Transaminase Biocatalysis: Applications and Fundamental Studies

Federica Ruggieri

Doctoral thesis in Biotechnology

KTH Royal Institute of Technology
School of Engineering Sciences in Chemistry, Biotechnology and Health
Stockholm 2019
©Federica Ruggieri

KTH Royal Institute of Technology
School of Engineering Sciences in Chemistry, Biotechnology and Health
Department of Industrial Biotechnology
AlbaNova University Center
SE-106 91 Stockholm, Sweden

TRITA TRITA-CBH-FOU-2019:47
Printed by Universitetsservice US-AB 2019

Akademisk avhandling som med tillstånd av KTH i Stockholm framlägges
till offentlig granskning för avläggande av teknisk doktorsexamen torsdagen
den 24 oktober kl. 10:00 i Kollegiesalen, KTH, Brinellvägen 8, Stockholm.
Dedicated to Antonio and Carla

Do not go where the path may lead, go instead where there is no path and leave a trail

- Ralph Waldo Emerson -
Abstract

Biocatalysis is the branch of science at the intersection between chemistry and biology and specifically dedicated to the application of natural evolvable catalysts, i.e. enzymes, in human-designed chemical processes. Among the array of promising biocatalysts, transaminases (EC 2.6.1.x) are possibly one of the enzyme classes with the largest unrealized potential. Fast inactivation, poor acceptance towards unnatural substrates and limited tolerance to cosolvents are some of the main factors hampering their implementation in chemical synthesis. In the present thesis work advances in both transaminase application and molecular understanding are presented. Indeed, these two topics are deeply interconnected, as a better molecular understanding is expected to ease the generation of novel enzyme variants suitable for new desired applications.

From the application perspective, the design of an effective one-pot transaminase-based racemization system offers new possibilities for the design of fully biocatalytic dynamic kinetic resolutions of valuable chiral amines. Similarly, the successful structure-guided redesign of the small substrate binding pocket of the Chromobacterium violaceum (S)-selective transaminase (Cv-TA) granted access to a new enzyme variant active on semi-preparative scale towards the unnatural substrate 1,2-diphenylethylamine.

From the molecular understanding perspective, the combination of crystallographic and computational techniques led to the formulation of a dimer dissociation model valid for Cv-TA and possibly for other enzymes belonging to the same fold type. This model, which aided the improvement of the Cv-TA stability by structure-based engineering, will hopefully enable similar results in other structurally related enzymes.
Sammanfattning

Biokatalys är den del av naturvetenskapen mellan kemi och biologi som är specifikt inriktad på tillämpningar av naturligt utvecklade biokatalysatorer, dvs enzymr, i av människan skapade kemiska processer. Av alla lovande biokatalysatorer som existerar, är transaminaser (EC 2.6.1.x) kanske den enzymkategorin som har störst outnyttjad potential. Snabb inaktivering, låg acceptans mot icke-naturliga substrat och en begränsad tolerans mot organiska lösningsmedel är några av de huvudsakliga faktorerna som begränsar användningen i kemiska synteser. I denna avhandling presenteras både framsteg inom transaminas-tillämpningar och molekylär förståelse. Dessa är i grunden djupt sammanhängande, då en bättre molekylär förståelse kan förväntas underlätta skapandet av tidigare okända enzymer som kan användas i nya önskade tillämpningar.

Från ett applikationsperspektiv, skapar designen av ett effektivt enkärls transaminasbaserat racemiserings-system nya möjligheter att designa biokatalytiska kinetiska resolveringar av värdefulla kirala aminer. På ett likartat sätt skapar den nya strukturbaserade designen av en liten substratbindande ficka i det (S)-selektiva transaminaset från Chromobacterium violaceum (Cv-TA) en ny enzymvariant som är aktiv i semi-preparativ skala med det icke-naturliga substratet 1,2-difenyletylamin.

La biocatalisi è quel ramo delle discipline scientifiche situato all’intersezione tra chimica e biologia e dedicato all’applicazione di enzimi in processi chimici messi a punto dall’uomo. La classe enzimatica delle transaminasi (EC 2.6.1.x) è forse tra quelle con il più grande potenziale irrealizzato in questo settore. La rapida inattivazione, scarsa attività verso substrati non naturali e limitata stabilità in presenza di cosolventi ne limitano infatti l’uso in applicazioni di sintesi organica. Il lavoro presentato in questa tesi è dedicato all’avanzamento delle possibilità applicative delle transaminasi e al raggiungimento di una loro migliore comprensione a livello molecolare. Questi due aspetti sono sostanzialmente interconnessi, dal momento che una più profonda comprensione teorica può contribuire al successo di strategie di ingegnerizzazione dirette all’ottenimento di varianti enzimatiche adatte per particolari applicazioni.

Nell’ambito applicativo, l’idea evizione di un sistema basato sull’uso di transaminasi per la racemizzazione in one-pot di ammine chirali offre nuove possibilità per la risoluzione cinetica dinamica di questa classe di molecole. Allo stesso tempo, l’ingegnerizzazione della tasca minore del sito attivo della transaminasi (S)-selettiva isolata da Chromobacterium violaceum (Cv-TA) ne ha ampliato l’utilizzo per la risoluzione su scala semi-preparativa del substrato non-naturale 1,2-difeniletilamine.

Nell’ambito dell’avanzamento teorico, la combinazione di tecniche cristallografiche e computazionali ha portato alla formulazione di un modello in grado di descrivere il processo di dissociazione dell’organizzazione dimerica valido per Cv-TA e, verosimilmente, per altri enzimi appartenenti alla medesima classe strutturale. Questo modello ha guidato con successo l’ingegnerizzazione di Cv-TA per migliorarne la stabilità ed è possibile che possa contribuire al raggiungimento di risultati simili in altri enzimi strutturalmente correlati.
Public defence of dissertation

This thesis will be defended on October 24th 2019 at 10:00 am in Kollegiesalen, Brinellvägen 8, Stockholm, Sweden.

Respondent:
Federica Ruggieri
School of Engineering Sciences in Chemistry, Biotechnology and Health
Department of Industrial Biotechnology
KTH Royal Institute of Technology, Stockholm, Sweden

Faculty opponent:
Associate Professor Mariarita Bertoldi
Department of Neurosciences, Biomedicine and Movement Sciences
Università di Verona, Verona, Italy

Evaluation committee:
Professor Rajni Hatti-Kaul
Department of Chemistry
Lund University, Lund, Sweden

Dr. Mats Sandgren
Department of Molecular Sciences
Swedish University of Agricultural Sciences, Uppsala, Sweden

Professor Mikael Widersten
Department of Chemistry
Uppsala University, Uppsala, Sweden

Chairperson:
Professor Antonius Van Maris
School of Engineering Sciences in Chemistry, Biotechnology and Health
Department of Industrial Biotechnology
KTH Royal Institute of Technology, Stockholm, Sweden

Respondent’s main supervisor:
Professor Per Berglund
School of Engineering Sciences in Chemistry, Biotechnology and Health
Department of Industrial Biotechnology
KTH Royal Institute of Technology, Stockholm, Sweden
List of publications and manuscripts

This thesis is based on the following publications or manuscripts, which are referred to in the text by their roman numerals:

**PAPER I**
**Ruggieri F**, van Langen LM, Logan DT, Walse B, Berglund P.
Transaminase-Catalyzed Racemization with Potential for Dynamic Kinetic Resolutions.

**PAPER II**
Land H, **Ruggieri F**, Szekrenyi A, Fessner, W-D, Berglund, P.
Engineering the Active Site of an (S)-Selective Amine Transaminase for Acceptance of Doubly Bulky Primary Amines.
*Manuscript.*

**Paper III**
**Ruggieri F**, Campillo-Brocal JC, Chen S, Humble MS, Walse B, Logan DT, Berglund P.
Insight into the Dimer Dissociation Process of the *Chromobacterium violaceum* (S)-Selective Amine Transaminase.

**Paper IV**
**Ruggieri F**, Ljungqvist E, van Langen LM, Logan DT, Walse B, Berglund P.
Stability Determinants in a sub-group of fold type I PLP-dependent enzymes.
*Manuscript.*
Contributions to appended papers

PAPER I
Contribution to the preliminary substrate acceptance screening, re-design of the cascade, design of the Ao-ATA expression plasmid, execution of experiments, synthesis of references, development of analytical methods, optimization and use of enzyme expression and purification protocols, analysis of data, preparation of the manuscript.

PAPER II
Contribution to F88A/L59A expression and purification, biochemical characterization ($T_m$ measurements, pH screen, DMSO screen), development of analytics, KR reactions, data analysis, revision of the manuscript.

PAPER III
Contribution to Cv-ATA (wild-type) expression and purification, crystallization, data collection, data processing, structure solving and refinement, development of reduction protocols, collection of stability and UV-Vis data, MD simulations, data analysis, preparation of the manuscript.

Paper IV
Contribution to development of the reduction method and of the expression and purification methods. Contribution to protein expression, purification, DSF characterization and structure-based search and analysis. Preparation of the manuscript.
List of abbreviations

TA  Transaminase
API  Active pharmaceutical ingredient
PLP  Pyridoxal-5’-phosphate
PMP  Pyridoxamine-5’-phosphate
KR   Kinetic resolution
AS   Asymmetric synthesis
DKR  Dynamic kinetic resolution
PGBC Phosphate group binding cup
PDB ID Protein data bank identification
HEPES 2-[4-(2-hydroxyethyl)piperazin-1-yl]ethanesulfonic acid
IMAC  Ion metal affinity chromatography
IPTG Isopropyl β-D-1-thiogalactopyranoside
ee   Enantiomeric excess
$t_{\text{rac}}$ Racemization time
DSF  Differential scanning fluorimetry
$T_{m}$ Melting temperature
MD   Molecular dynamics
Table of contents

1 Introduction 1

1.1 Evolution at work 1
1.2 Enzymes: the catalytic proteins of the cellular metabolism 2
1.3 The biomimicry of biocatalysis 4
1.4 Enzyme engineering 6
1.5 Transaminases 7

1.5.1 Transaminases: classification, mechanism and potential for biocatalytic applications 7
1.5.2 Transaminases in biocatalytic applications 11
1.5.3 Structural aspects of transaminases and prior-art in transaminase engineering 14

2 Present investigation 18

2.1 Transaminase-catalyzed racemization of primary amines (PAPER I) 20

2.1.1 Selection of biocatalysts 21
2.1.2 Production of biocatalysts 23
2.1.3 Proof of concept of the racemization of 1-methyl-3-phenylpropylamine 25
2.1.4 Determination of racemization time 26
2.1.5 Conclusions from PAPER I 28

2.2 Expanding the biocatalytic toolbox for the synthesis of 1,2-diphenyl chiral amines (PAPER II) 30

2.2.1 Identification of candidate sites for mutagenesis 31
2.2.2 Fluorescence-based screening method for substrate walking 32
2.2.3 Biochemical characterization of the L59A/F88A mutant 33
2.2.4 Kinetic resolution of 1,2-diphenyldiphenylethylamine from lab to semi-preparative scale

2.2.5 Conclusions from PAPER II

2.3 Understanding the dimer dissociation process of the Chromobacterium violaceum (S)-TA: the catalytic lysine as conformational switch (PAPER III)

2.3.1 PLP-dependent structural reorganizations in Cv-TA

2.3.2 The structural rearrangement of the catalytic lysine K288

2.3.3 MD simulations investigating the K288 rearrangement in the Cv-TA monomer in the presence and in the absence of covalently bound PLP

2.3.4 MD simulations investigating the K288-interfacial loop* interplay

2.3.5 Minimization of the K288-Y322* clash by mutagenesis and evaluation of the Schiff base-dependent stabilization

2.3.6 Evaluation of the role on stability of different PLP substructures

2.3.7 Conclusions from PAPER III

2.4 Structural determinants affecting the multimeric stability of a sub-group of fold type I PLP-dependent enzymes (PAPER IV)

2.4.1 NaCNBH$_3$-mediated reduction of a set of (R)- and (S)-selective TAs

2.4.2 Structural analysis of the tested (S)-TAs

2.4.3 PDB-wide structure search for structural analogues to Cv-TA

2.4.4 Structural analysis of Lb (EC 5.1.1.21) and Pj (EC 2.6.1.1)

2.4.5 Structural analysis of the remaining available crystal structures
2.4.6 Assessment of amino acid variation among the Cv-TA Y322 structural analogues and the alignment outlier Ht

2.4.7 Conclusions from PAPER IV

3 Concluding remarks and outlook
4 Acknowledgements
5 References
1. Introduction

1.1 Evolution at work

The existence of life as we know it is deeply linked to the series of complex and finely balanced biochemical transformations that are overall referred to as “metabolism”.\[1,2\] The metabolism of all organisms, just like all other inheritable traits, has been slowly and steadily moulded in the long evolutionary process that has rewarded those organisms capable of harvesting and transforming more efficiently than other competitors the energy and the matter available in their environment.\[1,3–6\]

The same principles of “random variation and selection” first formulated by Charles Darwin to explain the phenotypical variation of animals dominate the evolution of metabolisms.\[7,8\] Random mutations in the DNA sequences coding for catalytic proteins called enzymes can alter their amino acid sequence and hence their 3D shape, thus affecting their function within the complex cellular metabolic network and ultimately changing the “system cell”.\[8\] During the evolutionary process and under selective pressure, these mutations are selected for or against based on the metabolic advantages or disadvantages that they bring to the host organism in their competition for resources with other life forms.\[3\]

This simple evolutionary algorithm has moulded life over billions of years and has led to the astounding variation that we see at all levels in the biological and biochemical world.
1.2 Enzymes: the catalytic proteins of the cellular metabolism

Most of the biochemical transformations of the metabolic network are accelerated to rates that are useful for biological processes thanks to proteins called enzymes.\(^9,10\) The unique 3D structures of enzymes, optimized through evolution, make enzymes excellent catalysts suitable to perform a specific type of chemical transformation on a limited number of substrate molecules,\(^3,4\) with or without the aid of small non-protein molecules called cofactors.\(^11\) Although many hypotheses have been formulated over time to explain the catalytic nature of enzymes, it is now accepted that the pre-organization of the enzyme active site is the main contributor to the transition state stabilization leading to the enzyme-dependent enhancement of reaction rates.\(^10,12\) Indeed, the type and location of the amino acid residues lining the enzyme active site not only determine the substrate acceptance profile of an enzyme, but also contribute to form those stabilizing interactions with the transition state intermediate that lower the energy barrier necessary to form or break specific chemical bonds during the course of a reaction. This results in faster reaction rates (kinetics) with no effect on the equilibrium of the chemical transformation, which is solely determined by the free energy difference between the substrates (reactants) and the products ($\Delta G$ in Figure 1).

Based on the type of reaction that they catalyze and on the nature of the substrates that they accept, enzymes can be classified according to the EC (Enzyme Commission) system. In this system, a decision tree approach is used to assign to every enzyme a four-digit EC number (expressed in the form $x.y.z.w$) that defines the type of catalyzed reaction (first digit) and the reaction species involved in the reaction (second to fourth digits).\(^6\)
Figure 1. Schematic representation of the reaction diagram (free energy G vs reaction progress) of a general spontaneous reaction in which the reactants A and B (represented as black 2D shapes) are condensed to form the product C through formation of the transition state intermediate [AB‡] (uncatalyzed reaction, in blue) or [E·AB‡] (enzyme-catalyzed reaction, in red). The corresponding activation energies are represented with $\Delta G^{[AB]}$ and $\Delta G^{[E·AB]}$, respectively. The free energy difference of the reaction is represented with $\Delta G$. The active site of the enzyme E is shown as a red 2D shape. The energies of the enzyme-substrates and enzyme-product complex are intentionally omitted.

As the 3D structures of enzymes, and hence their functions, are altered in the evolutionary process, diversity is generated in the reactions that are catalyzed, in their efficiencies and in the types of molecules that are accepted as substrates.\textsuperscript{[3,6,8,13–15]} Thus, as reaction- and/or substrate promiscuities are generated, the evolutionary process is fueled.\textsuperscript{[4,5,15,16]} The stochastic nature of evolution has led to an astounding variety of enzyme-catalyzed reactions (classified according to the EC system) and enzyme structures (grouped into structural folds). Enzymes belonging to the same EC class can have
different structures (due to functional convergence or structural divergence) just like enzymes with similar structures can catalyze different reactions and belong to different EC classes (due to functional divergence or structural convergence).\textsuperscript{[17]} Functional divergence and convergence, jointly with the widespread existence of promiscuity phenomena,\textsuperscript{[3,18]} constantly pose a challenge for our habit to classification.

1.3 The biomimicry of biocatalysis

Nature has always been a great source of inspiration for the development of new technologies in many different fields, a phenomenon known as “biomimicry”. Within the realm of chemistry, the field of biocatalysis can be defined as the use of enzymes, in isolation or in combination, for the synthesis of particular compounds of interest for human applications. For thousands of years humankind has unknowingly used biocatalysis in fermentative processes by exploiting the complex metabolisms of small organisms invisible to the naked eye.\textsuperscript{[19,20]} Following the discovery of enzymes and the characterization of metabolic pathways, chemists began exploring the possibility of using enzymes to access valuable chemicals.\textsuperscript{[6,19,21–23]}

The renewability, cost-efficient production and biodegradability of enzymes, jointly with milder and safer operation requirements, are some of the significant advantages of biocatalytic methods over traditional synthetic processes.\textsuperscript{[21,24]} especially in consideration of the goals of chemistry with respect to sustainability.\textsuperscript{[19,21,24–27]}

Possibly, the most attractive features of enzymatic catalysis are high reaction- and substrate selectivity, high enantioselectivity, the possibility to combine enzymes from different sources to design enzymatic cascades\textsuperscript{[28]} and natural evolvability. Indeed, both the naturally
evolved selectivity of enzymes and their promiscuity can be exploited in synthetic processes to access natural and man-made compounds in one or multiple reaction steps.\textsuperscript{[15,29]} When multiple steps are performed in one-pot using multiple selective biocatalysts assembled in cascades, the isolation of intermediates becomes unnecessary, with a positive impact on process time and yields.\textsuperscript{[27]} Nonetheless, the major advantage of biocatalysts over traditional chemo-catalysts is possibly their evolvability. Nowadays, well-established recombinant technologies can be used to modify the DNA sequence coding for an enzyme of interest and produce changes in its amino acid sequence and hence in its properties.\textsuperscript{[19,30–32]}

Nonetheless, in spite of the many advantages of biocatalysis over traditional chemical synthesis, enzyme-based processes are not yet the rule in industry. The areas of the chemical industry that have seen the largest implementation of enzyme-based processes are the food industry and the pharmaceutical industry,\textsuperscript{[19,27,33,34]} while the bulk chemistry sector has seen so far sparse (yet significant) examples.\textsuperscript{[20,23,24,27,35–37]} For a fact, biocatalytic approaches advance slowly against established chemical methods. Problems such as poor enzyme stability, poor acceptance towards non-natural substrates of interest, difficult process intensification and/or poor tolerance towards (co-)solvents, possible need for cofactor regeneration and incompatibility issues in the design of multi-enzymatic cascades are only some of the problems that need to be overcome on a case-to-case basis in the development of biocatalytic processes.\textsuperscript{[20,23,36,38,39]} These issues can be mitigated and even solved completely by either adapting the process to the needs of the enzyme(s) (process engineering), or by adapting the enzyme to the needs of the process (enzyme engineering).\textsuperscript{[27]}
1.4 Enzyme engineering

Enzyme engineering is the process of modifying existing enzymes to create novel biocatalysts suitable for a particular application. Indeed, while the enzymes found in nature have evolved to be excellent catalysts in the metabolism of their host organisms, they often perform poorly in the intensive artificial systems devised for human-determined purposes. Poor operational stability and low tolerance to (co-)solvents, sensitivity to substrate- and product inhibition and limited acceptance towards non-natural substrates are only some of the main limitations that should be overcome to enable the implementation of enzymes as catalysts for organic synthesis on industrial scale.\textsuperscript{[19,20,27]}

In what has been defined as the “second wave of biocatalysis” (as opposed to the first, where natural enzymes, \textit{i.e.} wild-type, were used in intensified processes to access their natural products) the 3D structures of enzymes with unrealized industrial potential were inspected to identify amino acid positions that, once mutated, could lead to enzyme variants better suited for a particular application.\textsuperscript{[19]}

This approach, still used today under the name “structure-based enzyme re-design” or “rational design”, is usually characterized by a low success rate\textsuperscript{[40]} for two main reasons: i) any inspected crystal structure shows only one (or a few) of the possible conformations that the enzyme can adopt, providing very little information about its possible structural motions\textsuperscript{[41]} and ii) our understanding of the effects of a mutation on the structure, dynamics, stability and catalysis of an enzyme are very limited, let alone predictable.\textsuperscript{[20,42–44]}

During the “third wave of biocatalysis” started in the 1990s, new enzyme variants have been successfully generated by what has been called “directed evolution”.\textsuperscript{[19,31,32]} In this approach, the evolutionary
algorithm of random mutation and selection is exploited iteratively in the laboratory under artificially established selective pressure, leading to new quickly evolved enzyme variants capable of accepting novel substrates and even of performing completely new chemical reactions.\[45\] Enzyme variants evolved through this method have shown that amino acid positions difficult to correlate with enzymatic catalysis merely based on structure can, once altered, affect the catalytic behavior and other properties of an enzyme in ways that we cannot fully understand yet.\[40\] In spite of the success of this method, the need to implement complex and usually expensive high-throughput screening (HTS) platforms to characterize the vast collections of enzyme variants generated by directed evolution hampers the ready exploitation of this strategy for enzyme engineering.\[19,36,46\]

In conclusion, enzyme engineering has proven to us that we still cannot predict the behavior of an enzyme merely based on its 3D structure nor easily understand the exact effect of amino acid mutations on its properties. As Nobel laureate Frances Arnold very elegantly put it, enzyme engineering has taught us that “we can read, write, edit… but we cannot compose”.\[45\]

1.5 Transaminases

1.5.1 Transaminases: classification, mechanism and potential for biocatalytic applications

Transaminases (TAs, EC 2.6.1.x) are enzymes with great potential for the synthesis on industrial scale of chiral amines,\[47–49\] i.e. organic molecules containing an amino group (-NH$_2$) as one of the four distinct substituents arranged around a carbon atom, the stereocenter. The four substituents decorating a stereocenter can be
arranged around it in any of two configurations (i.e. 3D arrangements), resulting in different chemical species called enantiomers and defined (R)- or (S)- according to a set of rules (Figure 2).\textsuperscript{50} Chirality, i.e. the property of an object of existing in different enantiomeric forms, is common among bioactive organic molecules and the vast majority of enzymes found in nature are capable to discriminate between different enantiomers of a substrate molecule, a property called enantioselectivity.\textsuperscript{51,52} The enantioselective fashion in which enzymes perform their catalytic function is of outstanding interest for large-scale synthetic purposes, especially for the synthesis of complex chiral active pharmaceutical ingredients (APIs).\textsuperscript{33,53}

Transaminases are of interest in biocatalysis because of the high enantioselectivity with which they transfer an amino group from a donor molecule (usually an amine or an amino acid) to an acceptor molecule (usually a ketone or an aldehyde) following a reaction mechanism based on the participation of the covalently bound cofactor PLP as a transient amino donor/acceptor. (Scheme 1).\textsuperscript{11,54,55}

**Figure 2.** General schematic representation of two amine enantiomers, where the group priorities are assigned in the order NH$_2$>2>3>4 according to the Cahn-Ingold-Prelog set of rules. The clockwise or counter-clockwise arrangement of the three highest priority groups around the carbon stereocenter determine the (R)- or (S)-configuration of the stereocenter.
The fact that both (R)- and (S)-selective transaminases with high enantioselectivities and broad substrate scopes have been found to exist in nature\cite{47,56} has added to the interest toward this enzyme class. Although a further classification into \( \alpha \)-, \( \omega \)- and amine-transaminases has been proposed in the past based on the substrate acceptance of these enzymes\cite{57,58}, their substrate promiscuity\cite{48,59–62} has limited the utility of this classification, which, for this reason, will be avoided in this PhD thesis in favor of the more general classification based on enantioselectivity.

Regardless of classification, all known transaminases are believed to follow a common reaction mechanism.\cite{54,55} In their catalytically competent resting state (holo-state), the cofactor PLP is bound via a covalent Schiff base linkage to the side chain nitrogen of a catalytic lysine to form the so-called “internal aldimine”. In the first reaction step, upon attack of the aldimine carbon by the amino group of the substrate (the amino donor), the PLP-lysine bond is broken and a new Schiff base is formed between the PLP and the substrate molecule, the so-called “external aldimine”. After a series of bond rearrangements across the PLP ring and a proton abstraction step involving the terminal nitrogen of the catalytic lysine, the amino group of the substrate is transferred on the cofactor to form the non-covalent cofactor species pyridoxamine-5’-phosphate (PMP) with concomitant release of the carbonyl product (ketone or aldehyde). The second half of the transamination reaction, starting upon binding of the carbonyl co-substrate (the amino acceptor) and proceeding with the transfer of the amino group of PMP onto the carbonyl moiety of the ketone, ends with the release of the amino product and the regeneration of the PLP-lysine internal aldimine (Scheme 1).
By virtue of the high enantioselectivity of the transaminase-catalyzed reaction, the last 25 years have witnessed an increasing interest in transaminase research, especially in relation to the synthesis of chiral primary amines for the pharmaceutical sector. In 2015 it was estimated that about 40% of the marketed active pharmaceutical ingredients (APIs) contained at least one amino stereocenter in their chemical structure, a percentage that is not expected to decrease in the future under the effect of the ever tightening regulatory requirements for the approval and commercialization of chiral drugs. While these policies protect the patients against the possible different pharmacological profiles of
API enantiomers, they also challenge the drug manufacturing processes, faced with the limited availability of tools for the stereoselective synthesis of chiral compounds on a large scale.\[^{57,65}\]

In this context, the high enantioselectivity of the transaminase-catalyzed reaction has drawn the attention of those chemists interested in controlling the stereoselectivity of key chemical steps in the synthesis of chiral APIs containing amino stereocenters.

### 1.5.2 Transaminases in biocatalytic applications

The constantly expanding toolbox of highly enantioselective (\(R\))- and (\(S\))-selective transaminases is being explored and exploited in a variety of different approaches to synthesize chiral amines, resolve amine racemates and invert the chirality of amine stereocenters.\[^{49,66,67}\]

The most straightforward reported transaminase application is the resolution of amine racemates, \textit{i.e.} 1:1 mixtures of enantiomers, \textit{via} “kinetic resolution” (KR) (Scheme 2, panel a).\[^{47,49,57,67}\] In this approach, an enantioselective transaminase is used to selectively convert the undesired amine enantiomer in the racemate into its corresponding carbonyl-containing species, which can then be easily separated from the unconverted (slow-reacting) desired amine enantiomer. Although straightforward, the low amine yields (maximum 50 %), the consumption of an amino acceptor in the process and the generation of large amounts of co-product are significant drawbacks in the large-scale exploitation of kinetic resolutions.

A second approach, called “asymmetric synthesis” (AS),\[^{47,49,57,67}\] takes advantage of the reversibility of the transamination reaction to access enantiopure amines in one step through stereoselective re-
Scheme 2. Overview of the applications of transaminases in biocatalysis. Panel a: kinetic resolution (KR); panel b: asymmetric synthesis (AS); panel c: deracemization; panel d: racemization for dynamic kinetic resolutions (DKRs).
ductive amination of aldehydes or pro-chiral ketones (Scheme 2, panel b). In 2007, five of the largest pharmaceutical manufacturers worldwide listed this reaction as one of the most desirable reactions to be prioritized in chemical research.\cite{25} Despite the many advantages of transaminase-base AS over traditional chemo-catalysts for this type of application, the often unfavorable thermodynamic equilibrium\cite{68–70} has not led to a quick and massive implementation of transaminases in large-scale chemical processes, nor, for that respect, of stereoselective reductive amination reactions. While equilibrium-displacement strategies such as large excess of cheap amino donors\cite{69} and carbonyl co-product removal by physical,\cite{71–75} chemical\cite{69,76–82} or enzymatic\cite{68,69,83–85} methods have proved successful on the lab-scale, their potential for implementation on industrial scale is quite limited.

A third, less explored and more complex application of transaminases is the de-racemization of amines in a one-pot sequence of transaminase-catalyzed KR and AS (Scheme 2, panel c). In this type of application, two enantiocomplementary transaminases, or one transaminase combined with a different enzyme,\cite{86} are used in either one\cite{87} or two steps\cite{88,89} to i) deaminate the undesired amine enantiomer in a KR and ii) stereoselectively re-introduce the amino group on the carbonyl-containing product in an AS. As the substrate overlap between (R)- and (S)-TAs is limited, the identification of a suitable couple of catalysts with compatible operation conditions is a limiting factor, as is the necessity to find a suitable equilibrium displacement method to drive the AS step of the deracemization.

The last transaminase-based application reported to date is the one-step racemization of amines, in which two enantiocomplementary transaminases are coupled, similarly to the deracemization case, to
invert the chirality of an amine stereocenter (Scheme 2, panel d). The only example of this type reported in the literature at the beginning of this PhD project was based on the use of expensive and unstable achiral amino donor/acceptor couples that could not prevent the reaction to halt at the ketone stage.\cite{90} Indeed, a racemization of this type can only proceed with internal recycling of amino donors and acceptors if the racemization equilibrium constant between substrates and co-substrates, both of which must be accepted by the two transaminases, is 1. The design of successful transaminase-based racemization systems is important for the development of amine dynamic kinetic resolutions (DKRs) in which racemic mixtures are resolved by coupling in one-pot of a stereoselective enzyme-catalyzed KR and a racemization system that constantly re-equilibrates the racemate from the slow-reacting enantiomer.

Due to the harsh, enzyme-incompatible reaction conditions needed to racemize amines with traditional chemical methods,\cite{91} DKRs are difficult to implement. A fully enzyme-compatible transaminase-catalyzed racemization system would therefore ease the assembly of DKR reactions to resolve amine racemates with theoretical yields of 100%.

1.5.3 Structural aspects of transaminases and prior-art in transaminase engineering

All known transaminases catalyze the stereoselective transfer of an amino group from a donor to an acceptor in a reaction mechanism that involves the participation of the covalent cofactor pyridoxal-5’-phosphate (PLP) as building block delivery system.\cite{11,54} The cofactor PLP, derivative of the vitamin B6, is one of the most versatile cofactors in nature, reportedly involved in 238 distinct enzymatic activities.\cite{92} It is thought that PLP-dependent enzymes govern and
steer the reactivity of their cofactor by providing active site environments that are optimized to stabilize in different ways the transition state intermediates formed in the first common step of their catalytic cycles. Evolution fulfilled this requirement by selecting seven distinct fold-classes within the PLP-dependent enzymes superfamily (I-VII), each optimized to achieve both a specific orientation of the substrates with respect to PLP and a particular stabilization of the resulting transition state intermediate.

Therefore, it is not surprising that (R)- and (S)-TAs meet the requirement of substrate coordination on opposite faces of the PLP ring, necessary to achieve opposite enantioselectivities, thanks to different foldings of the protein backbone, i.e. fold type I for (S)-TAs and fold type IV for (R)-TAs (Figure 3). As a consequence of their distinct folds, their active site architectures are also radically different. According to a commonly accepted model, the relative orientation of two active site pockets (one big and one small) around the cofactor PLP governs the coordination of the small and large groups of the substrate with respect to the internal aldimine bond, and thus the enantioselectivity of the reaction (Figure 4). This model explains the opposite enantioselectivities of (R)- and (S)-TAs with an opposite orientation of the big and small active site pockets around the reactive Schiff base. Although both TA groups are known to be active as homo-dimers or homo-tetramers, only in the (S)-TAs the two identical active sites found in the dimeric assembly are located at the monomer-monomer interface, with residues belonging to both monomers lining the active site cavity and participating in PLP binding. An identical active site architecture is found in the homo-tetrameric (S)-transaminases (fold type I), where the tetramer is formed in the side-to-side association of two homo-dimers. On the other hand, the active site of (R)-transaminases (fold type IV) is most-
Figure 3. Homo-dimeric assembly of holo-(S)-transaminases (on the left, fold type I) and holo-(R)-transaminases (on the right, fold type IV). In each dimer, the two monomers are shown in green and pink. The cofactor PLP is shown as black sticks in all active sites. (PDB IDs: 4A6T, Chromobacterium violaceum (S)-TA; 4CMD, Nectria haematococca (R)-TA)

ly buried within each of the monomers, with only one loop belonging to the neighboring monomer contributing to shape the access tunnel. Recent studies have demonstrated that the stability of fold type I (S)-TAs, both dimeric and tetrameric, correlates with the association of the cofactor moiety. Structural studies have shown that in fold type I (S)-TAs the phosphate group of PLP is coordinated at the monomer-monomer interface, where residues belonging to both monomers line the so-called phosphate group binding cup (PGBC). Based on this, the extensive hydrogen bonding network established in the PGBC between the protein and the phosphate group of PLP has been proposed to contribute to the stability of the dimeric assembly. On the contrary, not much is known about the role of PLP on the stability of (R)-TAs.
Figure 4. Virtual representation of the organization of the small (S) and large (L) active site pockets in (S)- and (R)-selective transaminases. The volumes of the two substrate binding pockets are delimited by black discontinuous lines. The atoms of the lysine-PLP covalent species are shown as yellow sticks. The atoms of the model substrate 1-phenylethylamine, considered in its (R)- and (S)-configuration, are shown as pink sticks, with the amine functionality coloured in blue.

Enzyme engineering strategies have been successfully used in a number of cases to alter the substrate scope, enantioselectivity, reaction promiscuity, tolerance to solvents and sensitivity to inhibition of both (R)- and (S)-selective transaminases. Stability improvement remains challenging despite recent progress in this direction.\textsuperscript{[66,102–106]} Due to the great diversity of TAs, engineering strategies and outcomes are still to be evaluated on a case-to-case basis.
2. Present investigation

Despite their attractiveness from the environmental friendliness and chemical selectivity perspective, enzyme-catalyzed syntheses of commercially important chemicals have been poorly implemented on industrial scale. In what could be defined as the “biocatalysis paradox”, the same enzyme properties that have been progressively fine-tuned by evolution and that make enzymes selective and regulatable catalysts in the natural world also limit their performance in artificially designed systems. Indeed, processes operated over a long period of time in conditions that are significantly different from those of the cellular environment can challenge the stability and the catalytic profile of enzymes.

The biocatalysis paradox can be overcome by devoting efforts in three directions (Figure 5): i) broadening of the biocatalytic toolbox through the discovery of new enzymes with unique properties and the development of enzyme variants with altered properties; ii) deeper understanding of the biochemical determinants of enzyme structure, stability, dynamics and catalytic behavior and iii) design of new processes and cascades (both chemo-enzymatic and fully enzymatic) that realize the full potential of enzymes and open new possibilities for their use.

In the present thesis work all three directions have been explored with regards to transaminases (EC 2.6.1.x), particularly (S)-selective.

In PAPER I the usefulness of transaminases as catalysts was expanded with the design of a system relying solely on enantiocomplementary transaminases for the racemization of primary amines.

In PAPER II the Chromobacterium violaceum (S)-selective TA was engineered by rational design to accept an unnatural bulky-bulky
amine, thereby retaining its enantioselectivity. The variant was successfully employed in both homogeneous and heterogeneous mixtures for the kinetic resolution of the target amine at varying concentrations of the cosolvent DMSO.

In PAPER III the (S)-selective transaminase from *Chromobacterium violaceum* was investigated by means of structural and computational methods to elucidate its PLP-dependent dimer dissociation process. The resulting model aided the improvement of the enzyme stability via rational design.
In PAPER IV the investigation of the role of PLP in the stability of (R)- and (S)-TAs revealed that only fold type I transaminases show a Schiff base-dependent stability. A broader analysis across other structurally-related fold type I enzymes suggested that dimer dissociation processes similar to that formulated in PAPER III might apply to other enzymes belonging to this fold type.

2.1 Transaminase-catalyzed racemization of primary amines (PAPER I)

Dynamic kinetic resolution (DKR) approaches consist in the coupling in one pot of an enantioselective enzyme and a racemization system. By enabling the enantioconvergence of readily accessible racemates with yields of the desired enantiomer close to 100\%, DKRs are approaches with high potential for the pharmaceutical manufacture, where single enantiomers are still mostly accessed with a maximum 50 % yield via simple racemate resolution. Hence, the higher yields and lesser waste of DKRs make them a desirable alternative to traditional resolutions.\[107\]

The development of DKRs for the enantioconvergence of amines is highly desirable since chiral amines are one of the classes of compounds with the highest impact in the pharmaceutical manufacture, where they are extensively used as chiral auxiliaries and synthons.\[57,108\] Despite the extensive research that followed the first successful amine DKR report by Reetz and Schimossek,\[109\] the advances in amine DKR design and application have halted due to the scarcity of enzyme compatible amine racemization methods.\[110\]

In fact, the traditional amine racemization methods usually require expensive transition-metal chemo-catalysts\[110–112\] (mostly Raney-Ni,\[111–113\] Shvo-type,\[114–118\] palladium-\[113,118–121\] or ruthenium-based\[122\]), expensive non-selective reducing agents\[123\] and/or harsh
conditions\textsuperscript{[108,124]} that can hardly be combined in one pot with an enantioselective enzyme.

In PAPER I the proof-of-concept of a fully transaminase-based racemization system is reported. In this racemization method, two wild-type enantiocomplementary transaminases are coupled via an internal co-product removal system to invert the configuration of the substrate stereocenter (Scheme 3). The internal co-product removal system, based on the use of minimal amounts of the broadly accepted and easily accessible co-substrate couple pyruvate/D,L-alanine, also proved effective in displacing the equilibrium of the asymmetric synthesis (AS) step of the racemization.

2.1.1 Selection of biocatalysts

The first step in the cascade design process included the choice of the target amine compound and the selection of two enantiocomplementary transaminases capable of accepting it as a substrate. The chiral model compound 1-methyl-3-phenylpropylamine (1) (Scheme 3) was selected as racemization target due to its slow reported chemo-catalyzed racemization,\textsuperscript{[108,120]} which makes it a good gold standard for the comparison of chemo- and enzyme-catalyzed racemizations.

Suitable enantiocomplementary transaminases accepting 1 as a substrate were selected among a panel of 10 enzymes, 3 (S)-selective and 7 (R)-selective (Table 1), including the (S)-selective TA from Chromobacterium violaceum [Cv-(S)-TA], both wild-type (WT) \textsuperscript{[62]} and W60C,\textsuperscript{[125]} and the enzymes available in the ACS-ATA-KIT (Enzymicals AG). All biocatalysts were tested for activity using the variant of the acetophenone assay\textsuperscript{[126]} shown in Scheme 4, where the co-substrate 1-phenylethylamine, either with (R)- or (S)-configuration, is consumed to aminate the ketone 4-phenyl-2-butan-
Scheme 3. Transaminase-catalyzed racemizations of 1S (panel a) and 1R (panel b) (15 mM). Both reactions were performed using equal amounts of purified Cv-(S)-TA and Ao-(R)-TA (50 mU) as catalysts in the presence of amino acceptor pyruvate (4) (4.5 mM) and amino donor D-alanine 3R (panel a) or L-alanine 3S (panel b) (10 mM) in HEPES pH 7.5 (50 mM), glycerol (20% v/v) and PLP (0.1 mM) at room temperature in the dark for four days.

Based on the results of the screening (Figure 6), the TA couple consisting of *Chromobacterium violaceum* (S)-TA and *Aspergillus oryzae* (R)-TA [Ao-(R)-TA] was selected for use in the proof-of-concept of the racemization of 1.
Table 1. Panel of enzymes screened for the selection of a suitable couple of enantiocomplementary TAs.

<table>
<thead>
<tr>
<th>Enzyme (code in the ACS-ATA_KIT)</th>
<th>Organism of origin</th>
<th>Enzyme (code in the ACS-ATA_KIT)</th>
<th>Organism of origin</th>
</tr>
</thead>
<tbody>
<tr>
<td>Af-(R)-TA (ATA01)</td>
<td>Aspergillus fumigatus</td>
<td>Pc-(R)-TA (ATA06)</td>
<td>Penicillium chrysogenum</td>
</tr>
<tr>
<td>Fg-(R)-TA (ATA02)</td>
<td>Fusarium graminearum</td>
<td>Mv-(R)-TA (ATA07)</td>
<td>Mycobacterium vanbaalenii</td>
</tr>
<tr>
<td>Nf-(R)-TA (ATA03)</td>
<td>Neosartorya fischeri</td>
<td>Rp-(S)-TA (ATA08)</td>
<td>Ruegeria pomeroyi</td>
</tr>
<tr>
<td>Ao-(R)-TA (ATA04)</td>
<td>Aspergillus oryzae</td>
<td>Cv-(S)-TA</td>
<td>Chromobacterium violaceum</td>
</tr>
<tr>
<td>At-(R)-TA (ATA05)</td>
<td>Aspergillus terreus</td>
<td>Cv-(S)-TA-W60C</td>
<td>Chromobacterium violaceum – W60C mutant</td>
</tr>
</tbody>
</table>

2.1.2 Production of biocatalysts

In order to have complete control over both the amount of active catalysts and the concentrations of pyruvate and D- and L-alanine used in the reactions, the enzymes were used in their purified form after having experimentally determined the specific activity of the purified enzyme formulations in the variant of the acetophenone assay (Scheme 4). Both Cv-(S)-TA and Ao-(R)-TA were hence produced recombinantly in E. coli BL21(DE3) and purified in one step on IMAC (immobilized metal affinity chromatography). The N-terminal His-tagged Cv-(S)-TA was expressed and purified starting from the E. coli BL21(DE3) strain previously described in the literature\[127\] following established procedures.\[101,127\] The pET-29b(+) expression plasmid containing the C-terminal His-tagged Ao-(R)-TA sequence, retrieved by BLAST search and codon-optimized for E. coli, was inserted into the E. coli BL21(DE3) strain. The Ao-(R)-TA expression was optimized by maximizing the amount of soluble pro-
Scheme 4. Variant of the acetophenone assay used to screen the panel of \((R)\)- and \((S)\)-TAs for acceptance towards the target substrate 1.

Figure 6. Results of the initial screening for the identification of a suitable couple of enantiocomplementary TAs accepting 2 as a substrate. Initial activity is shown increase in $\text{Abs}_{245}$ vs $t\,[\text{min}]$. Data referring to selected catalysts are shown as continuous lines. The $Cv\-(S)\-TA$ W60C, although showing good initial catalytic rates in the screening, showed poor stability in preliminary reactions and was not investigated further.
Figure 7. SDS-PAGE results for the optimization of the expression of soluble Ao-(R)-ATA. The two SDS-PAGE gels show the amount of soluble Ao-(R)-ATA (MW = 37 kDa) normalized on a fixed amount of cell suspension. The combination of variables tested for each sample are indicated on the top of the gel pictures.

The soluble protein expressed at 20 °C in TB medium for 24 hours in the presence of 0.3 mM IPTG was purified in the holo-form in one step of IMAC after cell disruption by sonication. Removal of imidazole triggered protein precipitation, and for this reason the purified enzyme was stored directly after elution from the Ni-NTA column. The specific activity of both purified enzymes was measured in the standard acetophenone assay\cite{126} and expressed as mU/mg.

2.1.3 Proof of concept of the racemization of 1-methyl-3-phenylpropylamine

The proof-of-concept racemization of 1-methyl-3-phenylpropylamine (1) (Scheme 3) was performed starting from either its (S)-
enantiomer 1S or from its (R)-enantiomer 1R (15 mM) in HEPES pH 7.5 (50 mM) in the presence of both Ao-(R)-TA and Cv-(S)-TA (50 mU) and minimal amounts of co-substrates pyruvate 4 (4.5 mM) and L- or D-alanine (3S and 3R, 10 mM). In order to prevent enzyme precipitation, glycerol (20% v/v) was also added to the reaction mixture. A constant cofactor supply was ensured by either adding excess PLP only at time zero (0.1 mM) or by also supplementing PLP daily (to final 1.7 mM) during the whole reaction time (4 days).

The qualitative and quantitative composition of the reaction mixture at day 4, analyzed by chiral normal phase HPLC and achiral GC, confirmed near-completion racemization in all cases (ee < 10%) (Table 2). Our results also show that daily addition of cofactor PLP is both unnecessary for the purpose of the racemization and undesirable with respect to both accumulation of intermediate ketone and final amine ee. We suspect that PLP could participate stoichiometrically in the reaction by removing amino groups from the system in the form of leaked PMP, thereby affecting the reaction equilibrium and promoting the accumulation of ketone 4. This effect would be larger and more significant in the case of TAs that, like Cv-(S)-TA, show a low affinity for the cofactor molecule.¹⁹⁹

These results demonstrate that the transaminase-based racemization of 1 shown in Scheme 3, performed in water, at room temperature and at near-neutrality pH, is fully comparable, in terms of performance on 1, to the already known chemo-catalyzed racemization methods.

2.1.4 Determination of racemization time

For the assembly of successful DKRs a highly enantioselective resolving enzyme needs to be coupled to a much faster racemization system, typically progressing at least ten times faster than the resolu-
Table 2. Qualitative and quantitative results for the transaminase-catalyzed one-step racemization of 1S and 1R under tested conditions.

<table>
<thead>
<tr>
<th>Starting amine</th>
<th>Catalytic system[a]</th>
<th>Total amount of PLP [μmol (mol%)]</th>
<th>ee of amine after 4 days[b]</th>
<th>Absolute concentration of ketone after 4 days [% (μmol)]</th>
</tr>
</thead>
<tbody>
<tr>
<td>1S</td>
<td>No enzymes</td>
<td>1.7 (11)</td>
<td>&gt;99% (S)</td>
<td>3 (0.5)</td>
</tr>
<tr>
<td>1S</td>
<td>Ao-(R)TA + Cv-(S)-TA</td>
<td>1.7 (11)</td>
<td>9% (S)</td>
<td>36 (5.4)</td>
</tr>
<tr>
<td>1R</td>
<td>No enzymes</td>
<td>1.7 (11)</td>
<td>&gt;99% (R)</td>
<td>4 (0.6)</td>
</tr>
<tr>
<td>1R</td>
<td>Ao-(R)TA + Cv-(S)-TA</td>
<td>1.7 (11)</td>
<td>5% (R)</td>
<td>42 (6.3)</td>
</tr>
<tr>
<td>1R</td>
<td>No enzymes</td>
<td>0.10 (0.7)</td>
<td>&gt;99% (R)</td>
<td>4 (0.6)</td>
</tr>
<tr>
<td>1R</td>
<td>Ao-(R)TA + Cv-(S)-TA</td>
<td>0.10 (0.7)</td>
<td>3.8% (R)</td>
<td>28 (4.2)</td>
</tr>
</tbody>
</table>

[a] Substrate amine 1S or 1R (15 mM), amino donor of opposite chirality 3R or 3S (10 mM) and amino acceptor 4 (4.5 mM) in HEPES pH 7.5 (50 mM) supplemented with PLP (0.1 mM). PLP was supplemented daily (up to 1.7 mM) unless otherwise specified. Equal amounts of the two catalysts (50 mU) were used, when present. Reaction volume 1 ml.
[b] values obtained from chiral normal phase HPLC (qualitative analysis)
[c] values obtained from achiral GC (quantitative analysis)
[d] PLP was not supplemented daily after time zero

The racemization time ($t_{\text{rac}}$) of our transaminase-based system (i.e. the time that would be needed to achieve complete racemization when the initial racemization rate is held constant) was therefore evaluated for the two cases, starting from either 1S or 1R, when cofactor PLP (0.1 mM) was only added at time zero.

Samples collected over time and analyzed qualitatively and quantitatively on chiral normal phase HPLC showed that the racemization rate decreases as the racemization equilibrium is approached. In DKRs, where the resolution step keeps the system far from the racemization equilibrium, the racemization is expected to reach completion when $t = t_{\text{rac}}$, a value that was calculated for both reactions as the abscissa of the intersection point between $y = 1$ (enantiomeric ratio at the racemic equilibrium) and the linear fitting of the variation of enantiomeric ratio over the first 1.5 h of reaction.
(Figure 8). We think that the different racemization times (20 h for the reaction from 1S to racemate and 14 h for the reaction from 1R to racemate) calculated for the two thermodynamically identical reactions could be a result of the different structural stabilities and sensitivities to inhibition of the two enantioselective enzymes. Indeed, depending on the starting (excess) enantiomer of 1 and on the different inactivation rates of the two TAs, the amination and deamination of 2 can proceed at different rates, which would also be reflected in the value of $t_{rac}$ if the racemization progress is evaluated based on the variation of enantiomeric ratio of 1 over time.

2.1.5 Conclusions from PAPER I

In conclusion, in PAPER I a novel biocatalytic method for the racemization of primary amines is presented. The method, that relies on the use of wild-type non-proprietary enantiocomplementary transaminases, proved effective in water and at room temperature without the need for complex or wasteful equilibrium displacement systems. The presented enzymatic cascade, although demonstrated on lab scale and only for one model amine substrate, proved comparable in performance to other known racemization methods that are incompatible with enzyme catalysis due to their harsh reaction conditions. The transaminase-catalyzed racemization presented in PAPER I poses itself as a valuable alternative for the development of DKR reactions targeting valuable chiral amines.
Figure 8. Determination of the racemization kinetics starting from 1S (panel a) or 1R (panel b). The experimentally determined values of enantiomeric ratio (i.e. absolute relative concentrations) are plotted against time in a scatter plot. The values of $t_{rac}$ (in red) was calculated as the abscissa of the intersection point between $y = 1$ and the linear functions $f(x)$, shown as dotted lines, interpolating $(0;0)$ and the first two points of the scatter plots ($t = 0.5$ h and $t = 1$ h).
2.2 Expanding the biocatalytic toolbox for the synthesis of 1,2-diphenyl chiral amines (PAPER II)

The different size of the two substrate binding pockets in the active site of transaminases is commonly regarded as the main factor determining their typically high enantioselectivity. While the large pocket (L) can accommodate bulky aromatic and aliphatic substituents, the small pocket (S) can generally accommodate groups not much bigger than a methyl group, thus limiting the scope of possible substrates accessible via transamination to compounds that bear significantly differently sized substituents around the carbon stereocenter. The enlargement of the small binding pocket S via enzyme engineering could, in perspective, expand the utility of transaminases for the synthesis of chiral amines bearing two large groups decorating the carbon stereocenter. Diphenylethylamines are one of the groups of compounds of pharmaceutical importance containing this feature in their chemical structure. Their derivatives have been approved in a number of countries for use as anaesthetics and analgesics and for this reason the development of routes for their synthesis in enantiopure form is of commercial importance.

In PAPER II the size of the small binding pocket of the Chromobacterium violaceum TA (Cv-TA) was successfully modified using rational design approaches to accommodate the 2-phenyl group of the compound 1,2-diphenylethylamine. Contrary to the wild-type enzyme, the engineered variant L59A/F88A proved capable of accepting the target bulky-bulky chiral amine and showed an improved tolerance towards DMSO. This variant, which retained (S)-selectivity, was used successfully in lab- and semi-preparative scale KR reactions both in the presence and in the absence of DMSO,
both in homogeneous and heterogeneous reaction mixtures, at substrate loadings as high as 20 g/L.

2.2.1 Identification of candidate sites for mutagenesis

As in the case of most rational design approaches where structural information is available for the target enzyme, the deposited structure of the holo-Cv-TA (PDB ID: 4A6T and 4AH3) was exploited to identify candidate sites for mutagenesis. A preliminary structural analysis revealed that only two hydrophobic residues with different side chain sizes, i.e. L59 and F88, contribute to determine the volume this pocket (Figure 9). Both amino acid positions were therefore mutated by site-directed mutagenesis to generate the single mutants L59V, L59A, L59G and F88A.

Notably, in the case of the F88 position only one mutation type was attempted (F88A) as this mutation was expected to result in the largest possible increase in the size of the S pocket without introducing the backbone rotational freedom associated with glycine.
mutations. Indeed, due to the large aromatic side chain of F88 in the wild-type, the introduction of a glycine in this position was expected to have a much larger effect on the flexibility of the backbone (and possibly protein conformation) compared to the L59 position, for which a glycine mutation was instead attempted.

2.2.2 Fluorescence-based screening method for substrate walking

The effect of the mutations on the size of the S pocket was evaluated by taking advantage of a previously described sensitive fluorescence-based transamination assay \([132]\) (Scheme 5) in which 1-(6-methoxynaphth-2-yl)alkylamines are used as amino donors in transamination reactions that consume equivalent amounts of inexpensive amino acceptor pyruvate. In the case of transaminases that, like Cv-TA, can accommodate the napthyl group of these substrates in the L pocket, the chemical variation around the carbon stereocenter can be used to probe the acceptance of groups of different sizes in the S pocket in a substrate walking approach. The structurally related acetonaphthone species produced during transamination can then be easily detected by fluorescence \((\lambda_{ex} = 330 \text{ nm}; \lambda_{em} = 460 \text{ nm})\), thus providing a powerful and sensitive tool for the detection of (trace) transaminase activity.

Compounds 5, 6 and 7 (Scheme 5) were therefore used to probe the effect of the introduced mutations on the size of the S pocket. The results (Figure 10) show that in all of the generated mutants the acceptance of compounds 6 and 7 was improved at the expenses of the acceptance of 5, possibly due to a less stable enzyme-substrate interaction upon broadening of the S pocket. Further combination of the F88A mutation with each of the L59 mutations led to enzyme variants showing trace activity towards both 5 and 6 jointly with drastically increased activity toward the bulky substrate 7, suggesting
Scheme 5. Substrate walking fluorescence-based screening method used for the enlargement of the small substrate binding pocket of Cv-TA (top). Chemical variation can be introduced by changing the substituent R. The three substrates used for the screening of the increased S pocket size are shown at the bottom, with the group accommodated in this pocket shown in magenta. The blue and magenta portion of the substrates corresponds to the substrate amine screened for (5 for 1-phenylethylamine; 6 for 1-phenylbutan-1-amine; 7 for 1,2-diphenylethylamine).

an agonistic effect of the mutations at these two amino acid positions. The variant with the highest initial rate towards 7, i.e. the double mutant L59A/F88A, was characterized biochemically and explored as suitable biocatalyst for the synthesis of 1,2-diphenylethylamine.

2.2.3 Biochemical characterization of the L59A/F88A mutant

The biochemical characterization of the Cv-TA L59A/F88A included the determination of i) its melting temperature ($T_m$) by differential scanning fluorimetry (DSF); ii) its pH optimum and operational pH range, iii) its DMSO tolerance and iv) its enantioselectivity.
Figure 10. Results of the substrate screening assay for the broadening of the Cv-TA substrate binding pocket using compounds 5, 6 and 7 as probes (see Scheme 5). For each mutant set, i.e. single or double mutants, data are shown as percentage of the highest specific activity (in U/mg) measured for that set. For comparison, the activity of the wild-type enzyme (WT) towards the three tested substrates is also shown.

The \( T_m \) value for L59A/F88A, corresponding to the temperature value at which the folded and unfolded forms of the protein are at thermodynamic equilibrium, was experimentally determined to be 75 °C, 4 °C higher than the one reported for the wild-type enzyme. \[^{101}\] The standard acetophenone assay \[^{126}\] performed in the presence of equimolar amounts (5 mM) of 1-(S)-phenylethylamine and pyruvate at different pH values allowed to establish that the L59A/F88A variant has a pH optimum at pH 7.5, about one pH point lower than for the wild-type enzyme, with residual activities above 50% at both pH 7.0 and 8.0 (Figure 11, panel a).
Figure 11. Results of the biochemical characterization of Cv-TA L59A/F88A: pH profile (panel a), DMSO screen (panel b) and enantioselectivity (panel c) shown for the KR of 1,2-diphenylethylamine performed with Cv-TA wild-type (top) or L59A/F88A (bottom) as catalyst (Reaction conditions: 1 mM amine, 2 mM sodium pyruvate, 50 mM HEPES pH 7.5, room temperature, 0.5 mg/ml purified enzyme, 1.5 h. Reaction volume: 1ml. Retention times: 8.0 min for (R)-enantiomer, 11.0 min (S)-enantiomer).

Tolerance to DMSO was evaluated by performing the same type of assay on reaction mixtures containing varying amounts of cosolvent and HEPES buffer (50 mM, pH 7.5). Differently from the wild-type enzyme, which shows a steady decrease in activity from 0% to 50% v/v DMSO, the L59A/F88A variant has a near-symmetrical activity optimum around 20% v/v DMSO, with similar specific activity retained at 0% and 50% v/v cosolvent (Figure 11, panel b). This trait is particularly beneficial for the intended application of the L59A/F88A variant because of the poor solubility in water of both 1,2-diphenylethylamine (5 mM) and 2-phenyl-acetophenone (<0.3 mM).
Finally, the (S)-enantioselectivity of this enzyme variant was confirmed in preliminary KR reactions of 1,2-diphenylethylamine performed in water at low substrate concentrations (1 mM) (Figure 11, panel c).

2.2.4 Kinetic resolution of 1,2-diphenylethylamine from lab to semi-preparative scale

The biochemical characterization of the (S)-selective Cv-TA L59A/F88A mutant showed that this variant is at least as stable at room temperature as the Cv-TA wild-type enzyme and that it can be employed for synthetic purposes at neutral to slightly basic pH (pH 7.0-8.0) in the presence of 0-50% v/v DMSO. Its potential in the KR of 1,2-diphenylethylamine was therefore evaluated in different conditions (Figure 12). The results show that in the presence of low substrate concentrations (5 mM, 3 ml) and 0% v/v DMSO the KR reached completion (> 99% ee (R)) in 4 h using 0.2 mg/ml of purified enzyme. By increasing the concentration of substrate and enzyme (20 mM and 0.7 mg/ml, respectively) in the presence of 20% v/v DMSO, the KR was completed in 2 h with excellent final enantiomeric excess (> 99% ee (R)). The KR could also be performed successfully in 4 h (> 99% ee (R)) in a non-homogeneous reaction mixture containing undissolved substrate (100 mM) in buffer (500 mM HEPES) without addition of cosolvent in the presence of 1.8 mg/ml of purified enzyme.

2.2.5 Conclusions from PAPER II

The work presented in PAPER II is a clear example of the advantages of rational design approaches for the generation of enzyme variants with altered properties. Available structural information for Cv-TA guided the selection of two candidate sites for mutagenesis to broa-
Figure 12. Cv-TA L59A/F88A catalyzed KR of 1,2-diphenylethylamine (7, Scheme 5, page 33). Panel a: 5 mM 7, 7.5 mM sodium pyruvate, 0.5 M HEPES pH 7.5, 0.2 mg/ml Cv-TA wild-type, reaction volume 3 ml. Panel b: 5 mM 7, 7.5 mM sodium pyruvate, 0.5 M HEPES pH 7.5, 0.2 mg/ml Cv-TA L59A/F88A, reaction volume 3 ml. Panel c: 20 mM 7, 30 mM sodium pyruvate, 20 % v/v DMSO, 0.5 M HEPES pH 7.5, 0.7 mg/ml Cv-TA L59A/F88A, reaction volume 3 ml. Panel d: 100 mM 7, 200 mM sodium pyruvate, 0.5 M HEPES pH 7.5, 1.8 mg/ml Cv-TA L59A/F88A, reaction volume 3 ml.

den its small substrate binding pocket S. The generation of seven enzyme variants (four single mutants and three double mutants), easily screened using a sensitive fluorescence-based assay, sufficed to isolate one variant (Cv-TA L59A/F88A) capable of accepting the unnatural di-phenyl substrate 1,2-diphenylethylamine. Although changes in other biochemical properties of this mutant, such as pH
optimum and tolerance to DMSO, cannot be currently explained, we found that the (S)-selective Cv-TA L59A/F88A enzyme is a robust catalyst for the preparation of 1-(R)-1,2-diphenylethylamine via kinetic resolution of racemates in the presence or absence of cosolvent at substrate loadings as high as 100 mM (20 g/L).

2.3 Understanding the dimer dissociation process of the *Chromobacterium violaceum* (S)-TA: the catalytic lysine as conformational switch (PAPER III)

One of the main factors hampering the implementation of transaminase-based processes in industry is their usually low stability, both during storage and extended operation in unnatural conditions. Despite that enzyme engineering approaches have been successfully applied in a number of cases to alter the substrate specificity, enantioselectivity and sensitivity to inhibition of these enzymes,[66] the improvement of their stability has generally proven more challenging. Although different engineering approaches have been described recently for this purpose,[103–105,133–135] our still limited understanding of the molecular processes affecting transaminase stability hinders the possibility to design more stable enzyme variants without relying on directed evolution.

In PAPER III the dimer dissociation process of the well-characterized (S)-selective transaminase from *Chromobacterium violaceum* was investigated. Previously reported biochemical,[99] biophysical[99] and structural[101,136] data, jointly with new structural, experimental and mutagenesis data, support an inactivation model centered on the role of the catalytic lysine as a conformational switch sensitive to the presence of the Schiff base linkage with the cofactor PLP.
2.3.1 PLP-dependent structural reorganizations in Cv-TA

The starting point in the formulation of the model was the in-depth analysis and comparison of the deposited crystal structures of the apo- and holo-Cv-TA (PDB IDs: apo: 4A6R, 4BA4; holo: 4A6T and 4AH3). The authors that first solved and compared the structures of the two forms of the Cv-TA observed that significant structural rearrangements occurred upon PLP binding (in the apo-to-holo transition) or release (in the holo-to-apo transition). These structural rearrangements, considered in the holo-to-apo transition, include: i) the disordering of the N-terminal domain (residues 5-35); ii) the reorientation of the catalytic lysine K288 from its conformation stretched inside the active site (herein called “forward conformation”) to a conformation pointing toward the back of the active site (herein called “backward conformation”); iii) the recoiling of a loop located at the dimer interface (the “interfacial loop”, residues 311-326) and iv) the outward rigid-body swing of what will be here defined “outer loop” (residues 81-96).

While the conformational changes of these four regions were initially described with reference to the PLP-binding site, we observed that these four regions also organize into an interlock at the dimer interface (Figure 13), with the interfacial and outer loop of one monomer being “clamped” between the K288 and the N-terminal domain of the neighboring monomer. In each active holo-Cv-TA homodimer (PDB IDs: 4A6T and 4AH3) two such interlocks are formed, each contributing to shape the so-called phosphate group binding cup (PGBC) of one of the active sites. Indeed, the PGBC of each active site is lined by residues belonging to the interfacial loop of the monomer that does not contribute its K288 to
Figure 13. Reciprocal spatial interlock organization of the four regions undergoing structural rearrangements upon PLP binding or release. Regions belonging to different dimer subunits are shown in different shades of blue and regions belonging to the monomer that does not contribute the K288 residue to the active site are also marked with *. The K288-PLP moiety is represented as sticks. (PBD ID: 4A6T)

Indeed, upon PLP release and transition to the apo-enzyme (PDB IDs: 4A6R and 4BA4), each of the interfacial loops participating in the PGBC interactions is observed to recoil, thereby causing the opening of two deep solvent-accessible clefts from opposite sides of the dimer, one for each active site (Figure 15). The opening of these clefts, jointly with the disordering in solution of the N-terminal do-
main, would favor the dimer dissociation of the apo-Cv-TA, in agreement with previously published observations. The close spatial organization of the regions undergoing structural rearrangements and the correlation between enzyme conformation, PLP binding and dimer association prompted us to further investigate the role of PLP in the structure dynamics of Cv-TA.

2.3.2 The structural rearrangement of the catalytic lysine K288

In the holo-Cv-TA, two of the four regions observed to rearrange upon PLP loss, i.e. the catalytic lysine and the interfacial loop, esta-
blish strong direct contacts with the cofactor moiety. Specifically, the catalytic lysine K288 is involved in the formation of a Schiff base linkage with the 4’ aldehyde group of PLP and the interfacial loop participates in the extensive H-bonding network at the level of the PGBC. Previous studies\cite{99} focusing on the effect of phosphate coordination on Cv-TA stability showed that, while both PLP and inorganic phosphate contribute to promote the stability of the C\textsubscript{v}-TA dimer, the effect of PLP on dimer stability is more prominent than the effect of inorganic phosphate.\cite{99} This led us to hypothesize a previously unsuspected contribution of the Schiff base to the C\textsubscript{v}-TA stability, which prompted us to investigate further the conformational dynamics of the catalytic lysine K288 involved in the formation of the Schiff-base.

Indeed, K288 is known to rearrange from its forward to its backward conformation in the holo-to-apo transition, \emph{i.e.} upon PLP loss.
Figure 16. Forward (in blue) and backward (in red) conformation of the catalytic lysine K288. The lysine and PLP, when present, are shown as sticks. The transition from forward to backward conformation results in a 8.1 Å displacement at the terminal side chain nitrogen and in a 2.7 Å displacement at the Cα. (PBD ID: holo: 4A6T; apo: 4A6R\textsuperscript{[101]})

Importantly, due to its direct participation in the catalytic cycle (Scheme 1, page 10), the terminal amino group of the K288 side chain must necessarily be in close proximity to the 4’ reactive group of PLP. While this proximity requirement is met in the holo-\textit{Cv}-TA, where the K288 is present in the forward conformation, the backward conformation characteristic of the apo-\textit{Cv}-TA holds the terminal K288 amino group about 8 Å away from the PLP ring (Figure 16).

In order to establish whether the K288 rearrangement occurs during the catalytic cycle as the Schiff base is broken, the crystal structure of the complex between \textit{Cv}-TA and PMP (PMP·\textit{Cv}-TA) was solved at
1.67 Å resolution. The complex, generated by briefly exposing holo-Cv-TA crystals to the substrate (S)-1-phenylethylamine during cryoprotection, showed that, at least in the crystal, the K288 rearrangement to the backward conformation does not occur upon transient Schiff base disruption. This result led us to conclude that this K288 rearrangement could be the outcome of a more complex dynamic process.

The analysis of the Ramachandran plots calculated from the available crystal structures revealed that the forward K288 conformation, contrary to the backward K288 conformation, is associated with backbone strain. Indeed, in the holo-Cv-TA, K288 and its preceding residue A287 fall consistently in unfavoured regions of the Ramachandran plot. The conformational rearrangement to the K288 backward conformation dissipates this backbone tension and brings the \( \phi \) and \( \psi \) torsion angles of K288 and A287 back to favored regions of the Ramachandran plot.

2.3.3 MD simulations investigating the K288 rearrangement in the Cv-TA monomer in the presence and in the absence of covalently bound PLP

To test whether this backbone tension would suffice to trigger the rearrangement of K288 to the backward conformation upon PLP loss (and hence Schiff base disruption), molecular dynamics (MD) simulations were performed on the isolated holo-Cv-TA monomer after having removed the atom coordinates corresponding to the PLP moiety. The results of the simulation showed that the rearrangement of K288 to the backward conformation is indeed hindered by the presence of the K288-PLP Schiff base and enabled by its absence (Figure 17 and Figure 18).
Figure 17. Results of the MD simulation performed of the holo-Cv-TA monomer complete of PLP. The top part of the figure shows the superposition of the MD snapshots (0 – 40 ns) in a progressive color palette from blue to red. Regions known for undergoing conformational changes are represented as tubes. The bottom part of the figure shows the comparison between the K288 conformation at the end of the simulation (in red) and its conformation in the apo-Cv-TA (in green, PDB ID: 4A6R101).
Figure 18. Results of the MD simulation performed of the holo-Cv-TA monomer depleted of PLP. The top part of the figure shows the superposition of the MD snapshots (0 – 40 ns) in a progressive color palette from blue to red. Regions known for undergoing conformational changes are represented as tubes. The bottom part of the figure shows the comparison between the K288 conformation at the end of the simulation (in red) and its conformation in the apo-Cv-TA (in green, PDB ID: 4A6R[101]).
2.3.4 MD simulations investigating the K288-interfacial loop* interplay

Further structural analysis revealed that the rearrangement of K288 to the backward conformation would cause a collision with the side chain of Y322* located on the interfacial loop* lining the PGBC (with * indicating elements belonging to the neighboring monomer) (Figure 19).

Additional MD simulations performed on the complete dimer depleted of the two N-terminal domains (Figure 20) showed that the clash generated by the rearranging K288 would suffice to trigger the recoiling of the interfacial loop* to the conformation observed in the apo-Cv-TA. This recoiling would in turn possibly cause the displacement and loss of order of the N-terminal domain in a cascade of conformational rearrangements initiated by the disruption of the K288-PLP Schiff base linkage. The role of the K288-Y322* clash in Cv-TA stability was therefore tested in mutagenesis studies targeting the side chain size of Y322*.

2.3.5 Minimization of the K288-Y322* clash by mutagenesis and evaluation of the Schiff base-dependent stabilization

The truncation of the bulky side chain of Y322 in the Y322A mutant, which would create more space for the K288 movement, correlates with an increased stability of both the apo- and the holo-form of the Cv-TA. The exact extent of the stabilization was evaluated as positive shift in the melting temperature ($T_m$) value measured by differential scanning fluorimetry (DSF) experiments ($\Delta T_m$(apo-Y322A)-(apo-WT) = 7.9 °C; $\Delta T_m$(holo-Y322A)-(holo-WT) = 7.0 °C) (Figure 21a).

The stabilizing role of the K288-PLP Schiff base, which would provide the constraint necessary to keep K288 in the forward confor-
Figure 19. Reciprocal lysine K288 and interfacial loop conformations in the holo-(in shades of blue) and apo- (in shades of red/pink) Cv-TA homodimers. Different subunits belonging to the same dimer are represented in different shades of the same type of color. Structural elements belonging to the subunit that is not contributing the K288 are also indicated with *. K288, Y322* and PLP, where present, are shown with a stick representation. (PBD ID: holo: 4A6T; apo: 4A6R)[101]

formation, was then evaluated by treatment with the mild and selective reducing agent NaCNBH3.

In agreement with our hypothesis, the holo-Cv-TA, both wild-type and Y322A, is stabilized upon reduction of the Schiff base to a non-hydrolyzable secondary amine (Figure 21b).

As expected, this stabilizing effect is more prominent for the Cv-TA
Figure 20. Results of the MD simulation performed of the whole holo-Cv-TA homodimer depleted of PLP and N-terminal domains, shown limited to the K288 and the Y322* of one of the chains. Time-resolved MD snapshots (0 – 40 ns) are represented in a progressive color palette from blue to red. The reference conformation in the apo-Cv-TA (PDB ID: 4A6R) is shown in red. The K288 and Y322* residues are shown as sticks.

wild-type, where less space is available for the K288 reorganization upon Schiff base hydrolysis. A further confirmation of the role of the Schiff base constraint on K288 and Cv-TA stability came from the K288A mutant, which is inherently unsuitable to form the Schiff base
Figure 21. DSF results for the untreated (panel a) and treated (panel b) Cv-TA wild-type, Y322A and K288A.

with PLP and whose A288 rearrangement, due to the much smaller side chain, is expected to affect less the conformation of the interfacial loop*. The stability data measured by DSF showed that i) the K288A mutation has a mild destabilizing effect with respect to the Cv-TA wild-type, and this effect is larger for the holo-form ($\Delta T_m$(apo-K288A)-(apo-WT) = -1.5 °C; $\Delta T_m$(holo-K288A)-(holo-WT) = -3.2 °C) (Figure 21a-b) and ii) the NaCNBH3-mediated Schiff base reduction does not affect significantly the stability of the K288A mutant in the holo- nor in the apo-form ($\Delta T_m$(apo-K288A)-(apo-reducedK288A) = -0.1 °C; $\Delta T_m$(holo-K288A)-(holo-reducedK288A) = -0.1 °C).

It was previously shown[100] that Cv-TA incubation in the presence of an amino donor leads to Cv-TA inactivation, possibly due to diffusion of the non-covalently bound PMP away from the active site. This observation is in agreement with our inactivation model, as the prolonged Schiff base disruption in the absence of a suitable amino acceptor would enable the rearrangement of K288 and the cascade of structural reorganizations leading to Cv-TA inactivation. Therefore, we also assessed the effect of the prolonged incubation in
the presence of L-alanine on the stability of Cv-TA, considered in its wild-type, Y322A and K288A variants.

As expected, DSF experiments confirmed that the incubation with L-alanine destabilizes the holo-forms of the Cv-TA variants competent for the formation of the Schiff base (i.e. wild type and Y322A) by an extent that is larger for the Cv-TA wild-type (Figure 21b), where the movement of the bulky K288 side chain is likely to trigger the series of conformational rearrangements possibly leading to dimer dissociation.

On the other hand, and possibly due to non-covalent stabilizing interactions, the same incubation with L-alanine mildly increased the measured $T_m (\Delta T_m < 0.3 \, ^\circ C)$ of the apo-forms of all of the Cv-TA variants (i.e. wild-type, Y322A and K288A) (Figure 21b). A similar explanation could be provided for the increased stability of the holo-Cv-TA K288A upon incubation with L-alanine (Figure 21). This mutant, unsuitable to form a Schiff base with PLP, could still bind and retain the cofactor non-covalently after the initial PLP saturation thanks to the PGBC interactions. Upon incubation with L-alanine, an external Schiff base could form spontaneously and be retained in the active site as a large ligand that cannot be further converted due to the lack of K288-mediated proton abstraction. The non-covalent nature of this stabilization was supported by the finding that apo- and holo-Cv-TA K288A are equally stabilized when incubated in equimolar amounts of PLP and L-alanine, while the apo-forms of the Cv-TA wild-type and Y322A are stabilized more than their corresponding holo-forms (Figure 22).

Our results, in agreement with previous studies,\cite{99} suggest a positive effect of non-covalent phosphate-containing species on Cv-TA stability. As previously described,\cite{101} the interfacial loop partly con-
Figure 22. Cv-TA stabilization, assessed by DSF, by protein incubation in the presence of equimolar amounts of PLP and l-alanine.

tributes to the H-bonding network spanning the PGBC, which could contribute to its rigidity and make its K288-mediated displacement less likely. Therefore, we decided to use MD to evaluate separately the effect of the PGBC interactions and of the Schiff base on the K288 and interfacial loop conformations.

2.3.6 Evaluation of the role on stability of different PLP sub-structures

The results of the simulation (Figure 23) suggest that the whole PLP structure might be necessary for the stabilization of the holo-Cv-TA homodimer. When the atoms of the phosphate group of PLP are removed, the K288, although bound via the Schiff base to the cofactor pyridine ring, was still seen to partly reorient to the backward conformation (Figure 23b) and when the PLP pyridine ring is removed, the remaining methyl phosphonate can diffuse out of the PGBC and out of the active site (Figure 23a). In the latter case, and in agreement with the previously described mild stabilizing role of inorganic phosphate,\(^{[99]}\) the loose coordination of the methyl phos-
Figure 23. Results of the MD simulations performed on the whole Cv-TA dimer depleted of N-terminal domains and of one of the two main cofactor substructures, either pyridine ring (panel a, yielding a methyl phosphonate) or 5’ phosphate (panel b, resulting in PLP-desphosphate bound via the Schiff base to K288). For both cases, the MD results relative to both active sites of the dimer AB are shown.
phonate, partly overlapping with the rearrangement trajectory of K288, is sufficient to slow down the K288 movement and to prevent, at least in the time covered by the simulation, the recoiling of the interfacial loop* away from the PGBC.

From these results we concluded that both cofactor substructures (i.e. the 4’ group participating in the Schiff base and the 5’ phosphate group) are important, under normal conditions, to prevent the K288 rearrangement and to promote the stability of Cv-TA.

Nonetheless, based on the crystal complex between Cv-TA and the inhibitory adduct formed in the reaction between PLP and the compound gabaculine (PDB ID: 4BA5),\textsuperscript{[136]} we suspect that, of the two, the K288 strain might be the predominant factor affecting stability. In the complex, produced by co-crystallization, Cv-TA is present in the conformation typical of the apo-enzyme, with the K288 in the backward conformation, destructured N-terminal domains and recoiled interfacial loops*, hence showing that the coordination of the big phosphate-containing PLP-gabaculine adduct was not sufficient to prevent the rearrangement of K288 nor the subsequent conformational changes. If this were true, the 4BA5 complex would be the true complex of the gabaculine adduct and the apo-Cv-TA.

\textbf{2.3.7 Conclusions from PAPER III}

In conclusion, the investigation presented in PAPER III shows how, in the (S)-selective Cv-TA, the geometrical properties of the active site determine the mode of inactivation of this PLP-dependent fold type I enzyme.

The Cv-TA presented inactivation model (Figure 24), with its PLP-
dependent and dynamic nature, suggests that, in vivo, the yet unknown metabolic role of this enzyme might be regulated through the intracellular level of available PLP. This inactivation route, which is in agreement with other published studies describing the inactivation of other enzymes belonging to the same fold class,\textsuperscript{137–139} might apply to other non-transaminase enzymes belonging to this group and offer a new perspective to design more stable enzymes for a variety of different biocatalytic applications. The specific active site architecture that this inactivation mechanism has evolved for makes us suppose that, despite the similar catalyzed reactions, (R)-transaminases belonging to fold group IV do not obey this same inactivation mechanism (see PAPER IV). Although to date we do not know the natural metabolic function(s) of (R)- and (S)-transaminases, a tighter PLP-dependent regulation for (S)-TAs would not be surprising. Indeed, all natural proteinogenic amino acids are (S)-amino species and it is plausible that any metabolic process affecting the intracellular pool of amino acids is finely regulated.

Figure 24. Schematic representation of the proposed Cv-TA inactivation model.
2.4 Structural determinants affecting the multimeric stability of a sub-group of fold type I PLP-dependent enzymes (PAPER IV)

The cofactor pyridoxal-5′-phosphate, or PLP, is one of the most versatile cofactors existing in nature, involved in more than 230 distinct catalytic functions.\textsuperscript{[92]} It is commonly accepted that the high reactivity of this cofactor is steered thanks to a complex interplay between the cofactor moiety and the protein structure\textsuperscript{[58,93,95,140]} and that this is possibly the evolutionary drive that has led to PLP-dependent enzymes evolving into 7 different, phylogenetically unrelated\textsuperscript{[141–144]} structural folds (fold types I-VII).\textsuperscript{[98,145]} Each fold includes many enzymatic activities\textsuperscript{[96,145]} that diverged gradually from an ancestor function by mutation of residues involved in the stabilization of the transition state intermediate(s).\textsuperscript{[144]} \((R)\)- and \((S)\)-selective TAs govern the stereochemistry of the PLP-dependent transamination by coordinating the cofactor PLP in protein scaffolds belonging to different fold types, \textit{i.e.} fold type I for \((S)\)-TAs and fold type IV for \((R)\)-TAs (Figure 25).

In PAPER III a dimer dissociation process was proposed for the \((S)\)-TA isolated from \textit{Chromobacterium violaceum} which suggests a significant role of the Lys-PLP Schiff base in preventing the conformational changes associated with protein inactivation and ultimately responsible for dimer dissociation. In PAPER IV the role of the Lys-PLP Schiff base was experimentally evaluated for \((R)\)-TAs using the NaCNBH3-mediated Schiff base reduction combined with DSF as “diagnostic” tool. After having confirmed that the Schiff base does not contribute to the stability of \((R)\)-TAs, a broader investigation based on structural analysis allowed to establish that the dimer dis-
Figure 25. Structural differences between (R)- (in green, fold type IV) and (S)-TAs (in blue, fold type I). In the two enzyme groups the fold of the monomer, the organization of the dimer and the position of the PLP (represented as sticks) with respect to the monomer-monomer interface differ (panel a). Panel b shows the (R)- and (S)-TA active site organization up close. Panel c shows the 2D structural topology of the two fold types, with the position of the catalytic lysine indicated with asterisks. (PDB IDs: Aspergillus fumigatus (R)-TA 4CHI, Chromobacterium violaceum (S)-TA 4A6T).
sociation process proposed for \( \text{Cv-TA} \) could be valid for other fold type I enzymes, even non-transaminase. It was observed that, while the conformation of the interfacial loop is highly conserved among the considered enzymes, its reorganization upon PLP release is very diverse. It was found that the dimer interface architecture differs among the enzymes classified within fold type I and that only a sub-group of these, herein included \( \text{Cv-TA} \), do have a flexible interfacial loop prone to rearrangement. We propose that for the enzymes belonging to this sub-group, regardless of their catalytic activity and possibly with a few exceptions, a dimer dissociation process similar to the one proposed for \( \text{Cv-TA} \) could apply.

2.4.1 \( \text{NaCNBH}_3 \)-mediated reduction of a set of \((R)\)- and \((S)\)-selective TAs

The \( \text{NaCNBH}_3 \)-mediated reduction protocol presented in PAPER III (see paragraph 2.3.5) was used to establish whether the Lys-PLP covalent Schiff base contributes to the stability of dimeric \((R)\)-TAs as well as other \((S)\)-TAs sharing the same structural fold of \( \text{Cv-TA} \). This method, combined with DSF measurements, can indeed provide important data regarding the contribution of the Schiff base to protein stability despite the lack of structural information for the apo and/or holo forms of a given transaminase of interest.

Similarly to what was done in PAPER III, a small set of \((R)\)- and \((S)\)-selective transaminases (herein included \( \text{Cv-TA} \) and its W60C variant \([125]\)) was subjected to Schiff base reduction and DSF analysis. The DSF results (Figure 26) show that, while all of the tested \((R)\)-TAs are mildly destabilized or otherwise not significantly affected by the chemical treatment, all tested \((S)\)-TAs are stabilized by the Schiff base reduction.

This suggests that the Schiff base-dependent stabilization is unique,
among transaminases, to fold type I TAs and it opens the possibility that dimer dissociation processes similar to the one proposed for Cv-TA might apply to other enzymes belonging to this group. The available structures of the experimentally tested (S)-TAs (Vf-TA and Sp-TA) were therefore inspected, especially with respect to those regions proposed to be dynamically linked through the series of conformational rearrangements leading to dimer dissociation, i.e. the catalytic lysine and the interfacial loop.

2.4.2 Structural analysis of the tested (S)-TAs

Both Vf-TA apo and holo crystal structures (PDB IDs: apo 5ZTX, holo 4E3R) were compared to the corresponding Cv-TA structures (Figure
The holo forms of the two enzymes show nearly identical interfacial loop conformations around the Cv-TA Y322 position (whose structural analogue is an alanine in Vf-TA) and a higher degree of structural divergence at the N-terminal portion of this loop. Upon PLP loss, the interfacial loop of both enzymes is also observed to reorganize, although differently (Figure 27).

In PAPER III it was shown that in Cv-TA the mutation of the bulky Y322 to an alanine improves the thermodynamic stability of the enzyme, presumably by allowing more space for the K288 movement and thereby preventing the steric clash triggering the conformational rearrangement of the interfacial loop. In Vf-TA, an alanine residue in the same position does not seem to prevent the reorganization of the interfacial loop away from the PGBC, which seems to affect the conformation of a neighboring loop (residues 161-176) forming the core of the crystallographic holo Vf-TA tetramer. Although the quaternary structure of the holo Vf-TA in solution has not been addressed experimentally, it is possible that this enzyme exists in a dimer/tetramer equilibrium and that PLP affects this equilibrium just like it has been shown to affect the monomer/dimer equilibrium of Cv-TA.\textsuperscript{[99]}

The inspection of the apo-\textit{Sp}-TA (PDB ID: 3HMU) revealed that this structure resembles more the holo structures known for Cv-TA and Vf-TA in terms of catalytic lysine (K290) strain and extended interfacial loop conformation. This led us to conclude that this structure should be regarded as the true holo \textit{Sp}-TA structure despite the lack of bound PLP. Indeed, a tighter crystal packing and the coordination of a sulfate ion in the PGBC could have contributed to disfavor any apo-to-holo conformational rearrangement in the crystallization process.
Figure 27. Structural analysis of the catalytic lysine and interfacial loop conformations of the apo and holo crystal structures of Cv-TA (in blue, PDB IDs: apo 4A6R, holo 4A6T), Vf-TA (in orange, PDB ID: apo 5ZTX, holo 4E3R) and Sp-TA (in purple, PDB ID: “holo” 3HMU). Interfacial loops are shown as tubes, the catalytic lysines, the sulfate ion and the structural analogues to Cv-TA Y322 are shown as sticks.
The positive $T_m$ shift measured upon Schiff base reduction (Figure 26), the strained K290 conformation and the presence of a tyrosine on the interfacial loop (residues 313-328) as structural analogue to the Cv-TA Y322 (Figure 27) suggest that this enzyme could obey a dimer dissociation process similar to that proposed for Cv-TA. Nonetheless, the lack of additional structural information does not allow any further conclusion.

2.4.3 PDB-wide structure search for structural analogues to Cv-TA

The structural analysis of Vf-TA, which revealed that the rearrangement of the interfacial loop occurs upon PLP loss in other fold type I transaminases, prompted a broader structural analysis to be performed on the whole RCSB PDB using the Dali server and the holo-Cv-TA structure as query. The 2187 redundant hits were manually curated to select a set of 22 structures showing 59 to 12% structural identity to Cv-TA and belonging to 17 transferases and one epimerase. The structures belonged to enzymes isolated from mesophiles, thermophiles and halophiles for a total of 8 apo structures and 14 holo structures presenting varying crystallographic quaternary assembly (monomers, dimers, tetramers and dimers of dimers) (Table 3).

2.4.4 Structural analysis of $Lb$ (EC 5.1.1.21) and $Pj$ (EC 2.6.1.x)

For two of the enzymes retrieved in the structure search, i.e. $Lb$ and $Pj$, both the apo and the holo structures are available in the RCSB PDB for inspection and direct comparison with Cv-TA (Table 3).

$Lb$, a leucine 2-epimerase (EC 5.1.1.21) crystallized as a tetramer in both enzyme forms, has a threonine as structural analogue to the Cv-TA Y322 on the interfacial loop (Figure 28). This loop, which in the holo form adopts a 3D arrangement similar to that of its structural analogues in Cv-TA and Vf-TA, is found to be disordered.
<table>
<thead>
<tr>
<th>Organism</th>
<th>Code</th>
<th>Class</th>
<th>PDB</th>
<th>PBD Holo</th>
<th>Structural Identity %</th>
<th>Interfacial loop residue</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chromobacterium violaceum WT*</td>
<td>Cv</td>
<td>2.6.1.62</td>
<td>4A6R (D)</td>
<td>4A6T (D+D)</td>
<td>100</td>
<td>Y</td>
</tr>
<tr>
<td>Vibrio fluvialis*</td>
<td>Vf</td>
<td>2.6.1.2</td>
<td>5ZTX (D)</td>
<td>4E3R (T)</td>
<td>37</td>
<td>A</td>
</tr>
<tr>
<td>Lactobacillus bacthneri</td>
<td>Lb</td>
<td>5.1.1.21</td>
<td>5LL2 (T)</td>
<td>5LL3 (T)</td>
<td>30</td>
<td>T</td>
</tr>
<tr>
<td>Pseudomonas jersenii</td>
<td>Pj</td>
<td>2.6.1.x</td>
<td>6G4B (D)</td>
<td>6G4D (D)</td>
<td>39</td>
<td>G</td>
</tr>
<tr>
<td>Halomonas elongata</td>
<td>He</td>
<td>2.6.1.x</td>
<td>-</td>
<td>6GW1 (D)</td>
<td>57</td>
<td>Y</td>
</tr>
<tr>
<td>Mesirhizobium loti</td>
<td>Mi</td>
<td>2.6.1.1</td>
<td>-</td>
<td>3GJU (M)</td>
<td>37</td>
<td>Y</td>
</tr>
<tr>
<td>Virgibacillus sp.</td>
<td>Vs</td>
<td>2.6.1.x</td>
<td>-</td>
<td>6FYQ (M)</td>
<td>37</td>
<td>Y</td>
</tr>
<tr>
<td>Paracoccus denitrificans</td>
<td>Pd</td>
<td>2.6.1.-</td>
<td>-</td>
<td>4GRX (D+D)</td>
<td>37</td>
<td>A</td>
</tr>
<tr>
<td>Silicibacter sp.</td>
<td>Ssp</td>
<td>2.6.1.x</td>
<td>-</td>
<td>3FCR (M)</td>
<td>36</td>
<td>Y</td>
</tr>
<tr>
<td>Uncultivated Pseudomonas</td>
<td>P</td>
<td>2.6.1.x</td>
<td>-</td>
<td>5LH9 (T)</td>
<td>34</td>
<td>Y</td>
</tr>
<tr>
<td>Rhodobacter sphaeroides</td>
<td>Rs</td>
<td>2.6.1.62</td>
<td>-</td>
<td>31ST (D)</td>
<td>33</td>
<td>Y</td>
</tr>
<tr>
<td>Pseudomonas aeruginosa</td>
<td>Pa</td>
<td>2.6.1.18</td>
<td>-</td>
<td>4BQ0 (T)</td>
<td>31</td>
<td>Y</td>
</tr>
<tr>
<td>Salmonella typhimurium</td>
<td>St</td>
<td>2.6.1.11</td>
<td>-</td>
<td>2PB0 (D)</td>
<td>25</td>
<td>Y</td>
</tr>
<tr>
<td>Hydrogenimonas thermophilia</td>
<td>Ht</td>
<td>2.8.1.7</td>
<td>-</td>
<td>5ZSP (D)</td>
<td>12</td>
<td>-</td>
</tr>
<tr>
<td>Pseudomonas sp.</td>
<td>Ps</td>
<td>2.6.1.1</td>
<td>5T18 (D)</td>
<td>-</td>
<td>59</td>
<td>Y</td>
</tr>
<tr>
<td>Pseudomonas putida</td>
<td>Pp</td>
<td>2.6.1.82</td>
<td>6HX9 (D)</td>
<td>-</td>
<td>55</td>
<td>Y</td>
</tr>
<tr>
<td>Silicibacter pomeroyi*</td>
<td>Sp</td>
<td>2.6.1.1/62</td>
<td>3HMU (D)</td>
<td>-</td>
<td>53</td>
<td>Y</td>
</tr>
<tr>
<td>Bacillus anthracis</td>
<td>Ba</td>
<td>2.6.1.62</td>
<td>3N5M (T)</td>
<td>-</td>
<td>29</td>
<td>F</td>
</tr>
</tbody>
</table>

[a] Completed, when possible, according to search in the BioCyc database
[b] Four-letter PDB code. Crystallographic quaternary structure in brackets: M monomer; D dimer; D+D dimer of dimers; T tetramers.
* Tested in the Schiff base reduction experiment
Figure 28. Structural analysis of the catalytic lysine and interfacial loop conformations of the apo and holo crystal structures of Lb (in green, PDB IDs: apo 5LL2, holo 5LL3) and Pj (in red, PDB ID: apo 6G4B, holo 6G4C). Interfacial loops are shown as tubes, the catalytic lysines and the structural analogues to Cv-TA Y322 are shown as sticks. Portions of missing electron density are represented with straight, discontinuous lines.

in the apo form (Figure 28).

Pj, an aspartate transaminase (EC 2.6.1.1) crystallized as a dimer in both the apo and holo form, presents a glycine both as structural analogue to the Cv-TA Y322 (Figure 28) and as residue preceding the
geometrically strained catalytic lysine K287. In this enzyme, the loss of PLP is not accompanied by any structural rearrangement around the PGBC nor close to the monomer-monomer interface (Figure 28).

2.4.5 Structural analysis of the remaining available crystal structures

For all of the remaining enzymes identified by structure search either the apo or the holo structure is known. These structures where exploited in a all-against-all comparison (apo vs apo and holo vs holo). In general, the interfacial loop conformation is identical in the holo enzymes around the position occupied by Y322 in Cv-TA and diverges at the N-terminus of the loop (Figure 29, panel b). The only exception is represented, among the considered structures, by St, where the interfacial loop is disordered in the holo form. This high degree of structural similarity at the monomer-monomer interface is absent in the considered apo structures (Figure 29, panel a), where in most instances the interfacial loop is either at least partly
disordered or organized away from the PGBC region. In the remaining cases, where the interfacial loop around the Y322 structural analogue is stretched inside the active site to shape the PGBC in the apo form, either other parts of the loop are disordered (\(Lb\)), or the Y322 analogue is a glycine (\(Pj\)) or the enzyme exists as a tetramer in the apo form (\(Ba\)).

2.4.6 **Assessment of amino acid variation among the Cv-TA Y322 structural analogues and the alignment outlier \(Ht\)**

To assess the extent of variation at the amino acid position occupied by Y322 in \(Cv\)-TA, a sequence alignment was performed. Seventeen of the eighteen unique sequences considered in this study were successfully aligned to the \(Cv\)-TA one. The alignment revealed that in 13 instances the \(Cv\)-TA Y322 structural analogue is a tyrosine, in 2 cases it is an alanine and in 3 cases it is either a threonine, a glycine or a phenylalanine (Table 3). The limited amino acid variation at this position, found to be important for the stability of the active dimer in PAPER III, suggests that the structural interplay at the dimer interface proposed in that same study might be important for other enzymes belonging to fold type I and not necessarily belonging to the transaminase class.

Only one of the sequences corresponding to the structures considered in this study could not be aligned to the \(Cv\)-TA reference sequence. The cysteine desulfurase (EC 2.8.1.7) isolated from the thermophilic organism *Hydrogenimonas thermophilia* (\(Ht\)) shows the lowest structural identity to \(Cv\)-TA and its sequence cannot be aligned to the sequence of the query \(Cv\)-TA. From the structural perspective, \(Ht\) presents the deletion of the N-terminal domain (Figure 30a), a more compact core assembly (Figure 30b) and a distinct organization of the monomer-monomer interface (Figure
Figure 30. Structure comparison between the holo Ht (in green on the left-hand side, PDB ID: 5ZSP) and the holo Cv-TA (in blue on the right-hand side, PDB ID: 4A6T). Panel a shows the 2D topology of the two enzymes. The N-terminal region of Cv-TA and missing in Ht is enclosed in a black frame. Panel b shows the 3D structure of the Ht and Cv-TA homodimers. Panel c shows the monomer monomer organization of the two enzymes. The structure segments corresponding to the interfacial loop or occupying the same region in space are shown as tubes. The Lys-PLP moiety, Y322 and its closest structural analogue are represented as sticks. 

30c), where no exact structural analogues to the Cv-TA interfacial loop or Y322 can be identified (Figure 30c). This and other similar
fold type I PLP dependent enzymes likely obey a different dimer dissociation process than the one proposed for Cv-TA and possibly followed by other enzymes belonging to this structural class, as hinted at by the structural analysis carried out in this study.

2.4.7 Conclusions from PAPER IV

In PAPER IV the investigation of the factors affecting dimer stability in fold type I PLP-dependent enzymes has been continued starting from the findings reported in PAPER III. Not surprisingly, it has been shown that the dimer dissociation mechanism initially proposed for Cv-TA is only valid among fold type I transaminases showing (S)-selectivity. The different fold type, PLP coordination and interface architecture in fold type IV transaminases showing (R)-selectivity suggest that the determinants of dimeric stability in these enzymes are different. Nonetheless, a progressively broader analysis of the apo and holo enzyme structures of a sub-group of fold type I enzymes showing structural similarity to the Cv-TA has shown that the structural determinants around the PLP-binding region found to be important for the dimeric stability of Cv-TA are conserved across this sub-group of enzymes.

While the interfacial loop conformation is remarkably preserved across the considered structures in the holo form, the release of PLP seems to trigger a wide variety of conformational rearrangements, which are likely to be determined by the particular sequence and structural surroundings of the interfacial loop in each enzyme. These reorganizations may play different roles in the multimeric equilibrium of these enzymes. Although fragmented, the structural data considered suggests that the dimer dissociation process initially proposed for Cv-TA could be obeyed by a number of enzymes (even non-transaminase) sharing the same fold and showing a similar
monomer-monomer interface architecture. In this study we also found that other fold type I PLP-dependent enzymes (of which the enzyme Ht of known structure was considered as representative) are characterized by different structural determinants at the dimer interface and do not possess exact structural analogues to the interfacial loop nor to the Cv-TA Y322 residue. Future more collective efforts directed towards the investigation of a larger number of PLP-dependent fold type I enzymes shall lead to a better understanding of the stability determinants within this group of enzymes and to the development of suitable strategies for the improvement of their stability, regardless of their catalytic activity.
3. Concluding remarks and outlook

The presented thesis work has focused on the investigation of several aspects of transaminase biocatalysis, from the development of novel applications (PAPER I) to the broadening of our current understanding of the biochemical basis determining transaminase stability (PAPER III), through the design of a new transaminase variant showing altered substrate acceptance and improved cosolvent tolerance (PAPER II). In PAPER IV, an attempt was made to extend the findings about the multimeric stability of the model (S)-selective transaminase from *Chromobacterium violaceum* to other (S)-TAs as well as other enzymes belonging to the same fold type, with promising preliminary data. The findings compiled in this thesis contribute therefore to advance the utility and the basic understanding of transaminases, particularly with respect to the (S)-selective ones.

The design of a novel effective transaminase-based racemization system for primary amines is expected to prompt further research efforts both in industry and academia. The possibility to accomplish amine racemization under the mild reaction conditions used is promising for the development of highly desired DKR systems targeting valuable chiral amines. Such methods double the yield of enantiopure amine compared to “simple” KR systems with concomitant consumption of half equivalent of cheaper chiral amino acids. DKRs could therefore be competitive with established approaches under both the safety and atom economy perspectives. As additional benefit, the use of naturally accepted amino
acids/ketoacids as coupled amino donors/acceptors simplifies the substrate engineering (when necessary) to the target valuable amine.

The successful structure-guided engineering of the small substrate binding pocket of Cv-TA, which led to a robust enzyme variant capable of accepting 1,2-diphenylamines, will hopefully be followed up to broaden the available biocatalytic toolbox dedicated to access similar compounds. Although in this thesis the generated enzyme variant could only be used in preparative KR reactions, its exploitation in AS setups should be possible and further research should be dedicated to the achievement of this goal.

The engineering of (S)-selective TAs, both in terms of stability and substrate acceptance, has proven so far more challenging than for (R)-selective transaminases, possibly due to their more complex multimeric architecture. In this thesis work, the dynamic interplay between the subunits of (S)-TAs multimers has been acknowledged for the first time and recognized as one of the main factors undermining the multimeric stability of these enzymes. This new perspective on (S)-TAs stability will hopefully aid the development of targeted methods to improve the operational stability of this group of biocatalysis as well as others with similar multimeric architectures.

In conclusion, the content of this thesis will hopefully direct future research in both industry and academia to access chiral amines through safer and simpler routes relying on more stable, versatile and efficient transaminase variants engineered via more rational approaches. The lessons learned from the transaminase case will hopefully come useful to direct future efforts in the understanding and exploitation of other PLP-dependent enzymes with unrealized industrial potential.
4. Acknowledgements

About five years ago I wrote the acknowledgement section of my master’s thesis on a train ride between home and university. At that time I had so many doubts about what was going to happen after the university chapter. I remember writing that I had always felt just like a climber hanging from a cliff, slowly ascending and constantly looking for the next grip. I remember having a special “thank you” for all the people that had been cheering for me on my way up.

Five years later, as I take a moment to look down at the last stretch of vertical cliff, I have a clear view of all the cheering faces that have encouraged me, helped me, pointed at the best next grip and that have made it possible for me to go this far without giving up. This time my “thank you” must be so much larger, so much louder and spoken out in so many different languages.

Il grazie più grande va ai miei fantastici genitori Carla e Antonio. Non solo mi avete sempre sostenuto nelle mie scelte, ma mi avete incoraggiato nel perseguirle anche quando avreste forse preferito che il meglio per me fosse diverso o fosse altrove. Quindi grazie per essere in quella esigua percentuale di genitori capaci di tanto. Un grande grazie anche a tutta quella parte della mia famiglia che è riuscita a capire le mie scelte e che non mi ha mai fatto sentire sola, nemmeno nelle giornate più buie dei bui inverni di qui.

The second largest “thank you” goes to all the people and institutions that have given me the gigantic chance of embarking in this adventure. Thank you to all the members of the Marie-Curie ITN-EID BIOCASCADES network (H2020-MSCA-ITN-2014, grant number 634200) that have participated with energy and enthusiasm in this initiative. Thank you to all of the ESRs for sharing this
adventure and many others outside of the lab. Thank you to my supervisors Per Berglund (KTH), Björn Walse (SARomics Biostructures AB), Derek Logan (SARomics Biostructures AB) and Luuk van Langen (Viazym BV) for the huge chances, for direction, for advice and for support.

Huge thanks to all of the fantastic people at Viazym BV, SARomics Biostructures AB and KTH with whom I have had the luck to share the best part of the last 4 (and a half) years. I would need some 73 extra pages to name you all one by one, so forgive me if I won’t try. You know who you are. A special thank you also goes to all the past and present members of the Biocat group at KTH and to all of my co-authors.

These four years have certainly been very busy with things to do, places to go and challenges to face. A whole universe of fantastic people have filled the gaps between one challenge and the next, between one day in the lab and the next, between one Friday and the following Monday. To you, fantastic people, I owe another huge thank you.

I always think that these four years have given me way more than what they have taken, and, David, you are the prime example of that. Thank you for having found me. Thank you for having been so determined and so supporting, especially when I needed it the most. Thank you for your attempts to learn Italian, for forcing me to take better care of myself and for sending me a never-ending series of mood-boosting puppies (among all the other other things). I do not know what these last years would have been without you. You brighten up my sky.

Stora tackar till Lena, Stiegge och Lars, eftersom aldrig skulle jag ha kunnat föreställa mig en varmare välkommande i Örnsköldsvik. Jag
kände mig en del av ert hus från första ögonblicket. Tack för att ni accepterade mig i er familj. Tack för era nyfikenhet, öppenhet och tålamod. Jag lovar att mitt svenska kommer att förbättras...

Thank you all. I could have not won over this particularly challenging vertical cliff without your support.

I will stop here for a bit, catch my breath, enjoy this moment and gather my strengths before keeping on climbing.
5. References


74. Börner, T., Rehn, G., Grey, C. & Adlercreutz, P. A Process Concept for High-Purity Production of Amines by Transaminase-Catalyzed


