
vided invaluable information that led to a better understanding of the catalytic mechanism\textsuperscript{74,75,125} and to identification of functionally important residues to be targeted by various engineering approaches.\textsuperscript{126} Having a solved crystal structure for the wild-type enzyme also makes crystallization of slightly modified variants easier as the conditions under which they crystallize may be identical or very similar and also facilitates solving the crystal structures of the variants by providing a search model for molecular replacement.

Once structures are solved these can be used to improve theoretical models of the reaction mechanism, not having to manually modify the residues. However, one needs to remember that these are rigid structures without ligands bound.

Theoretical modeling for experimentalists

Theoretical modeling is increasing in use and importance for enzyme systems due to an improvement in computing powers now allowing for modeling of also bond-making and breaking in large systems like proteins. This also leads to improvement in accuracy making the theoretical models able to reproduce subtle effects like enantio- and regioselectivity. What is used to compare calculations with experiments is Equation 3 based on the transition state theory,

\[ k = \frac{k_B T}{h} e^{-\frac{\Delta G^*}{RT}} \]  

(3)

where \( k \) is the rate constant (s\(^{-1}\)), \( k_B \) is Boltzmann’s constant (1.38 \( \times \) 10\(^{-23} \) J/K), \( T \) is the temperature (K), \( h \) is Planck’s constant (6.626 \( \times \) 10\(^{-34} \) Js), \( \Delta G^* \) is Gibbs free energy difference between the reactant and the transition state, and \( R \) is the gas constant (8.314 J/K).

From this equation it can be seen that a large rate enhancement of 10 will result in an energy difference of only 1.4 kcal/mol, often within the error of the calculations. This means that theoretical modeling is by no doubt a powerful instrument in trying to explain behaviors seen from experiments and to suggest amino acid residues contributing to rate enhancement, however, not yet a tool for determining rates of the enzymatic reaction.

Different methods for theoretical modeling demands different computational power and can therefore be applied for larger or smaller systems (Table 3). Molecular mechanics (MM) and molecular dynamics (MD) can be used to study all atoms in a protein where MM is used for docking experiments. In MD, sampling according to Newtonian physics allows the investigation of equilibrium properties, but the increasing complexity of the calculations also increases the computational demand. If the interest is in the actual chemistry,
thus to study the kinetics of the enzyme-substrate system, quantum mechanics (QM) needs to be involved. Here, the computational demand is very high and this system is limited to study only a few atoms. Thus, QM cannot be used to study the whole protein but often a combination of QM/MM is used where QM is applied for the substrate and the closest atoms of the enzyme which is involved in catalysis, and MM is applied for the rest of the protein. For the QM part the empirical valence bond theory (EVB) was applied. This model is based on the introduction of individual valence bond states for all reacting species and states. Using a well described reference state, this method allows the investigation of changes to the electrostatic environment of the reaction, as might be experienced for shifting a reaction from solvent to a protein active site.\textsuperscript{138–140}

<table>
<thead>
<tr>
<th>Parameters included</th>
<th>What system could be applied</th>
</tr>
</thead>
<tbody>
<tr>
<td>Molecular Mechanics MM Bonded and non-bonded interactions</td>
<td>Whole protein</td>
</tr>
<tr>
<td>Molecular Dynamics MD Dynamics included</td>
<td>Whole protein</td>
</tr>
<tr>
<td>Quantum Mechanics QM Bond breaking/making</td>
<td>A limited number of atoms</td>
</tr>
</tbody>
</table>

\textit{Table 3. Different methods for calculations and what can be obtained for these.}
Towards multistep biocatalysis

Now enzyme variants and methods fill up the toolbox from which one can pick and match to combine them in different modes depending on the reaction of choice. Structural information from crystallization is combined with known preferences and activities from characterization and calculations to get a more educated guess. And this is where my work ended. What is important when combining two or more different enzymes is to find optimal reaction condition for buffer choice and pH. In cases where co-factors are needed these should be re-generated in order to keep the costs down. Further, the more compounds you include the more complex the reaction gets in inhibition and downstream processing. More choices to make are whether to use free enzymes or whole cells, if to immobilize them and in that case by what method.
My work

Aim and strategy

My thesis is based on two parts: A) Directed evolution of StEH1 and understanding of structure-function relationships and B) Towards multi-step biocatalysis.

In the major part (A), directed evolution was applied for creation of a number of enzyme variants with different enantio- and regioselectivity to try to understand structure-function relationships. Experimental data was together with solved crystal structures used to train mathematical models with the aim of being able to in silico predict mutations needed to get a desired catalytic function.

Modifying an enzyme to perform a certain reaction is often a bottleneck in various applications, to overcome this, understanding why an enzyme behaves like it does is of great importance. If the structure-function relationship for a set of enzyme variants is mapped there is a possibility of foreseeing which mutations to introduce a certain property to be achieved. In this way one can create a platform of enzyme variants where one could relatively easy and fast find a suitable candidate, with the long-term aim of building up a toolbox consisting enzyme variants with a range of different properties.

To be able to create a platform like this, understanding the enzymes preferences for one or the other stereoisomer of a substrate and why the attack is on a certain carbon of an epoxide ring, is needed. Thus we have studied enantio- and regioselectivity of different enzyme variants and through collaboration with crystallographers and theoretical chemists we are approaching an understanding of structure-function relationships.

In all my work the enzyme used has been an epoxide hydrolase from Solanum tuberosum and variants thereof. This enzyme has been thoroughly studied over the past decade resulting in a large amount of experimental data for a range of epoxide substrates. In addition, the crystal structure for the wild-type enzyme was solved and thus the potential for finding a structure-function relationship seemed promising. The enzyme had also been shown to be susceptible to changes in both enantio- and regioselectivity.
In the second part (B) engineered enzymes are put together *in vitro* to perform a two-step biocatalytic reaction. Thus, the long-term goal will be a toolbox constituting not only enzyme variants with a range of properties but also techniques and a platform for optimization of reaction conditions.
Obtaining Optical Purity for Product Diols in Enzyme-Catalyzed Epoxide Hydrolysis: Contributions from Changes in both Enantio- and Regioselectivity

At the start of this project we had the crystal structure for the wild-type enzyme, a suggested and well-studied mechanism and an enzyme variant showing a reverted enantioselectivity compared to the wild-type enzyme. Directed evolution was performed to obtain enzyme variants with further increased product purity from racemic (2,3-epoxypropyl)benzene. In addition, activity and selectivity with two structurally similar epoxides were studied to see if any patterns in structure-activity relationships could be observed.

The wild-type enzyme shows a modest 2.5-fold preference for the (S)- compared to (R)-(2,3-epoxypropyl)benzene (1), the main substrate of this work. However, previous studies do describe isolated variants with reverted enantioselectivity. This finding was the stepping-stone here, with the aim of trying to improve product purity even further. In addition, two structurally related substrates, styrene oxide (2) and trans-2-methylstyrene oxide (3), were examined to see if any conclusions on selectivity could be drawn using the same enzyme hits. If patterns could be seen, one might use these substrates for screening instead, leading to higher throughput due to product formation being possible to detect by spectrophotometrical measurements. In addition, one would also be able to use the same enzyme for different biocatalytic reactions.

An attempt to reduce the screening effort

To reduce screening effort CASTing was applied where residues were picked for their structural involvement in the active site and the anticipated involvement in positioning of the substrate but without touching the catalytic residues. Library residues were previously decided upon (A-D) with the exception of library E consisting of L266 and V267. These two residues are positioned at the back of the substrate-binding pocket possibly affecting substrate positioning by steric constraints and interaction with the phenyl ring of the substrates (Figure 9).
Figure 9. The residues for libraries A (yellow), B (pink), C (green), D (petroleum) and E (orange) are shown as sticks together with the catalytically active residues (wheat).

In an attempt to further reduce the library size a degenerate codon set (NDT) was used, allowing for 12 out of 20 amino acids. For libraries of 2 residues 144 instead of 400 different combinations are possible and for 95 %- coverage only 430 instead of 3068 clones need to be analyzed. The 12 amino acids do represent different physicochemical properties and were considered to be a reasonable trade-off, and indeed a hit rate of 8 % was obtained for the first round of screening. Notable is that NDT does not include a tryptophan codon thus W106 was by default modified for variants visiting library B.

Drawbacks with an incomplete ISM approach

The ISM approach was not strictly followed in that not all library combinations were visited. Instead the best variant found in another study (R-C1B1) was used as template for the A and D libraries so that in total all sites had been visited (Figure 10). When running incomplete ISM, potential hits might be missed due to some modifications only being beneficial in the presence of others, a phenomenon called epistasis. When moving from library D to E the number of hits decrease from 42 to 8, screening the same number of variants. This indicates a need to insert stabilizing mutations for future variants to be susceptible to further modifications. Yet, a range of hits with varying enantio- and regiopreferences was identified, serving as a good base for understanding the underlying mechanism behind selectivity for this epoxide hydrolase.
Chaperonins necessary for tolerance of more modifications

No hits were isolated for the next generation using R-C1B1 as template without co-expression of chaperonins GroEL/ES. Thus strains also carrying a plasmid with genes coding for expression of chaperons were used from here on. For library A all isolated hits showed silent mutations, thus, retaining F33 indicates the importance of this residue for function. For the D library a number of hits with increased selectivity were isolated thus the best variant here (R-C1B1D33) was selected as template for the E library. Table 4 show that there is a decrease in number of hits when going from D to E library. It was also found that modifications occurring more than once are not necessarily the ones with the best selectivity, a clear-cut example being R-C1B1D33 with a unique modification (Table S12, Paper I). The explanation to some modifications occurring more often could be that they are being less detrimental to protein folding. If this is the case, introduction of compensating/stabilizing modifications might allow for other, beneficial, combinations of amino-acid substitutions.

Table 4. Information about library residues, library size and number of hits for the R-selected variants.

<table>
<thead>
<tr>
<th>Template</th>
<th>Library</th>
<th>Residues</th>
<th>Library size</th>
<th>Number of hits</th>
</tr>
</thead>
<tbody>
<tr>
<td>R-C1B1</td>
<td>A</td>
<td>F33</td>
<td>168</td>
<td>2</td>
</tr>
<tr>
<td>R-C1B1</td>
<td>D</td>
<td>I180 F189</td>
<td>504</td>
<td>42</td>
</tr>
<tr>
<td>R-C1B1D33</td>
<td>E</td>
<td>L266 V267</td>
<td>504</td>
<td>8</td>
</tr>
</tbody>
</table>
“You get what you screen for”

The colorimetric method used for screening is rather fast with acceptable throughput where lysate from each colony picked was examined for activity with both R-1 and S-1. Hits were selected based on their ratio of activity for the two enantiomers, i.e. enantioselectivity. In addition, also lysate showing high activity with both enantiomers were considered hits. In this way, possible hits showing a changed regioselectivity are also isolated. Protein concentration in the screen is not normalized for, but since the screen is comparative, using the same lysate for both enantiomer reactions, this would not influence the outcome. There is however a risk of missing out on the ones having too low activity, which might be due to low expression levels. On the other hand, low expression is also a sign of instability, which is not a desired property. You get what you screen for† is an expression often heard and truly valid for this work. The screen used for hit identification involves high epoxide concentration (10 mM) thus low $K_M$ was not selected for, something that is clearly seen in the results. Hits characterized did all (except R-C1B1D33E3) show an increase in $K_M$ for both enantiomers of 1. Except this increase in $K_M$ also $k_{cat}$ was reduced, more for S-1 than for R-1 resulting in a higher $k_{cat}/K_M$ for the latter. E-values ($((k_{cat}/K_M)^R)/(k_{cat}/K_M)^S$) for these R-selected variants increased from 0.4 for wild type to 14 for R-C1B1D33E3.

E-values do not tell the truth about product outcome

The actual product outcome from reaction with racemic 1 was studied over time and presented as product ratios, the ratio of one enantiomer over the other. However, when comparing these product ratios to the E-values obtained for some of the enzyme variants, the data did not agree, suggesting a change in regioselectivity, from the strict behavior seen for the wild type. This was indeed the case, where variants, to a certain degree, started preferring attack at C1 for S-1. However, the attack for R-1 was still strictly at C2. This enantioconvergent behavior, although not complete, is a sought-after property due to the possibility of overcoming the otherwise 50 %-conversion limit for kinetic resolution of a racemic substrate mixture.

So product outcome does not solely depend on enantiopreference for the epoxide substrate, but also on the regioselectivity of the nucleophilic attack and thus describing selectivities by E-value would therefore be highly misleading.

Since, in many cases, two products are formed from one substrate enantiomer this motivates the use of the branched reaction scheme also for reactions with pure enantiomeric substrate (Scheme 3). This means that the steady-

† Frances Arnold 1999
state kinetics does in fact measure the sum of two, not necessarily equal, rates. Some sigmoidality for several substrates was indeed seen in experiments. The sigmoidality was however not to an extent that addition of an extra parameter was statistically justified. To get the two individual rates of formation, samples from the steady state kinetics could be analyzed with chiral HPLC to get the product ratio for each time point.

\[
\begin{align*}
E & \xrightarrow{K_{S}, [S]} ES \\ k_{-0} & \quad k_{0} & \quad k_{5} \\
ES & \xrightarrow{K_{S}, [S]} E+diol_{1} \\
E & \xrightarrow{K_{S}, [S]} E'+[S] \\
E' & \xrightarrow{K_{S}, [S]} E'+diol_{2}
\end{align*}
\]

Scheme 3. The reaction scheme chosen to describe the reactions involves a substrate-independent conformational step and reversible isomerizations of the two Michaelis complexes (ES) while the two alkylenzymes (EA) are considered to not be interconvertible.

Changes in selectivity also for structurally similar epoxides

The behavior of enzyme variants with epoxides 2 and 3 did not follow that of 1, thus none of these would successfully serve as substitute substrates for screening, something that would otherwise present a faster throughput. The primary screen used here has a high throughput while further measurement for activity is a true bottleneck in having to use HPLC for endpoint measurements for different time points to build up a saturation curve. Due to this effort in identification of potential hits obtaining activities also for variants not showing enantio- or regioselectivity were not prioritized. To also map modifications that are detrimental to activity and selectivity would add important information in trying to understand the mechanism behind selectivity.

Despite no clear pattern between activities with the different epoxides the isolated variants did show varying enantio- and regioselectivities also for substrates 2 and 3 (Figure 4). The most dramatic effect in enantioselectivity is seen for R-C1B1D33 with S-2, going from an E-value of 69 for wild type to 5800. Concerning regioselectivity the S enantiomers is attacked at the benzylic carbon as would be predicted for the reaction in solution while for R-2 this is only true for wild type and R-C1B1D33E6. Also, for R,R-3 a range of different preferences are seen thus an indication a successful choice of library residues suggesting this set of enzyme variants to be applicable also for other epoxide substrates.
Important residues in deciding regioselectivity

Due to lack of crystals structures of the isolated enzyme variants all conclusions about structure/function relationships are based on simple dockings of substrates to the wild-type structure. Thus one should bear in mind that the discussion here is mostly speculative. The most pronounced effect is traced to the modification of residues W106 and L109, both interacting with the substrate and the former also contributing to the oxyanion hole. W106 is a highly conserved residue previously found to influence enzyme selectivity in related epoxide hydrolases, supporting the theory of its involvement also here.\textsuperscript{81,85,142} However, since this modification was not introduced alone but in tandem with position L109 the effect might not be solely due to modifications of W106.

We isolated enzyme variants showing increased enantioselectivity towards both \textit{R-4} and \textit{S-4}. The underlying cause was shown to be not only due to a change in enantio- but also regioselectivity, thus showing tendency of enantioconvergence. Deeper understanding of this was, at the time, mostly speculative and to get further insight one would wish to get crystal structures of the newly constructed enzyme variants, preferably with an intermediate bound. In addition, data for transient kinetics together with calculations would also add valuable information. No patterns of selectivity could be seen between epoxides 1, 2 and 3 but different product outcomes were obtained also for the epoxides not primarily selected for, indicating that the targeted residues did indeed play an important role in enantio- and regioselectivity among this set of structurally similar epoxides.
Expanding the Catalytic Triad in Epoxide Hydrolases and Related Enzymes

Wild-type epoxide hydrolase together with a set of modified variants, constructed to pin-point residues involved in catalysis, were examined theoretically repeat experimental values for selectivities and activities. Here, activity with a bulky substrate, allowing for few binding modes, was studied.

*trans*-stilbene oxide - a suitable model substrate

The symmetric substrate, *trans*-stilbene oxide (5), is a suitable model substrate, large enough to fill up the active site thus allowing for only one possible binding mode resulting in less expensive calculations. Once a reliable set-up for the calculations was found the same set of parameters could be used also for smaller and more flexible substrates. This substrate is not expected to show any enantio- or regioselectivity for the uncatalyzed reaction in solution. For the enzymatic reaction, on the other hand, a preference towards *R,R*-5 is seen from experimental data while regioselectivity cannot be monitored since attack at either carbon, for either enantiomer, will result in the same *meso*-hydrobenzoin product.

QM/MM calculations were applied to see if experimental data for enantioselectivity could be reproduced and regioselectivity explained. A more detailed understanding of the catalytic mechanism is needed to obtain further evolved variants and towards the aim of reliably and efficiently predict the effects of mutations on function *in silico*. Thus, activity for enzyme variants, previously constructed to confirm specific residues’ involvement in catalysis (E35Q, H300N, Y149F, Y154F and Y253F), was confirmed by calculations to validate the reaction mechanism. In addition, a variant carrying two of previous deleterious modifications in combination (E35Q/H300N) was constructed with the aim of getting a variant truly inactive in the hydrolytic half-reaction.

Crystal structure of H300N shows no catalytic water

To aid calculations available crystal structures for wild type and Y149F was used together with the structure for H300N that was solved here (PDB entry 4Y9S, resolution 2 Å). For the rest of the enzyme variants studied, modifications of amino acid were added manually to the wild-type structure thus adding some inaccuracy. When comparing the crystal structures of wild type and H300N, the side chain of N300 has a different orientation and there is no density where the hydrolytic water normally is positioned resulting in a
stronger interaction between H104 and E35 that is not seen in the wild-type enzyme.

Support of E35 acting as a back-up base for H300

Steady-state kinetics for E35Q/H300N showed no product formation after five minutes of incubation with either of the enantiomers of trans-stilbene oxide, thus this enzyme variant was considered inactive. This does however only assure that hydrolysis/product release does not take place whereas alkylenzyme formation could occur. H300 and E35 are thought to mainly be involved in activation and positioning of the hydrolytic water but are suggested also to help in activating D105 via charge relay and proton transportation out from the active site. To investigate the impaired activity further, stopped-flow measurements were performed. Here, alkylenzyme-formation is monitored by a decrease in tryptophan fluorescence and the signal is regained upon hydrolysis and product release (or possibly decay of alkylenzyme back to free enzyme and substrate). A clear quenching of fluorescence for both enantiomers of 5 was detected but no regained signal (after more than five minutes). Thus no, or undetectable, hydrolysis occur. Supporting the theory of E35 acting as a back-up base in H300N where the hydrolytic step could not be modeled due to the absence of the catalytic water in the crystal structure.

Product outcome is decided in the hydrolytic step

Experimental values for catalytic activities were reproduced within 1-2 kcal/mol by calculations. For wild-type, enantioselectivity is reproduced for the alkylation step, showing a lower energy for the preferred R,R-5 while hydrolysis is lower for S,S-5 thought to be coupled to an error in getting the correct reference reaction for the R,R- enantiomer. Regioselectivity cannot be seen experimentally since attack of either of the two ring carbons of this symmetric epoxide will result in the same product and thus no data is available to support calculations. When only energies for alkylenzyme formation is studied the preferred attacks differ between the two enantiomers, C1 for S,S-5 and C2 for R,R-5. However, the only pathway that leads to product formation is via C2 attack also for S,S-5 due to a too high energy barrier for the hydrolysis of C1-alkylenzyme.

The residues that seem to influence regioselectivity mostly in the hydrolysis step is H104 (destabilizing), W106 and D265 (both stabilizing) where the sum show a favored hydrolysis for alkylenzymes formed from attack at C2 for both enantiomers of 5. From electrostatic interactions it can also be seen that both residues of the E library (L266 and V267) show up as important residues for electrostatic contributions in the active site. This suggests the
inclusion of these residues to be a good choice, supporting continued work in introducing stabilizing mutations and re-visit library E.

Expanding the catalytic triad
A new, previously ignored but highly conserved, residue (H104) was found to be also important for catalysis. Further support can be seen from sequence alignments where a replacement of H104 is coupled to a replacement of also the adjacent E35, to compensate for the loss of positive charge. Thus H104 and E35 most likely form an ion pair and H104 is also close in space to the hydrolytic water and H300, suggesting the protonation states of these two histidines to be coupled (Figure 11). Calculations for wild-type enzyme verify the deprotonated state of H300 allowing it to function as a base. In order to provide charge balance in the active site H104 must be protonated. This is further supported by calculations where experimental data were reproducible for this form. Interesting is that this residue seems to have different protonation states for different enzyme variants, being neutral for both E35Q and H300N.

![Figure 11](image)

*Figure 11.* The active site with the additional residues E35 and H104 included.

The involvement of H104 in catalysis might be the explanation of the acidic shift in pH dependence seen for E35Q and $R,R$-5 where the pK$_a$ for this residue drops several units in calculations. For the wild-type enzyme there is a strong interaction between E35 and H104 that blocks solvent access to the active site. A similar behavior is seen for E35Q variant with $S,S$-5 where Q35 interacts with the tetrahedral intermediate formed after hydroxide attack. For $R,R$-5 the substrate pushes Q35 out of the active site thus allowing solvent access to the active site. This could be a clue to the peculiar pH pro-
file for E35Q where \( R,R\)-5 is affected in both alkylation and hydrolysis while the negative effects are smaller for \( S,S\)-5.

Calculations pointed out the involvement of yet another residue (H104) important for the mechanism of epoxide hydrolysis and its tight relationship with E35. Regioselectivity was decided in the hydrolytic step, thus, even though the preferred carbon for alkylation varied for the two enantiomers, product formation will in both cases only happen for C2 attack due to a kinetically blocked hydrolytic step for C1.
Calculations were performed to explain the enantioconvergent behavior for the wild-type enzyme with a smaller epoxide substrate, allowing for more than one binding mode. In addition, a modified variant with changed regioselectivity was included to see if also this behavior could be pin-pointed.

The same procedure as in Paper II was applied here for the QM/MM calculations with the smaller and non-symmetrical substrate styrene oxide (2). Also, regioselectivity is experimentally determined where the wild-type enzyme shows almost complete enantioconvergence. Calculations were performed to reproduce experimental data for both enantio- and regioselectivity. In an attempt to try to explain the underlying mechanism, something that will aid in construction of new enzyme variants possibly showing enantioconvergence also for other epoxides (Table 5). In addition to wild-type enzyme a modified variant (R-C1B1) was included due to its loss in regioselectivity with one of the enantiomers.

Table 5. Regio- and enantioselectivity shown by StEH1 wt and R-C1B1 with 2.

<table>
<thead>
<tr>
<th>Regioselectivity</th>
<th>Enantioselectivity</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>R-2</td>
</tr>
<tr>
<td></td>
<td>(%) Preferred carbon</td>
</tr>
<tr>
<td>StEH1 wt</td>
<td>89±0.8 C2</td>
</tr>
<tr>
<td>R-C1B1</td>
<td>55±10 C2</td>
</tr>
</tbody>
</table>

Smaller substrate allowing for more than one possible binding mode

Two stable binding modes were found (stable as the substrate staying in the active site) where the phenyl ring of 2 interacts either with the imidazole ring of H300 (Mode1) or with the indole of W106 (Mode2) (Figure 12). The involvement of W106 in regioselectivity was previously suggested by simply studying the crystal structure with the substrate manually docked and the idea of different binding conformers to be involved in selectivity was discussed in earlier work. To verify this speculation calculations for nucleo-
philic attack by D105 at both carbons and enantiomers with substrate bound in either of the binding modes were performed.

Figure 12. The two binding modes found to be occupied for styrene oxide. On the left, Mode1 where the phenyl ring of the epoxide interacts with H300 and to the right, Mode2 where the interaction is instead with W106.

Solved crystal structure add to accuracy
The accuracy of the calculations is aided by the solved crystal structure for enzyme variant R-C1B1. Compared to the wild-type structure there are no major differences except for the substituted residues W106L, L109Y, V141K and I155V where the former two seems to influence selectivity the most. Y109 is oriented away from the binding site thus increasing the volume of the active site while W106 is involved in Mode2 binding thus this mode is for L106 expected to be less stabilized. In addition, V141K, located at the entrance could affect availability of the active site but due to the low density of this side chain this cannot be stated.

Enantio- and regioselectivity is reproduced by calculations
Both wild type and R-C1B1 prefer S-2 to R-2, which was reproduced by calculations where in both cases energies were significantly lower for reaction with the (S)-enantiomer. The explanation lies in the change of preferred binding mode for the different enantiomers, where stabilization from amino acids lining the active site and the availability of the respective carbons for ring opening are important.

Mapping regioselectivity is a more complicated issue due to the different possible binding modes. Despite the alkylation being favored for ring opening at one carbon the energy for the subsequent hydrolytic step might be sufficiently high to be considered blocked. For wild-type enzyme and R-2 Mode2 is preferred where both alkylation and hydrolysis being energetically
favorable for attack at C2. For S-2 on the other hand Mode1 is preferred but essentially blocked for hydrolysis. Thus, also here Mode2 is the productive binding mode. Ring opening at C1 will be preferred due to lower energy for the alkylation despite the hydrolytic step being less favorable. Thus, here regioselectivity can be reproduced already in the alkylation step and not in the hydrolytic step as for epoxide 5.

A change in preferred binding mode for the enzyme variant with varied regioselectivity

For R-C1B1 and R-2, Mode1 was instead preferred where alkylation is more energetically favorable for attack at C2 while hydrolysis is equal with either of the alkyl enzymes. For S-2 binding of either Mode1 or 2 is possible with a slight preference for the former. In both cases, ring opening at C1 is preferred for alkylation but not for hydrolysis but this is still the most likely reaction path. Thus, regioselectivity was reproduced from calculated energies for S-2 but not for the R enantiomer.

What is seen for R-C1B1 is a shift from Mode2, seen for the wild type, to Mode1. This is most likely due to the loss of interactions between the phenyl ring of the epoxide and L106 after substitution. The active site is larger for R-C1B1 where additional binding modes might be available and indeed a third binding mode did appear during modeling however not stable enough to be included.

The importance of W106 in regioselectivity

When studying electrostatic and van der Waals contribution for wild-type enzyme W106 show major influence for the (R)-enantiomer in C2 attack, a possible explanation to the different regioselectivities for the two variants. Since this residue is mutated for R-C1B1 this might be the reason why a different mode is preferred. However, calculated energies still suggest C2 attack to be highly preferred while experimental data show attack at both carbons to be equally probable. Despite this failure in reproducing regioselectivity for R-C1B1 it is important to remember that simulations used for calculations are only studied for 10 ns and does not allow for the real reaction to occur. Also, the energy difference in attack at C1 and C2 is very small; a ratio of 100 in relative rates giving an energy difference of about 2 kcal/mol, comparable to the error in the calculations. Based on this, pinpointing the origin of this small effect is challenging.

The new findings, that different binding modes are preferred for different enantiomers and enzyme variants need to be considered when designing libraries for directed evolution. Also, even though one conformer is pre-
ferred for binding and alkylation the subsequent hydrolytic step might be too high in energy to be feasible.

*Two stable binding modes were found for styrene oxide in these enzymes where different modes were preferred for different enantiomers and for different enzyme variants. Calculations were able to reproduce experimental data for the enantioselectivity of both wild-type enzyme and a modified variant thereof. For wild type also regioselectivity and the concomitant enanti-convergence could be described via calculated energy differences for each step.*
Laboratory Evolved Enzymes Provide Snapshots of the Development of Enantioconvergence in Enzyme-Catalyzed Epoxide Hydrolysis

Here we set out to try to explain the regioselectivities seen with (2,3-epoxypropyl)benzene (I) for the set of enzyme variants isolated in Paper I. Crystals structures were solved for all these variants and substrate was docked into each active site to see if the preference in carbon to attack could be rationalized.

From Paper I, regioselectivities showed a strict attack at C2 with R-1 for all variants, while for S-1 there was a stepwise change from C2 for wild-type enzyme to C1 for R-C1, R-C1B1 and R-C1B1D33. This preference is reduced again for variant R-C1B1D33E6 that is back to wild-type behavior. Thus the high product purity is in this case from kinetic resolution while for R-C1B1D33 this is due to enantioconvergence. Residues W106L and L109Y were suggested to have the most influence in regioselectivity from structural estimations, based on speculations from what was seen upon manual modifications of wild-type structure. This involvement was further supported by work with epoxide hydrolase from another source where selectivity was altered upon modifications of this site. Since then, in Paper III, calculations had been performed for a structurally similar epoxide (2) where W106 was pointed out as highly stabilizing for one of the possible binding modes, adding to this hypothesis. Here, crystal structures for all enzyme variants have been solved and used for docking experiments thus adding more accuracy to speculations.

Crystal structures show a step-wise increase in active-site volume

Here we present three solved crystal structures (R-C1, R-C1B1D33 and R-C1B1D33E6). In addition, the inactive variant E35Q/H300N, described earlier, was co-crystallized with S-1 in an attempt to obtain a structure with a bound intermediate. Crystal structures were solved where some density in the active site could be seen, however, not to the extent that any conclusions could be drawn. Also, no covalent bond is seen between this density and D105 and the overall structure do not differ much from that of the wild type, suggesting that crystallization conditions need to be further optimized. It is worth mentioning that in the active site of R-C1 there is a glycerol, from the crystallization medium, positioned in between the tyrosines and the nucleophilic aspartate possibly indicating product binding due to structural similarities.
When studying the structures, the backbone carbons do not differ significantly in modified variants compared to wild-type enzyme. Further, there is no large difference in active site volume for R-C1 compared to wild type and these two variants are also similar in activity and selectivity. The R-C1B1 structure on the other hand shows an increased active-site volume due to Y109 being oriented away from the binding site upon interaction with N241. Upon modification F189L for R-C1B1D33, even more space is introduced in the active site. For R-C1B1D33E6, L266 is modified to a glycine, increasing the active-site volume even further. This increase in active site is not surprising since we see an increase in $K_M$ values for the modified variants with this substrate compared to wild type. This increase in $K_M$ goes hand in hand with the screen used for identification of these hits where a high concentration of substrate was used (as described in Paper I). Thus low $K_M$ was not selected for but rather a decent $k_{cat}$.

Slightly changed position of lid tyrosines

Other changes that could be seen for the modified enzyme variants are varying positions for the lid tyrosines Y154 and Y235. I155 is in close proximity to these (and H153) where for R-C1 the I155V exchange could potentially contribute to the positional flexibility observed for the lid Tyr phenol rings. This difference is slightly more pronounced for Y235 where in R-C1B1 it differs by 1.2 Å, compared to wild-type StEH1. This movement of Y235 is suggested to be due to W106L modification, allowing for an increased flexibility otherwise prevented by the bulky side chain. It should be mentioned that a dioxane molecule is present in the binding site, possibly influencing side-chain locations. Also for R-C1B1D33, the tyrosine positions are changed as the observed distances between the hydroxyl groups of Y235 and Y154 is measured to 5.0 Å compared to 4.4 Å in R-C1B1. This positional flexibility is rather minor in extent, but since the lid Tyr position the epoxide ring and facilitate its opening, even a small change may affect regio- and stereoselectivity as well as microscopic parameters of the catalyzed reaction, something that has been seen in other studies.86–88

Increased flexibility of the substrate lead to a new binding mode

As mentioned, W106 was thought to have major contributions to the changed changed regioselectivies due to its involvement in one of the binding modes for styrene oxide (Paper II). However, the docking studies here revealed that this binding mode was rarely adapted for the epoxide studied here, not even for wild type. Instead a new binding mode was revealed (Mode3) involving interactions with F33 and F189. This involvement of F33 could be an explanation as to why this residue was conserved when targeted for modification (Paper I) where both hits found showed silent mutations indicating this residue’s importance for activity. The new binding mode is
probably due to the increased flexibility from the methylene carbon connecting the epoxide and the phenyl ring of 1. When comparing $K_M$ for this substrate to the isomer trans-2-methylstyrene oxide the values are significantly higher for the former, suggesting that free rotation of the epoxide and the phenyl ring contributes to less binding stability. This suggests that it is actually the dissociation constant ($K_S$) that is increased, further supported by MD simulations where neither of the substrate enantiomers of 1 could be trapped in the active site without adding restraints. Also activity is lower with 1 compared to 3 probably also due to the increased flexibility.

Pre-steady state kinetics suggests a change in rate-limiting step and binding mode

To pinpoint the varying regioselectivities further, transient kinetic data was wished for to possibly decide the rate-limiting step for this set of enzyme variants. Since the residue thought to have the strongest effect in tryptophan fluorescence (W106, due to its position in the active site close to negatively charged D105) was modified for R-C1B1 and further evolved variants, tyrosine fluorescence was measured instead. Tyrosine fluorescence has a lower quantum yield and thus weaker intensities were expected. However, R-C1B1D33 did show detectable amplitudes and therefore tyrosine fluorescence was considered a good alternative also for R-C1B1. Despite clear amplitudes for wild type and R-C1B1D33 the rate of alkylenzyme build-up was too fast (within the dead-time of the instrument) to measure. For R-C1 and R-C1B1 instead no rates could be determined, probably due to inefficient build up of alkylenzyme. The different behavior in pre-steady state measurement could indicate a change in rate limiting step between the different mutants, from limiting hydrolysis in the wild type to limiting alkylation in variants R-C1 and R-C1B1, and back to limiting hydrolysis in the R-C1B1D33 variant.

Altered regiopreferences could partly be rationalized from docking

The low efficiency in binding and speculations about fast chemistry suggest docking to be a suitable instrument for visualizing regioselectivity, once a favorable binding mode is met the rest of the reaction pathway will proceed leading to product formation. This motivated the use of simple docking tools to predict preferred attack for alkylation to possibly reflect product outcome. Having real crystal structures add in accuracy for the dockings despite the lack of bound intermediates. The data in Table 6 shows that in most cases the regio preference matches the shortest distances between the nucleophilic D105 and the preferred carbon for attack. It should however be kept in mind that the preferred attack for alkylenzyme formation do not always lead to
product formation, as seen in Paper II and III where hydrolysis of the alkyl enzyme were in some cases blocked and product formation would only be allowed after attack at the least preferred carbon. In such cases conclusions of regioselectivity from substrate docking would fail. It should also be mentioned that what is seen from docking is that the preferred binding mode differs between R-C1B1 and R-C1B1D33 where the latter adapt Mode2, edge on Y109, possibly the explanation to why there is a detectable fluorescence signal here and not for R-C1B1. However, speculations like these should be confirmed by QM/MM calculations. Docking also proposed previously neglected residues F191 and F158 to interact with R-1 and Y183 and F301 with S-1, suggested to be targeted for future rounds of modifications.

Table 6. Regioselectivity and distances between nucleophile (D105) and the two carbons of the epoxide ring for the preferred binding mode seen in docking experiments.

<table>
<thead>
<tr>
<th>Enzyme</th>
<th>R-1</th>
<th>S-1</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>C1 (%)</td>
<td>rC1 (Å)</td>
</tr>
<tr>
<td>StEH1 (wt)</td>
<td>5.6</td>
<td>3.6</td>
</tr>
<tr>
<td>R-C1</td>
<td>3.0</td>
<td>2.8</td>
</tr>
<tr>
<td>R-C1B1</td>
<td>2.0</td>
<td>2.8</td>
</tr>
<tr>
<td>R-C1B1D33</td>
<td>2.0</td>
<td>2.6</td>
</tr>
<tr>
<td>R-C1B1D33E6</td>
<td>2.0</td>
<td>2.6</td>
</tr>
</tbody>
</table>

*Experimental data from Paper I. Calculated distance between D105 and C1/C2. *Model1’ involves interaction of the phenyl ring and H300, Mode2 W106 and Mode3 F33 and F189.

The overall crystal structures for the variants do not differ significantly form the wild-type where most changes, except for the gradually larger active site, was the slight movement of the lid tyrosines, possibly influencing the positioning of the epoxide ring and thus also possible effect on regioselectivity. The larger volume of the active site goes hand in hand with the fact that \( K_M \) is increased, which is an effect from the increased flexibility of the substrate and the screening set-up, where high epoxide concentration was used. Previously identified binding modes for styrene oxide was not occupied for this more flexible epoxide thus the direct involvement of W106 could not be seen. Instead a new binding mode including interactions with F33 and F189 was adapted. Further, docking experiments were in most cases able to point out the preferred carbon for alkylation to reproduce experimentally determined
regioselectivities but also here other steps on the reaction pathway could be involved in deciding product outcome. What docking does show are the features from having a more flexible substrate thus aiding in getting information out from the solved crystal structures. To try to map the reason for the changed regiopreferences seen there is a need to apply QM/MM calculations.
Complex Kinetic Schemes are Required to Describe Epoxide Hydrolase Catalyzed Production of Diols

Pre-steady state kinetics were performed for studying wild-type enzyme, R-C1 and R-C1B1 (previously isolated for their ability of producing enriched R-4 from a racemic mixture of 1) with trans-stilbene oxide (5) and the smaller trans-2-methylstyrene oxide (3) in order to reveal the microscopic rates to explain product outcome.

Crystal structures for all enzyme variants were reported earlier (Paper III and IV) and regioselectivities were analyzed in Paper I for epoxide 3 (Table 7). For epoxide 5, attack at either of the two carbons will result in the same meso-hydrobenzoin product and thus experimental data for regioselectivity cannot be obtained. Instead, calculations for wild-type enzyme in Paper II show that for alkyl-enzyme formation the preferred attack is C1 for S,S-5 and C2 for R,R-5. However, the only pathway that leads to product formation is via C2 attack also for S,S-5 due to hydrolysis of C1-alkylenzyme being essentially kinetically blocked. During this work complementing pre-steady state kinetic measurements were performed with both 3 and 5. Calculations did further show two different binding modes for more flexible substrates, structurally similar to 3 (Paper III), suggesting that they may also be present here.

Table 7. The regioselectivities for 3, shown in preferred carbon for attack and the percentage thereof.

<table>
<thead>
<tr>
<th></th>
<th>Preferred carbon</th>
<th>(%)</th>
<th>Preferred carbon</th>
<th>(%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>StEH1 wt</td>
<td>C1</td>
<td>&gt;99</td>
<td>C1</td>
<td>55±0.3</td>
</tr>
<tr>
<td>R-C1</td>
<td>C1</td>
<td>&gt;99</td>
<td>C1</td>
<td>66±0.6</td>
</tr>
<tr>
<td>R-C1B1</td>
<td>C1</td>
<td>&gt;99</td>
<td>C2</td>
<td>89±0.4</td>
</tr>
</tbody>
</table>

Non-strict regioselectivity motivates the use of complex reaction schemes

As can be seen from both product outcome and calculations regioselectivity is not strict, thus a branched reaction scheme is for sure needed to explain enzyme behavior (Scheme 3). The proposed formation of two different alkylenzymes for (amongst others) S,S-5 should ideally be seen from pre-
steady state kinetics. However, here only one rate was detected. Since the preferred attack through C1 is kinetically blocked this would allow for quenching of the signal due to build up of alkylenzyme. The productive pathway, via C2 alkylation, on the other hand is followed by a fast hydrolytic step thus possibly leading to less build up. This suggests that what we measure in the stopped flow could be not product formation, at least not alone, but dead-end alkylation. To complicate the situation even more the decay of alkylenzyme back to Michaelis complex for C1 attack is much lower than the one for C2, thus this might reduce the signal again. To sum up, what is seen in the stopped flow could be a mix of the different microscopic rates while we assume the numbers belong to one single pathway and this is important to remember when analyzing data.

Different binding modes demands a further complex reaction scheme

Preliminary simulation results together with literature^{143,144} suggest that the two binding modes from Paper III can also be available here for epoxide 3. The use of the branched reaction (Scheme 3) is convenient in trying to visualize different scenarios and explain product outcome. The detected new binding modes, however, may prove that this model should be revised to include more branches to cover also these. The different possible conformations for substrate binding do play a role in this case, something that is not seen with the bulky epoxide 5.

Pre-steady state kinetics do not reveal product outcome

For some of the multiple-turnover measurements the curves looked nice although the intercept was put to zero and no reliable values could be obtained, most likely due to the fact that a mix of different transient rates are measured. There are different scenarios that can occur, that need calculations to be confirmed and here I try to show the problems one stumble upon during these measurements using the branched reaction scheme mentioned earlier and presented in Figure 13.

In case (A) where there are two rates, two allowed alkylenzymes and two products formed the outcome is intuitive. If instead (B) two rates are seen but only one product is formed the explanation could be one of the hydrolytic pathways being kinetically blocked. Due to the fact that different binding modes can result in two different rates for attack at the same carbon yet another case (C) occur where two rates are seen despite only one alkylenzyme being energetically allowed for. This means that a branched scheme (D) like in all cases is needed for one production pathway where each branch represent one binding mode.
Figure 13. The branched reaction schemes used to explain various reactions scenarios that might arise. (A) two rates and two products are formed, (B) two rates despite only one product, (C) two rates and two alkylenzymes but only one carbon for attack meaning that two different modes are available that react at different rates. The last scenario (D) suggests one branched reaction scheme for each branch of the reaction in A, B and C due to two different binding modes forming the same alkylenzyme and product.

For $R,R-3$ we do know that we have varying degrees of regioselectivity for most of the enzyme variants studied, resulting in two products, thus already adding to complexity when measuring multi-turnover kinetics. In addition, noisy signals were obtained and in some cases clear signals, indicating a substantial build-up of alkylenzyme, but troubles in determining the actual rates. In order to pinpoint these behaviors there is a true need for QM/MM calculations. Some systems could, however, still be explained as seen in the following section, where R-C1 and R-C1B1 are compared to wild-type enzyme.

**R-C1 behaves similar to the wild-type enzyme**

R-C1 behaves similar to wild type and also in $k_{\text{cat}}/K_M$ for $R,R-3$ due to a 5-fold increase in both $k_{\text{cat}}$ and $K_M$ (all kinetic parameters are presented in Table 8). The intrinsic rate constants of alkylenzyme formation is increased for both $S,S-3$ and $5$, while rates for decay are similar where the energy barrier for alkylenzyme formation is lowered by around 1 kcal/mol. Also regioselectivity for $S,S-3$ is similar but for R-C1 only accumulation for one alkylenzyme is suggesting a linear reaction pathway and if not, the other rate for alkylenzyme formation is low enough to be hidden.

**RC1B1 deviate more from wild-type and R-C1**

R-C1B1 differ more from the wild type where $k_{\text{cat}}/K_M$ was decreased 30-fold for $R,R-3$, 200-fold for $R,R-5$, and a smaller decrease could be seen also for the (S)-enantiomers that could be explained by the increased active-site volume. Effects were expected also due to modifications of W106 where a main-chain amide interacts to stabilizing the tetrahedral intermediate and for a similar epoxide this residue was involved in stabilizing of one binding modes. This residue is therefore involved in both alkylenzyme formation and hydrolysis.
Table 8. Kinetic parameters for hydrolysis of 3 and 5.

<table>
<thead>
<tr>
<th>Enzyme</th>
<th>$k_{\text{cat}}$ (s$^{-1}$)</th>
<th>$K_M$ (µM)</th>
<th>$k_{\text{cat}}/K_M$ (s$^{-1}$×mM$^{-1}$)</th>
<th>$K_S$ (µM)</th>
<th>$k_2$ (s$^{-1}$)</th>
<th>$k_2$ (s$^{-1}$)</th>
<th>$k_3$ (s$^{-1}$)</th>
<th>$k_5 + k_{-5}$ (s$^{-1}$)</th>
<th>Regio-preference$^g$</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>R,R-3</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>WT$^{124}$</td>
<td>4.7±0.7</td>
<td>490±100</td>
<td>9.7±2</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>8.1±3</td>
<td>55 % C-1</td>
</tr>
<tr>
<td>R-C1</td>
<td>25±4</td>
<td>3300±10</td>
<td>7.7±0.2</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>66 % C-1</td>
</tr>
<tr>
<td>R-C1B1</td>
<td>0.92±0.4</td>
<td>3400±200$^c$</td>
<td>0.26±0.01</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>89 % C-2</td>
</tr>
<tr>
<td><strong>S,S-3</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>WT$^{124}$</td>
<td>63±3</td>
<td>77±10</td>
<td>820±100</td>
<td>470±100</td>
<td>370±20</td>
<td>170±20$^e$</td>
<td>110±10$^e$</td>
<td>32±2</td>
<td>99 % C-1</td>
</tr>
<tr>
<td>R-C1</td>
<td>65±2</td>
<td>110±10</td>
<td>560±40</td>
<td>1500±400</td>
<td>1900±300</td>
<td>120±30$^d$</td>
<td>72±20$^a$</td>
<td>44±7</td>
<td>99 % C-1</td>
</tr>
<tr>
<td>R-C1B1</td>
<td>31±1</td>
<td>240±20</td>
<td>130±6</td>
<td>2700±2000</td>
<td>1400±800</td>
<td>10±30$^a$</td>
<td>32±20$^a$</td>
<td>-</td>
<td>99 % C-1</td>
</tr>
<tr>
<td><strong>R,R-5</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>WT</td>
<td>17±0.5</td>
<td>16±1</td>
<td>1100±50</td>
<td>7.8±13</td>
<td>180±100</td>
<td>13±10$^d$</td>
<td>42±0.2$^a$</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>R-C1</td>
<td>12±0.5</td>
<td>20±2</td>
<td>610±40</td>
<td>13±17</td>
<td>85±3</td>
<td>22±40$^d$</td>
<td>28±0.2$^c$</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>R-C1B1</td>
<td>0.51±0.09</td>
<td>170±40</td>
<td>3.1±0.1</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td><strong>S,S-5</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>WT</td>
<td>4.0±0.06</td>
<td>1.0±0.08</td>
<td>3900±200</td>
<td>5.5±0.9</td>
<td>23±0.7</td>
<td>4.5±0.8$^d$</td>
<td>3.9±0.005$^c$</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>R-C1</td>
<td>4.4±0.09</td>
<td>1.3±0.09</td>
<td>3300±20</td>
<td>16±5</td>
<td>120±9</td>
<td>4.2±6$^d$</td>
<td>5.4±0.01$^c$</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>R-C1B1</td>
<td>4.0±0.09</td>
<td>16±0.9</td>
<td>250±0.09</td>
<td>21±10</td>
<td>8.9±1</td>
<td>5.1±0.08$^a$</td>
<td>3.0±0.01$^e$</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

$^a$Calculated from the steady-state expression for $k_{\text{cat}}$ and the determined values of $k_{\text{cat}}$ and ($k_{-2}$+$k_3$). $^b$The product is a meso diol and, hence, the different product enantiomers are indistinguishable. $^c$Experimentally determined in single-turnover experiment. $^d$Value calculated from the extrapolated value of ($k_{-2}$ + $k_3$) from multiple-turnover experiment and the experimentally estimated $k_3$ from single-turnover experiments. $^e$Experimentally determined in single-turnover experiment but with not fully saturated enzyme. The determined value is therefore $\leq k_3$. $^f$The product is a meso diol and, hence, the different product enantiomers are indistinguishable. $^g$Data from Paper I.
When comparing microscopic rates of R-C1B1 to those of R-C1 the alkylenzyme formation rate for S,S-3 is similar while both rates of decay ($k_2$ and $k_3$) are lower here. This suggests a more stabilized alkylenzyme and a lower turnover. For S,S-5 the alkylation rate is 15-fold lower while the decay is similar that could explain the 10-fold increase in $K_M$. R,R-3 was modeled into the active site and the docking pose suggests that its interaction with the side chain of L109 changed upon modification to Y. This tyrosine side chain is turned away from the active site, allowing for R,R-3 to bind also in Mode1, pointing towards H300. This conformer allows for attack on C2, which can explain the shift in regioselectivity seen for R-C1B1. The modeling also indicate that the side-chain of L106 is shifted towards the substrate binding site, which could cause steric clashes with some of the binding modes and also hinder attack at C1.

Single-turnover experiments suggest a change in rate-limiting step

Single-turnover experiments result in the most well-determined rate constants for $k_3$. However, experiments like these demands a low dissociation constant ($K_S$) to allow for saturation of the enzyme where substrate concentration should be at least 10-fold higher than $K_S$ and enzyme concentration a somewhat higher than substrate concentration. We were able to run single-turnover experiments for both enantiomers of 5 with wild-type enzyme and for S,S-5 with R-C1 and R-C1B1. For S,S-5 the values for $k_{cat}$ and $k_3$ are similar and therefore hydrolysis is clearly the rate-limiting step. For R,R-5 on the other hand $k_3$ values are substantially higher than $k_{cat}$ instead suggesting another step to also be rate limiting. What is seen from calculations (Paper II) is that there is one clearly preferred carbon for alkylation of about the same energy as the subsequent hydrolysis thus the rate-limiting step must be earlier on the reaction pathway.

Different binding modes lead to possible problems in computational reproduction of experimental data for pre-steady state kinetics

QM/MM calculations are suggested as a solution for many of the challenges stumbled upon here, to get a clearer picture of the microscopic rates and dependencies. It is important to keep in mind that the build-up of alkylenzyme for pre-steady state measurements relies on the formation being much faster than the decay. When trying to reproduce experimental pre-steady state data by theoretical modeling this will lead to difficulties, one example being the already discussed wild-type reaction with S,S-5, where the productive route does not include the preferred alkylation and thus the signal we measure in the stopped-flow apparatus may not necessarily be coupled to
product outcome. Also, experimentally it is not possible to distinguish between two different $k_2$ or $k_3$ rate constants and what is monitored is a mixture if more than one contributes to the observed decay rate.

Pre-steady state kinetics fail in describing the actual product outcome due to a mixture of reaction pathways and formation of two products. In some cases rate of alkylenzyme formation is too fast leading to troubles in catching the real rate. In others, signals are too weak probably due to insufficient build-up of alkylenzyme suggesting that hydrolysis might no longer be the rate-limiting step or what we measure is not purely product formation. This complicated story would for sure benefit from QM/MM calculations where there is a possibility to map both the different binding modes and the preferred carbons for attack but also steps that are kinetically blocked.
Integrated action of *Solanum tuberosum* epoxide hydrolase I and *Rhodococcus ruber* alcohol dehydrogenase A in a two-step biocatalytic reactor

Two enzymes were immobilized on separate aluminum-oxide membranes to perform a two-step biocatalytic reaction. This was done to study the potential use in an industrial process, where immobilization is a beneficial approach both for stabilization of enzymes in organic solvents and for recycling of the catalysts.

My work is focused on one enzymatic step included in a larger reaction network (Figure 14) with the aim combining different enzymes to co-function in one pot. The product diols are useful chiral building block with both hydroxyls having similar reactivity thus making selectivity somewhat hard to direct.

![Figure 14](image)

*Figure 14.* The reaction starting from an epoxide that is hydrolyzed by epoxide hydrolase to a vicinal diol. The next step involves oxidation by an alcohol dehydrogenase, either specific for primary hydroxyl groups forming aldehydes or for secondary hydroxyl group forming ketones. The latter is the one studied here, involving ADH-A. For this study R = phenyl group.

The two-step reaction studied here starts from styrene oxide (2) that is hydrolyzed to phenyl-1,2-ethanediol (6) by the engineered StEH1 variant R-C1B1D33. An aldoholdehydrogenase, selective for secondary alcohols was used for the second step, to form 2-hydroxyacetophenone (7) (Scheme 4).

In addition to serve as a base for optimization of this two-step biocatalytic pathway the reaction also refines the cheap starting material in racemic styrene oxide to the more valuable product 2-hydroxyacetophenone.
Scheme 4. The reaction from styrene oxide (2) to phenyl-1,2-ethanediol (6) and further to 2-hydroxyacetophenone (7). The first step catalyzed by epoxide hydrolase and the second by alcohol dehydrogenase.

The alcohol dehydrogenase of choice is ADH-A from Rhodococcus ruber and catalyzes the reversible reduction/oxidation of ketones and secondary alcohols. It belongs to the zinc-dependent group I alcohol dehydrogenases and is a suitable enzyme for biocatalysis due to its tolerance to high concentrations of isopropanol and acetone. This tolerance makes it possible to use acetone for co-factor regeneration, however, the formed isopropanol is a preferred substrate as compared to the diol formed in step I (Scheme 4) thus engineering is required to reduce its activity with isopropanol. ADH-A was also shown to oxidize bulky substrates that are of importance for this project.

Optimization of reaction conditions of free enzymes

First, the reaction conditions were optimized for the two enzymes to co-function free in solution. Green thinking and theoretical scale up was kept in mind, thus ammonium bicarbonate buffer was used that can be evaporated. pH optima differ for the enzymes and one has to make certain that the correct conditions for the oxidation/reduction reaction are present. Since epoxide hydrolase shows high activity for a rather broad pH range while ADH-A displays rather low activity with this substrate already at its pH optimum of 8, this was the chosen pH.

Co-factor regeneration & Inhibition problems

For co-factor re-generation two different systems were studied and are presented in Figure 15 where (A) reduction of acetone by the internal ADH-A and (B) reduction of pyruvate by lactate dehydrogenase (LDH).

When using an internal enzyme in ADH-A for co-factor regeneration, a clean system is obtained, however, with the drawback of competition for the active site, where in this case both acetone and isopropanol are preferred substrates over diol 6.

The alternative of an external enzyme, which possibly avoid the competition for the ADH-A active site but adds the problem of an additional pH optimum
to take into consideration and additional components added that could possibly complicate downstream processing.

![Figure 15](image15.png)

**Figure 15.** The alternative systems for regeneration of NAD\(^+\) studied during this work: (A) ADH-A reduction of acetone to isopropanol and (B) LDH reduction of pyruvate to lactate.

The second alternative includes an external enzyme, LDH, where pyruvate is reduced to lactate. Lactate showed to be a minor inhibitor of ADH-A and, to a larger extent, also for LDH, leading to difficulties in obtaining high ketone concentrations. Also, the product of the main reaction itself, 7, was previously shown to inhibit LDH\(^+\). To be able to use this system for co-factor regeneration the ketone concentration needs to be kept low by constantly withdrawing product from the reaction mixture. So far, the LDH alternative is the best method for ketone production where a maximum yield of 40\% was seen for a reaction with 33 mM 2 as starting material (Figure 16).

![Figure 16](image16.png)

**Figure 16.** Ketone (7) formation with co-factor regeneration for acetone as external electron acceptor (dark grey) and LDH and pyruvate (black).

Since NADH is known to inhibit ADH-A it is important to include a co-factor regeneration system to keep this concentration low. Also 7 is a potential inhibitor thus a two-phase system where ketone can be removed upon formation is probably beneficial, particularly since this product is suggested to react with NAD\(^+\). Despite only two or three enzymes being included to co-
function there is already a challenge in finding optimal reaction conditions due to inhibition and different pH optima, resulting in a rather complex dependence network. When thinking of the environment and the cost-efficiency of the process acetone would be the best alternative, out of these two, since the system tolerates high acetone concentrations, acetone is cheap and an internal enzyme is used.

Characterization of immobilized enzymes

Enzymes were immobilized to alumina oxide membranes (Figure 17) and characterized one by one before being placed in a sequential mode. 

![Figure 17](image)

Figure 17. The nanoporous alumina-oxide membrane and a schematic view of the coupling of StEH1 to the linker via a lysine residue of the enzyme.

The same reaction conditions were applied for the immobilized system except that for initial studies no co-factor regeneration was applied. This allowed for product outcome to be monitored by following the formation of NADH by UV/VIS spectrophotometric measurements at 340 nm. What was seen from the results when comparing free and immobilized enzymes was a reduced activity for both R-C1B1D33 and ADH-A. A larger decrease for ADH-A could be expected due function being dependent on a dimeric enzyme. The two enzymes, R-C1B1D33 and ADH-A, were able to co-function also in a sequential mode although at lower rates. 6 is known to be a poor substrate for ADH-A thus enzyme engineering should be applied to increase activity.

Both epoxide hydrolase and ADH-A are active upon immobilization and the two enzymes did also function in sequential mode to generate product. Immobilization is not only beneficial for a possible increase in stability and the easy re-use of the catalyst and downstream processing but also to allow for tuning the reaction conditions in between the enzymes. Thus problems with optimization is possibly reduced where in the case of using acetone for co-factor regeneration this can be added after the epoxide hydrolase.
Conclusions & Future perspective

Great progress has been made towards understanding of the underlying mechanism of enantio- and regioselectivity shown by StEH1 and variants thereof. Theoretical models showed to be successful in reproducing experimental values for activity and selectivity but we are still far from the ultimate goal of being able to predict which modifications to introduce for a certain selectivity. What calculations did produce was suggestion of various binding modes and information of possible kinetically blocked steps of a reaction pathway. Thus, the combination of experimental data, crystal structures and theoretical modeling has been crucial for trying to understand these complicated systems.

We managed to get crystal structures for most of the enzyme variants isolated, which together with experimental kinetic data have improved the theoretical models used to simulate the catalytic reaction. Yet, assumptions still have to be made due to lack of structures with reaction intermediates bound, something that is desirable in order to get closer to the actual structural background to selectivity. An enzyme variant with inactive hydrolytic ability (E35Q/H300N) was constructed to serve this purpose, but so far no solved structure did contain a clearly bound intermediate. Structures are one important key for understanding structure-function relationships thus further optimization for crystallization would be useful. In addition, this set of modifications was introduced also to the other isolated enzyme variants from Paper I, and these should also be characterized and used in co-crystallization experiments. Another approach could be to modify H104, a residue previously not considered to be involved in catalysis and therefore not well studied.

Enzyme variants isolated here did show a clear enrichment of product enantiomers but for an enzyme to be useful in e.g. pharmaceutical industry the purity needs to get higher. Further in vitro evolution to search for variants with improved but also varying selectivity are wished for. For the furthest evolved variant (R-C1B1D33E6) six mutations were introduced. The hit rate was decreased and before moving forward to additional mutagenesis at the active site, stabilizing mutations should be introduced. It could be favorable to do an inventory of the introduced mutations where V141K has shown to have only minor effects on product outcome and this should be investigated in more detail. Going back in the evolutionary tree and restart from a revert-
ed variant would be interesting, to see if other mutations would be possible. Yet another benefit from reverting this mutation is to avoid the trouble this side chain cause for structure solving, due to its high flexibility.

To minimize library size, a degenerated codon set (NDT) was used allowing for twelve out of twenty amino acids. It would be interesting to also include the remaining eight by using an unbiased set of primers for all 20 amino acids. The number of variants would increase from $12^2$ to $20^2$ for libraries of two residues, which would also be possible to screen. However, an alternative screening method should be investigated allowing for a higher throughput and possibly to directly getting the selectivity and activity out. This would allow us to identify and characterize also variants with decreased activity or selectivity something that is highly informative in mapping structure-function relationships. A suggestion could be to couple the screening reaction to ADH-A to be able to measure NADH formation by following absorbance at 340 nm giving directly a measure of the efficiency of the two-step reaction.

In industry, the use of the same enzyme for several reactions would be beneficial. As a step towards this the use of the same libraries for screening also other substrates could be helpful as well as to measure activity for the enzyme variants isolated as hits here with a range of epoxides, to see what could be fished out. Here, the immobilized enzymes could be used in a sequential mode to detect epoxide hydrolase activity indirectly by measuring NADH formation for ADH-A.

I would also like to continue the work with the coupling of StEH1 and ADH-A, to optimize co-factor regeneration and other reaction conditions and try other combinations of variants and substrates. Further, it would be interesting to continue the work with co-expression of the enzyme in E. coli or to express them on the surface to avoid possible problems with transporting the epoxide in to the cell. I would like to do more research about present industrial processes and what parameters to consider when scaling up. I would like to find potential industrial processes including a step of epoxide hydrolysis where StEH1 could replace existing catalyst to improve enantiomeric excess of product and thereby the yield.
Populärvetenskaplig sammanfattning
Mot förståelse av selektivitet & enantiokonvergens hos ett epoxidhydrolas

"Jag får enzymer att göra saker de egentligen inte kan så att de skulle kunna användas för framställning av rena produkter i exempelvis läkemedelsindustrin."

Detta är vad min beskrivning har kokat ner till med åren. Att förklara att man jobbar med biokatalys säger inte folk i allmänhet vidare mycket. Vad är biokatalys? En katalysator vet många vad det är, man har det i bilen. Det är något som snabbar på reaktioner utan att själv förbrukas – mycket effektivt! Bio syftar till att det är naturens egna katalysatorer, i form av hela celler eller rena enzymer, och de kan snabba på reaktioner upp till miljarder gånger genom att positionera molekyler (så kallade substrat) perfekt för reaktion och genom att stabilisera vissa labila mellanprodukter (Figur 18). Enzymer är kedjor av aminosyror och är fantastiska små varelser som finns överallt; i naturen, i djur och människor där de ser till att reaktioner sker, lagom fort och ofta. Det fina är att man kan använda dessa utanför deras naturliga habitat för att snabba på reaktioner av olika slag inom exempelvis läkemedelsindustrin.

![Figur 18. En schematisk bild över hur enzymet epoxid hydrolas omvandlar substrat till produkt.](image)

Varför är biokatalys viktigt? Vi blir mer och mer medvetna om att vi måste värna om miljön och hushålla med jordens resurser. En stor industri som definitivt behöver göras mer miljövänlig är kemikalieindustrin där nuva-
rande processer genererar miljöfarligt avfall och ofta kräver extrema temperaturer och tryck. Man försöker på många vis göra dessa processer grönare genom att exempelvis försöka utföra reaktioner i vatten istället för i organiska løsningsmedel. Här kommer biokatalys och enzymer väl till pass då dessa redan arbetar optimalt i vattenlösningar och vid skonsamma temperaturer samt pH. Utöver detta är de biologiskt nedbrytbara och de flesta enzymer har liknande optimala arbetsmiljöer så att flera kan samverka i ett och samma kärl, något som reducerar mängden avfall samt antalet reaktionssteg.

Enzymer är även, som titeln säger, selektiva i att de ofta är utvecklade för att utföra en viss reaktion med ett visst substrat. Detta kommer ifrån att enzymer har ett så kallat aktivt sätte där reaktionen sker och detta är oftast skrådarsytt för ett visst substrat vilket innebär att de kan välja ut ett visst molekylslag från en blandning. Denna egenskap gör processerna mer effektiva då den till och med möjliggör framställning av enbart den ena av två spegelbildsmolekyler, något som är eftertraktat i läkemedelsindustrin (Figur 19).

Figur 19. Två spegelbildsmolekyler eller enantomerer som de kallas.

I de fall då enzymet enbart katalyserar reaktion för den ena substratmolekylen (kinetisk upplösning, övre delen i Figur 20) så erhålls ett utbyte på maximalt 50 % från en substratblandning. Detta innebär i och för sig att en ren produkt erhålls men även att hälften av ursprungsmaterialet slösas bort. Man kan blanda olika enzymer med kompletterande egenskaper för att överkomma detta men det finns även enzymer (däribland huvudrollsinnehavaren i denna avhandling) som har den fantastiska egenskapen att kunna göra en och samma spegelbildsprodukt från en blandning av två spegelbildssubstrat, ett fenomen som kallas enantiokonvergens (nedre delen i Figur 20).

Denna förmåga är viktig då de två produkterna kan, även om de egentligen är väldigt lika, ha helt olika egenskaper. Ett exempel är smakerna citron och apelsin (S- och R-limonen). Särskilt viktig är selektivitet när det kommer till läkemedel där det oftast bara är den ena av de två spegelbildsmolekylerna som har den önskade effekten medan den andra kan vara rent av skadlig. Så var fallet med neurosedynincidenten på 1960-talet, då gravida kvinnor rekommenderades en lugnande medicin som bestod av en blandning av två spegelbildsmolekyler där den ena visade sig ha fosterskadande effekter. Nu för tiden har man kännedom om de ofta skilda egenskaperna och ser till att
rena upp de två spegelbildsmolekylerna ifrån varandra. Alternativt kan man försäkra sig om att den andra är ofarlig, men av ekonomiska skäl önskar man så klart att framställa enbart den önskvärda molekylen, och särskilt på ett enantiokonvergent vis.

Figur 20. Överst visas kinetisk upplösning och under enantiokonvergens, pilen indikerar var attacken för reaktion sker.

Förr i tiden extraherades enzymer ur sina naturliga källor, i mitt fall potatisblasten, vilket kunde vara både tidsödande och ge ett lågt utbyte. Men med dagens rekombinanta genteknik kan man konstruera genen som kodar för det enzym man önskar och få en bakterie att uttrycka detta för att sedan renar upp enorma mängder enzym på ett enkelt vis - mycket revolutionerande! Följaktligen utgör själva tillgången på enzym oftast inte längre något hinder utan vad som är svårt är istället att hitta ett enzym som gör precis det man önskar. Naturens enzymer har utsatts för selektionstryck under miljontals år i syfte att effektivt utföra en viss reaktion och detta försöker vi återskapa under kontrollerade former genom att utföra evolution i provrör. Vi använder så kallad riktad evolution där vi snabbar på evolutionen genom att modifiera ett enzym i syfte att skapa nya egenskaper.

viktigt att försöka förstå mekanismen bakom den nyvunna egenskapen, något som kartläggs med den typ av grundforskning som bedrivs i akademin.

I syfte att kartlägga förekommande struktur-beteende-samband har vi kombinerat experimentella data och kristallstrukturer av enzymer med beräkningar. I detta sammanhang har beräknings kemisterna till uppgift att ta fram teoretiska modeller som beskriver reaktionen där de utgår ifrån tillgängliga enzynstrukturer och experimentella data för att träna modellerna. Under förutsättning att detta lyckas för en mängd olika enzym-substratkombinationer är tanken att, med hjälp av datorn, kunna förutsäga vilka modifieringar som krävs för att få en viss aktivitet eller selektivitet även för andra, liknande molekyler.

I mitt arbete började jag med att göra tusentals varianter av epoxidhydrolas från potatis (StEH1), ett enzym som består av 321 aminosyror. Målet var att skapa en modifierad variant som endast framställer den ena av två spegelbildsmolekyler. I detta avseende lyckades jag identifiera flertalet nya enzynvarianter med nya selektiviteter genom att bara byta ut upp till sex stycken aminosyror. Vissa var selektiva genom att bara reagera med det ena av de två spegelbildssubstraten medan andra visade sig vara enantiokonvergenta som beskrivits ovan. Med andra ord är det möjligt att producera en och samma spegelbildsmolekyl från båda substraten. Härvid kan tilläggas att dessa enzymvarianter visade sig fungera bra även med andra, liknande substrat vilket är önskvärt om enzymet ska användas i industrin.

Vi vet sedan tidigare hur strukturen för epoxidhydrolas från potatis ser ut och lyckades här även få strukturer för de modifierade varianterna (Figur 21). Utifrån dessa var det möjligt att härleda hur substratet band och vilka aminosyror som eventuellt bidrog mest till de nya egenskaperna (Figur 21). För att få en djupare förståelse för detta så tillämpades ovannämnda beräkningar, där beräknings kemister tillhandahöll underlag i form av kristallstrukturer och siffror för enzymernas beteende med olika substrat med avseende på katalytisk förmåga och selektivitet. Genom dessa beräkningar identifierades ytterligare några aminosyror som är involverade i mekanismen, vilket har bidragit till en djupare förståelse för hur reaktionen går till och hur enzymet bidrar till processen. Darutöver har vi, i vissa fall, kunnat identifiera vilket steg i reaktionen som bestämmer selektiviteten och vilka aminosyror som har störst inverkan på både selektivitet och reaktivitet. Så det är möjligt att med hjälp av datorn reproducera de data som erhållits i laboratoriet men ännu är det långt kvar till det långsiktiga målet att underlätta det praktiska arbetet genom att tillämpa teoretiska beräkningar för att förutsäga vilka modifieringar som krävs för att få en viss egenskap hos ett enzym men vi är på god väg. De teoretiska modellerna behöver tränas mycket mer, med mer data för fler varianter och fler substrat. Dessutom försöker vi att få fram strukturer med substrat bundet i för att se hur det verkligen ser ut när detta sker.
Detta skulle inte bara rent visuellt bidraga till förståelsen utan också under- 
lättan mer för beräkningskemisterna och på så vis leda oss an närmare en 
kartläggning av struktur-funktions-sambandet.

Figur 21. Till vänster ses strukturen för StEH1 där de aminosyrorna som är involve-
rade i katalysen är färgade svarta, de visas även förstörade till höger i bild. Man kan 
hänfri se att de katalytiska aminosyrorna är väldigt få i jämförelse med hela prote-
inet (321 stycken).

Sålunda har vi skapat en verktygslåda med en mängd strukturer, teoretiska 
modeller och enzymvarianter att valja ifrån för att utföra reaktioner med 
olika aktivitet och selektivitet för olika substrat. Enzynerna har modifierats 
och studerats i detalj (ett och ett) för att sedan kunna kopplas ihop med andra 
enzymer i syfte att utföra flerstegs-syntes. Fördelen med att ha rena enzymer 
ar att man själv sätter ihop dessa för att få just den reaktion man önskar och 
på så vis slipper andra störande metabola vägar som uppstår vid användning 
av hela celler. Fördelen, å andra sidan, med att ha hela celler är att cellen 
själv reglerar produktion av andra molekyler som kan vara nödvändiga (och 
dyra att framställa) för en effektiv reaktion.

Slutligen så tror jag inte att enzymer kommer att ersätta alla steg i kemisk 
industri utan att man låter de två alternativen samverka. Enzymer kan med 
fördel användas för de steg där selektivitet är ytterst nödvändig medan traditionell 
organisk syntes kan behållas för de reaktionssteg som fungerar väl 
utan användning av höga temperaturer eller kemikalier med negativ påver-
kan på miljön. Det bedrivs mycket forskning i syfte att göra den kemiska 
industrin mer miljövänlig men man använder fortfarande huvudsakligen 
organiska lösningsmedel så för att enzymer ska kunna implementeras i de 
existerande processerna krävs att de modifieras för att överleva i denna ona-
turliga miljö. Ett stabiliseringsalternativ innebär bindning av enzymet till 
tunna membraner, en smart metod då det enkelt går att återanvända systemet 
för andra reaktioner. Jag tror på att vi måste utnyttja det bästa av båda värl-
dar, ett samarbete mellan handgjorda och naturens egna katalysatorer, för en 
hållbar industri.
Acknowledgement

There are so many people I would like to thank, both inside and outside the BMC-building. Thank you all! Now I am ready for a life in the real world and hopefully people will benefit from what I have learnt here. Yey!

First I would like to thank my supervisor Mikael Widersten for allowing me to think my life through before I jumped on this train. For trying to guide an engineer with a what-could-we-use-this-for mind into the fascinating world of enzyme kinetics. For your unbelievable fascination for everything we do, being it a perfect SDS-PAGE gel or listening to the sound of stopped-flow for a whole afternoon. For your ability to always see possibilities in all results even though me myself sometimes lost all hope. Your enthusiasm has been invaluable.

I would also like to thank my co-supervisor Helena Danielson for always encouraging me to keep my engineering-hunger. For helping me conclude what I want and what I need to do to fulfill that when my own mind was sometimes a bit blurry. Thanks for pushing me forward in your own slightly brutal but caring way.

I would like to thank my co-workers Lynn, Paul, Beat, Fernanda, Doreen, Mikael N, Huan, Gunnar & Erika without you guys this thesis would have been awfully thin. I could not thank you enough!

Thank you ÅForsk & Liljevalchs & Wallenberg for allowing me to go to America, Japan and Hong Kong to present my research (and to do some sightseeing and cute-stuff shopping). I have learned a lot from interacting with researchers from all over the world.

I am also grateful for the help in the proofreading of this piece of work by Huan, Rickard & Christopher your comments highly improved the quality of the story.
and now i will completely skip grammar and spelling and first give a huge thank you to the widersten group. my second family for the last five or so years. m i c k e for that i lärde mig hur man labbar. eller hur man inte labbar. för att du alltid får mig att skratta. trasiga glas(ögon) och soppa i väskan. för att du antingen köpte dig en hjälm. på rea för att du aldrig slapp Per taget om din kaffekopp. forskningsidéer. favoritskjortan. nyfikenheten. radarbilderna. för att du alltid är positiv. för dina stora visioner (trädäck, JACS, babylskustralor till hundar). för att du skapar en fantastisk atmosfär i gruppen. som får en att känna sig som hemma från första stund. för att du höll dig mer informerad om min fotboll än jag själv. och för att du hjalpte mig att ro detta i hamn. a n n & d i a n a för att ha banat väg. lite svårt att axla era mantlar bara då det ställdes riktigt höga krav på mig. diana, sättet du presenterar en på är förundrande. c i s i for klättringsupplevelser. och för alla resor. dödsturer på manhattan. 4th of july-firande i hål. som gud glömde (rosetten är fortfarande den finaste present jag nånsin fått och sura tanten kan fortfarande ge mig rysningar). farglada flingor på super8. nakentur i japansk taxibil (fortfarande det syndigaste jag gjort). det kändes stort att släppa iväg dig till amerika. utan att bara vare sig ditt pass. din packning. eller dig. h u a n min glad häst. for all unforgettable moments. always with a cute touch. for afro classes. for being my tea supplier. wherever you go i will come visit you. for all etsy inspiration. for your fighting spirit. for always being there. in sad and happy times. future is soon. with awesome jobs star homes and whatever you wish for. for you and your petit prince. c i s i for klättringsupplevelser. and for being my tea supplier. wherever you go i will come visit you. for all etsy inspiration. for your fighting spirit. for always being there. in sad and happy times. future is soon. with awesome jobs star homes and whatever you wish for. for you and your petit prince. c i s i for klättringsupplevelser. and for being my tea supplier. wherever you go i will come visit you. for all etsy inspiration. for your fighting spirit. for always being there. in sad and happy times. future is soon. with awesome jobs star homes and whatever you wish for. for you and your petit prince.
have been what it is today. thank you • d o r e e n • for teaching me and guiding me all the way through the crystallization experiments • and for always having time and showing enthusiasm and for baking awesome cakes. you will become a supermom • d i r k • for taking care of my monster crystals • now, let us make them get you a laboratory engineer. everything will be fine. you will have a kick-ass introductional part • g u n n a r j • for at you vet allt. institutionens levande uppslagsverk. tack for all ovärderlig hjälp med immobiliseringsmanuskriptet. och • e r i k a • min stjärna. allt jobb du gjort for vårt lilla manuscript. tack! for att jag hittade min smörkompis i japan. jag kommer aldrig att glömma vår fantastiska resa. for all spinning och pump och andra campus-event. for alla viktiga samtal och sms-trådar. tack for allt engagemang i jobbkajken. hej linus! tack for leenden. födelsedagspresenter. och glöm inte att fortsätta med piff även efter minis ankomst. ser fram emot dagslånga fikor med våra kids. arbetslösheten kommer att vara en dröm ☺ • l y n n k a m e r l i n & g r o u p • for the endless hours of calculations and re-calculations. a special thank you to • p a u l i e • (i know you have always loved your nick name) for answering my questions over and over again. and again. hang in there ☺

i would also like to thank all the past and present members of the HD group. in particular • t o n y • for introducing me to climbing. and the hilarious world of german-swedish. haha. for all nice moments in uppsala and during trips. for your easy-going lifestyle. for laughing at me for not having friends here when I first arrived • e l d a r • tjenamors din gamla galosch! sorry for leaving you hanging but you turned out just fine. new job and all. commuting is fun! • c h r i s t i a n s • you and your muffins and LateX • j o h a n w • for den hemliga posten. för att du förgyllde tillvaron tillsammans med royalisten • s o f i a • och all-time-high-skattets ägare • a n g e l i c a • • s a r a s • för alla snickesnack i stopped flow-rummet. hoppas att ni har det bra i danmark.

thanks to the small but naggande goda YI group • g u s t a v • så ung. du har låååång tid kvar. tack for att du inte övergav oss vid lunchbordet. de ibland något opassande samtalsämnena till trots. men många gånger var det faktiskt du som började. johol! • y l v a • cool som få. å nyfiken. skönt att det kommer in lite unga forskare med livsgnistan i behåll ☻ å tack for att du tar på dig frågeställarkoftern.

i would also like to post a big thank you to all past and present organic chemists from the dark ■ and the bright side. for all hej hejs in the corridor • r i k a r d • for alla tips och goda råd. for alla lunchsamtal och mislånga facebook-konversationer med dig och @johanna. för att du tittar förbi då och då för att se att allt går som det ska. för att du hejade mig hela vägen över mållinjen. nu kan du läsa den svenska sammanfattningen ☻ när ska jag komma till nya boet med skumpa? • a n n a • för att allt är så enkelt. aldrig
några problem. för att du gör det du tror på och struntar i vad andra tycker. vi får ta en playdate allihopa snart! • s a r a • för all inspiration. sagolika foton. unge. allt. kanske att du tar det lite lugnare nu och inte hetsar upp dig för att någon cyklar en liten bit in i cykelbanan? tack för alla tips för att inte bli runerad under mammaledigheten. verkligen. • m a g n u s • för att du var i samma sits som jag. på många vis. skönt att ha någon att prata med. mitt hjärta är för evigt ristat av våra du-vet-vad. så mycket skitsnack har jag aldrig fått ur mig gällande en och samma sak. vad hände med vårt samarbete? jag har fortfarande en box med photoswitchisar i min frys. men fortfarande ingen signered avhandling å la blom • j o h a n v • välkommen hit. välkommen hem. ditt ständiga leende och fniss var saknat. känns bara lite ovant att vi inte kan bestämma över ditt huvud längre. men man vänjer sig nog. i sinom tid. du borde alltid ha med dig lite fairy dust i fix. • • • • c h r i s t i a n • vår egen cover-boy. måtte du hitta en ny pizzeria. å det är alls inget fel på åttiotalsmat. kasserlåda. köttfärslimpa. å lite after eight-paron på det. • t o b i a s • för alla lunchspinningpass. för pendlingsförståelse. för konferenssällskap på andra sidan jorden och för jobbmedlade • a l e x a n d r a • för sharing the stress and reminding me that life is good and everything will turn out just fine • c a r i n a • för nice walks to ångström and excellent teaching company • t h o m a s n • • • för att du förstått vikten av fika och karleken till råsunda. sushi & choklad – det håller magen glad.

For all people who made teaching a true pleasure • f r a n ç o i s e • alltid positiv och glad. och omtäcksam som få. leden för ditt första möte med • • • • n i får göra upp det där någon dag. merci för att du finns. som de säger i chokladreklamen • d a v i d r • för att du gjorde undervisningen av grundläggande kemi till ett rent näje. när ska vi göra raketbjörnar igen? jag fixar sones. efter recept • l i l i a n • för att du gjorde undervisningen till en dans på rosor. hoppas att du har det lika bra där borta på andra sidan villorna • j o h a n n a • • • • för all hjälp med att hitta lilians gomstållen samt stöd när man behöver prata av sig efter en tung dag på labbet. uppsala är en mysig stad, glöm inte det.

tack alla administratörer som hjälper till med allt måste-arbete • l i l i a • även om ni försöker outsourca det mesta på gemensam lönn. och ibland tar er lite väl stora friheter som i att ordna bokningssystem (gurra. visst är du fortfarande arg?) vad är det för fel på penna och post-it? • • j o h a n n a • • • • jojo. biffen. johanna sekreterare. master of coins. kärt barn har många namn. för att du lyser upp en grå institution med rosa mössa, filt och stövlar. tack för all hemnetinspo. spinning-fighting-spirit (föredrar jag framstått med mordblick där framför på cykelbanan). skönt att ha någon att diskutera gäddhäng å fettvolanger med. någon som förstår. den första som fått mig att tvivla på mitt motto att alltid vara längst. utan våra sms-trådar med erika vore den här avhandlingen på tock för långt. så tack • i n g e r • för att du all-
tid hade tid när man kom och knackade på. och hade access till kreditkortet hoppas att du har det bra även utan oss linaj för din sköna inställning till livet å ditt luriga leende. för att du kom utifrån och berättade att det är viktigt att ha semester och så. så att vi lite fick känslan av att vi har ett riktigt jobb eva för att man antingen fick ett ansipte på den där mailutsändaren. tack för allt fix med mp. allt finns på mp. det har jojo lärt oss. eller hur erika? gunnar s för alla snack. inte minst under min första sommar här då alla andra verkade ha semester. för all gladje du sprider som julafon varje dag! bosse för all sponsring till fotbollen och all hjälp med program hit och dit. även om du holl dig på säkerhetsavstånd från annat för kakboden. eller ja. egentligen gillade jag den inte alls men att ha en egen vending machine som tar rosa post-it är ren lyx.

tack till alla som bidrog till att jag valde att följa den här vägen. tack valter & nils-gustav för osynlighetsskrift och klurigheter. till alla som gjorde lundatiden till en lek p101 & c:o för tegelbyggnadshång och våra försök till att ta dö på både KRUT och overalltraditionen. gjorde vårt karaoketåligt verkliga succé? ett särskilt tack till maria & robban för att jag fick bo på kontoret trots att det var ämnat åt alla demoprylar från superföretaget. tack för alla skratt. för tights tv-tablåer. för nintendomaraton. för bakmaskinsuppvaknandet. tack maurizio för att du var en fantasisk handledare som fick mig att vilja doktorera. för all jobbtips. jag kommer om några år! uppsala t tack för att du är en så fin stad. för alla små jurister som bidrog till den mysiga fritidsgårdskänslan på kantorsgatan. tack fotbollen som tvingade mig att strukturera arbetet för att kunna cykla från den ena änden av stan till den andra och bara få vara. få lite skall av max och tro att försvarsarbete plötsligt är det enda viktiga. och att skratta. för alla kompisar där. särskilt de som även fanns utanför plan mia madde cornelia tilda hanna lina l tack helena för att du alltid säger att det kommer att gå bra. och sedan erkänner att du inte har en aning för att man kan prata med dig om allt. för att du bryr dig om. så glad för att vi träffades den där julen i den där albyn jennie för att du alltid är så vettig och genombäst. min stötpelare och vapendragare. att bli ihopblandad med dig under hela gymnasiets tiden var ett rent nöje. jag ska se till att vi kommer till göteborg så att det kan hända igen brunberg & peterson mottot "det ordnar sig" har jag haft med mig under hela skrivandet. mycket lugnande mammadagis denna brokiga skara som fann varann ändå. så skönt att ta en paus från allt vetenskapligt dravel å låsa ikapp lite logiska mamilunderingar i chatten. snart står jag åter till tett förfogande david g. för att du vet allt. å kan allt. å hjälper till med allt. för att du berättar att man kan jobba i morgon och att man inte ska skicka lamptips mitt i natten före jobbdag ladad. för att du alltid är så peppande och förstående. och för all involvering i min jobbjakt jonsson den bästa gudmodern och vännen man kan få. tack för att du fyller min garderob när jag inte har tid att
tänka på sånt. för att du får mig att skratta. med dig. fast kanske mest åt dig. för att du föreslår singelbidrag när möbler blir för dyra. nu blir det utflykter igen. tack ♪ tidy tidy tidy up. I’d like to sit down on my potty. no monsters living in our home. five little ducks. wind the bobbin up ♪ för att ni fick mig att slappna av och andas ♪ justin bieber. jag hoppas innerligt att min kärlek till din love yourself är av det övergående slaget.

sist men inte minst vill jag tacka min familj. tack för att ni inte är som andra. ♪ mamma ♪ ett gott skratt förlänger livet så det har vi dig att tacka för “eller hur. pappa?” för att du då och då ringer för att fråga vad det är du ska säga att jag jobbar med igen. det var verklig lätta när jag höll på med bagerijåst. ska tänka på det nu inför nästa jobb. tack för all envishet. att du visat att allt går. knutar är till för att lönas. och kryddburkar kan man klämma sig i ♪ pappa ♪ för att du kan allt. och hjälper till med allt. fixar-frasse. med en egen MacGyver i familjen är det klart att man lär sig lösa diverse tekniska problem på labbet. grädsvipsskubbren är utökad! ♪ kim ♪ för alla inspirerande samtal om andra saker än forskning. som filmer. serier. politik (ZzzZzzZzz). för att du ger ruth det vi inte kan. intellekt och skägg. ♪ tony ♪ annika ♪ & ♪ adrian ♪ för den familjeförebild ni är. vi må ha skaffat unge först men vi kommer aldrig ikapp. hus. robotgräsklippare. och utspädd juice ♪ jan ♪ & ♪ gunilla ♪ för allt ni gör för oss. för att ni tycker att vi är viktigast i världen ♪ mormor ♪ för alla telefonsamtal om allt och inget ♪ för din vassa tunga och rappa humor. alla korsordsstunder som gjort mig till den klurare jag är idag. morfar skulle bara veta. tack för lingon i mjölen och aromat på bordet. tack ♪ farfar ♪ för allt du och farmor har gett mig. alla kortspelsstunder. våflor. kolaglassar från källaren. för att du alltid visar enorm glädje när man hör av sig även om du har hundra andra barnbarn och barnbarnsbarn. för att farmor änglavaktade mig när jag körde av vägen ♪ mogg & jinger ♪ å minifamiljen i minihuset. för all trivsel ni bidrar med. för alla uppmuntrande ord. och skrattfester.

• christoffer & ruth • för att ni finns. snarkar när jag åker hemifrån. så glada att ni inte kan stå något när jag kommer hem. tack • gossip girl & a blog to watch • för att ni tar hand om • chrisse • som i sin tur tar hand om allt. nästan. tack för att du introducerade detfinnsingarenakläerdervimåsteköpanya-begreppet. trevligt ändå. tack för att du säger att jag inte ska curla mina studenter. för dina pappa-skills och rut(h)iner. för köttbullar och falukorv. för pyttipanna och kebab. vilken hemmaman! • ruth • ■ tack för timingen ♪ som gjorde att artiklar blev publicerade medan jag hängde i lekhagar och käkade majsbrokar. utan dig hade denna avhandling inte blivit vad den blev. nu blir jag skojig igen. lovar!

xoxo
References


doi:10.1110/ps.034173.107.corresponding


89


Acta Universitatis Upsaliensis

Digital Comprehensive Summaries of Uppsala Dissertations from the Faculty of Science and Technology 1378

Editor: The Dean of the Faculty of Science and Technology

A doctoral dissertation from the Faculty of Science and Technology, Uppsala University, is usually a summary of a number of papers. A few copies of the complete dissertation are kept at major Swedish research libraries, while the summary alone is distributed internationally through the series Digital Comprehensive Summaries of Uppsala Dissertations from the Faculty of Science and Technology. (Prior to January, 2005, the series was published under the title “Comprehensive Summaries of Uppsala Dissertations from the Faculty of Science and Technology”.)

Distribution: publications.uu.se
urn:nbn:se:uu:diva-286557