

University of Groningen

## Redirected evolution of a proline-based tautomerase

Biewenga, Lieuwe

DOI:  
[10.33612/diss.151476975](https://doi.org/10.33612/diss.151476975)

**IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.**

*Document Version*  
Publisher's PDF, also known as Version of record

*Publication date:*  
2020

[Link to publication in University of Groningen/UMCG research database](#)

*Citation for published version (APA):*  
Biewenga, L. (2020). *Redirected evolution of a proline-based tautomerase: New tools for carboligation reactions*. University of Groningen. <https://doi.org/10.33612/diss.151476975>

### Copyright

Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

The publication may also be distributed here under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license. More information can be found on the University of Groningen website: <https://www.rug.nl/library/open-access/self-archiving-pure/taverne-amendment>.

### Take-down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

*Downloaded from the University of Groningen/UMCG research database (Pure): <http://www.rug.nl/research/portal>. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.*

*Redirected evolution of a  
proline-based tautomerase*

New tools for carbonylation reactions

Lieuwe Biewenga

The research described in this thesis was carried out in the Department of Chemical and Pharmaceutical Biology (Groningen Research Institute of Pharmacy, University of Groningen, The Netherlands) and was financially supported by the European Union's Horizon 2020 research and innovation program under grant agreement No 635595 (CarbaZymes) and the Netherlands Organization of Scientific Research (VICI grant 724.016.002).

The research work was carried out according to the requirements of the Graduate School of Science, Faculty of Science and Engineering, University of Groningen, The Netherlands.

Printing of this thesis was financially supported by the University Library and the Graduate School of Science, Faculty of Science and Engineering, University of Groningen, The Netherlands.

Cover design: Lieuwe Biewenga

Printing: Ridderprint | [www.ridderprint.nl](http://www.ridderprint.nl)

Design: Elisa Calamita, [persoonlijkproefschrift.nl](http://persoonlijkproefschrift.nl)

ISBN: 978-94-034-2639-6

Electronic ISBN: 978-94-034-2638-9

Copyright © 2020 Lieuwe Biewenga. All rights are reserved. No part of this thesis may be reproduced or transmitted in any form or by any means without the prior permission in writing of the author.



university of  
 groningen

# Redirected evolution of a proline- based tautomerase

New tools for carbonylation reactions

**PhD thesis**

to obtain the degree of PhD at the  
University of Groningen  
on the authority of the  
Rector Magnificus Prof. C. Wijmenga  
and in accordance with  
the decision by the College of Deans.

This thesis will be defended in public on

Friday 13 November 2020 at 16:15 hours

by

**Lieuwe Biewenga**

born on 4 October 1990  
in Ede

## **Supervisors**

Prof. dr. G.J. Poelarends

Prof. dr. W.J. Quax

## **Assessment Committee**

Prof. dr. ir. M.W. Fraaije

Prof. dr. A.K.H. Hirsch

Prof. dr. U. Hanefeld

*The fear of the LORD is the beginning of Wisdom*

Psalm 111:10



## Table of Contents

	Aim and outline of this thesis	8
<b>Chapter 1</b>	New Trends in Enzyme Engineering: the Generation and Exploitation of Protein Mutability Landscapes for Enzyme Engineering	15
<b>Chapter 2</b>	Enantioselective Synthesis of Pharmaceutically Active $\gamma$ -Aminobutyric Acids Using a Tailor-Made Artificial Michaelase in One-Pot Cascade Reactions	35
<b>Chapter 3</b>	Tuning Enzyme Activity for Nonaqueous Solvents: Engineering of an Enantioselective ‘Michaelase’ for Catalysis in High Concentrations of Ethanol	101
<b>Chapter 4</b>	A Selective Colorimetric “Turn-on” Probe for Efficient Engineering of Iminium Biocatalysis	117
<b>Chapter 5</b>	In Situ Acetaldehyde Synthesis for Carbolygation Reactions	147
<b>Chapter 6</b>	Enantioselective Aldol Addition of Acetaldehyde to Aromatic Aldehydes Catalyzed by Proline-based Carbolygases	171
<b>Chapter 7</b>	Summary and Future Perspectives	241
<b>Appendices</b>	Nederlandse samenvatting voor de geïnteresseerde leek	254
	Acknowledgement	264
	List of publications	268



## **Aim and outline of this thesis**

Enzymes are macromolecules that are used by all organisms to catalyze reactions necessary to sustain life. The vast number of different chemical transformations that take place within cells has resulted in a plethora of unique, highly specialized enzymes. The application of enzymes (purified, as cell free extract or in whole cells) as catalysts for chemical synthesis is referred to as biocatalysis. Due to their biodegradability and high efficiency under mild reaction conditions, enzymes have become a strong alternative to traditional metal catalysts for the development of efficient and sustainable chemical processes. Especially for transformations that involve the formation of one or more asymmetric carbon atoms, enzymes are often the number one choice catalyst, as the enantio- and regioselectivity of enzymes are unparalleled. Despite the highly diverse reactions that can be catalyzed by enzymes, most enzymes are optimized towards the relatively uniform reaction condition of the cell. As a consequence, it is often possible to find reaction conditions that ensure effective catalysis of several vastly different chemical transformations in one pot using enzymes as catalysts. The highly specialized nature of enzymes also limits the cross-reactivity towards intermediate products, which maximizes the yield of the final product. This has enabled scientists to combine enzymes from various sources and construct artificial metabolic networks that can be used for the efficient synthesis of important pharmaceuticals. Recent advances in recombinant DNA technologies and large scale fermentations have facilitated the sustainable and cost-effective production of enzymes. Therefore, the application of enzymes in chemical synthesis is expected to further increase.

However, the application of enzymes for chemical synthesis has brought new challenges. Due to the highly specialized nature of enzymes, it can be difficult to find a suitable biocatalyst for a given desired chemical transformation. In fact, for a number of highly important classes of chemical transformations in organic synthesis, no natural enzymes have been discovered. This has prompted researchers to investigate the use of enzymes as biocatalysts for other reactions than their natural reactions. Indeed, many enzymes show a certain degree of promiscuity towards alternative substrates (substrate promiscuity) and some enzymes can even catalyze completely different reactions (catalytic promiscuity).

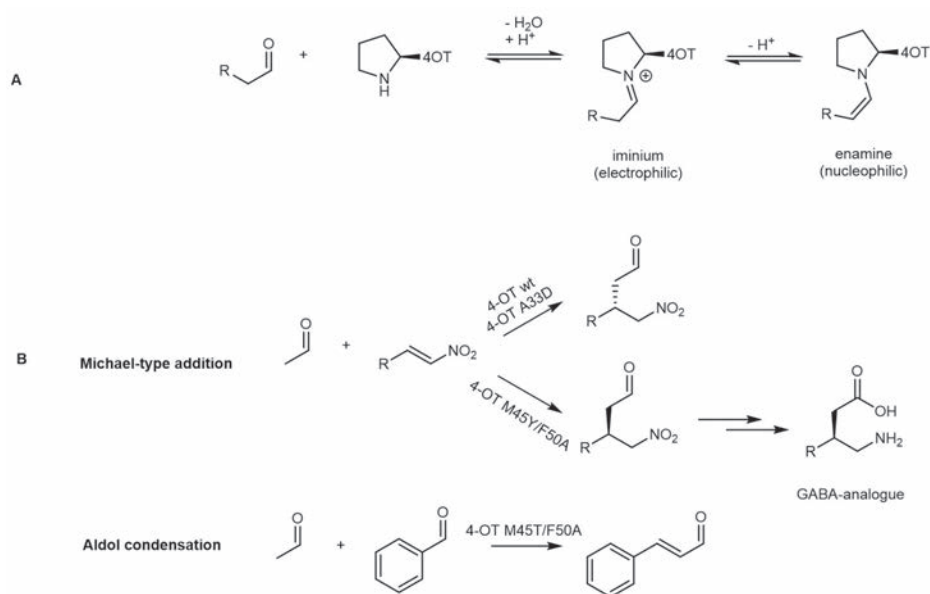
The catalytic efficiency of enzymes towards non-natural reactions is typically considerably lower compared to their natural reactions, as enzymes have not evolved for these reactions. To improve the catalytic efficiency of enzymes towards non-natural reactions, scientists have applied a protein engineering methodology that is referred to as directed evolution. Similar to natural evolution, directed evolution starts with

the introduction of variation by means of mutations. Mutants that show the desired characteristic (improved catalytic efficiency towards a non-natural reaction) are selected and subjected to a new round of mutagenesis and selection. This iterative process can be repeated many times until a mutant enzyme is found that shows a high catalytic efficiency towards the desired non-natural reaction. Directed evolution is not only applied to engineer the catalytic efficiency of enzymes, but can be used to engineer any characteristic of an enzyme (e.g. solvent tolerance, thermostability, enantioselectivity, etc.).

A prime example of a catalytically promiscuous enzyme is 4-oxalocrotonate tautomerase (4-OT). This exceptionally small enzyme (only 62 amino acids) naturally catalyzes the enol-keto tautomerization of 2-hydroxy-2,4-hexadienedioate to 2-oxo-3-hexenedioate, using the amino-terminal proline (Pro-1) as key catalytic base. However, because of the unusually low  $pK_a$  of Pro-1 ( $pK_a \sim 6.4$ ), it can also act as a nucleophile and react with aldehydes to form highly reactive enamine or iminium species (Figure 1). In the field of organocatalysis, iminium or enamine activation by small proline-derived catalysts has formed the basis for many different asymmetric C-C bond-forming transformations. This has inspired Poelarends and coworkers to investigate 4-OT as a potential biocatalyst for these reactions. Contrary to organocatalysts that typically only function in organic solvents due to their low solubility and relatively high  $pK_a$  in water, Pro-1 of 4-OT folds into a hydrophobic active site which lowers its  $pK_a$  and allows for catalysis in water. It was discovered that 4-OT can catalyze several promiscuous C-C bond-forming reactions, including the Michael-type addition of acetaldehyde to nitroalkene acceptors and the aldol condensation of acetaldehyde with benzaldehyde. Interestingly, there are no enzymes known that can naturally catalyze these reactions. The products of the 4-OT catalyzed Michael-type additions and aldol condensations form important chemical building blocks with various applications. For instance, the 4-OT catalyzed Michael-type addition of acetaldehyde to nitroalkenes yields  $\gamma$ -nitroaldehydes, which are important precursors to  $\gamma$ -aminobutyric acid (GABA) analogues, a class of abundantly prescribed pharmaceuticals.

Poelarends and coworkers have started directed evolution efforts on 4-OT to identify mutants with improved activity towards different promiscuous activities. For instance, 4-OT M45T/F50A was identified as a promising aldolase, showing a 3300-fold improvement in catalytic efficiency compared to 4-OT wild-type (wt). Similarly, 4-OT A33D showed a 3.5-fold increase in 'Michaelase' activity. An important goal has been to invert the enantioselectivity of the 4-OT catalyzed Michael-type addition of acetaldehyde to nitroalkenes, as 4-OT wt is enantioselective towards the pharmaceutically irrelevant

enantiomer. Important steps have been made, resulting in the identification of a double mutant, 4-OT M45Y/F50A, with inverted enantioselectivity compared to 4-OT wt albeit its enantioselectivity is not excellent. The aim of the work described in this thesis was to use enzyme engineering to further improve 4-OT for enantioselective C-C bond-forming reactions and apply the best enzyme variants in new (chemo)enzymatic reaction cascades towards the synthesis of important pharmaceuticals.



**Figure 1. Aldehyde activation by 4-OT and promiscuous reactions catalyzed by 4-OT.** A) Aldehydes reacting with 4-OT forming electrophilic iminium or nucleophilic enamine species. B) Michael-type addition and aldol condensation reactions catalyzed by 4-OT.

The success of directed evolution largely depends on the strategy applied to introduce mutations in a target enzyme and the method used for selection of improved enzyme variants. In **chapter 1**, we review an emerging enzyme engineering technique referred to as mutability-landscape-guided enzyme engineering. The principal behind this technique is to map the effect (detrimental, neutral and beneficial) of all, or nearly all, single mutations of an enzyme on a desired characteristic. This information can consequently be used to design smart libraries to efficiently engineer an improved enzyme.

In **chapter 2**, we describe our engineering efforts to further improve the enantioselectivity of 4-OT for the Michael-type addition of acetaldehyde to nitroalkene acceptors. The previously engineered 4-OT M45Y/F50A was used as a starting point. We show

that the enantioselectivity of 4-OT M45Y/F50A can be improved by performing the reaction in the presence of small diols, such as ethylene glycol and 1,3-propanediol, as co-solvents. Further structure-guided enzyme engineering afforded two triple mutants, 4-OT L8F/M45Y/F50A and 4-OT L8Y/M45Y/F50A, with improved activity and near perfect enantioselectivity. Comparison of a crystal structure of 4-OT L8Y/M45Y/F50A with a crystal structure of 4-OT M45Y/F50A in complex with nitrostyrene revealed that the volume of the binding pocket of 4-OT L8Y/M45Y/F50A is reduced. This is likely to restrict the rotational freedom of the substrate, which could explain the high enantioselectivity of this mutant. We show that this artificial ‘Michaelase’ can be combined with a natural aldehyde dehydrogenase and a chemocatalyst (nickel boride) to synthesize a range of GABA-analogues in one pot in high overall yields (up to 70%) and with excellent enantiopurity (e.r. up to 99:1).

In **chapter 3**, we report mutability-landscape-guided enzyme engineering to improve 4-OT for catalysis in elevated ethanol concentrations. We identified two “hot-spot” positions in the enzyme that contribute to ethanol tolerance, Ser-30 and Ala-33. Mutating “hot-spot” position Ala-33 in the context of the highly enantioselective, but ethanol-sensitive, mutant 4-OT L8F/M45Y/F50A afforded several highly solvent-stable mutants, that allow enantioselective catalysis in the presence of up to 40% v/v ethanol.

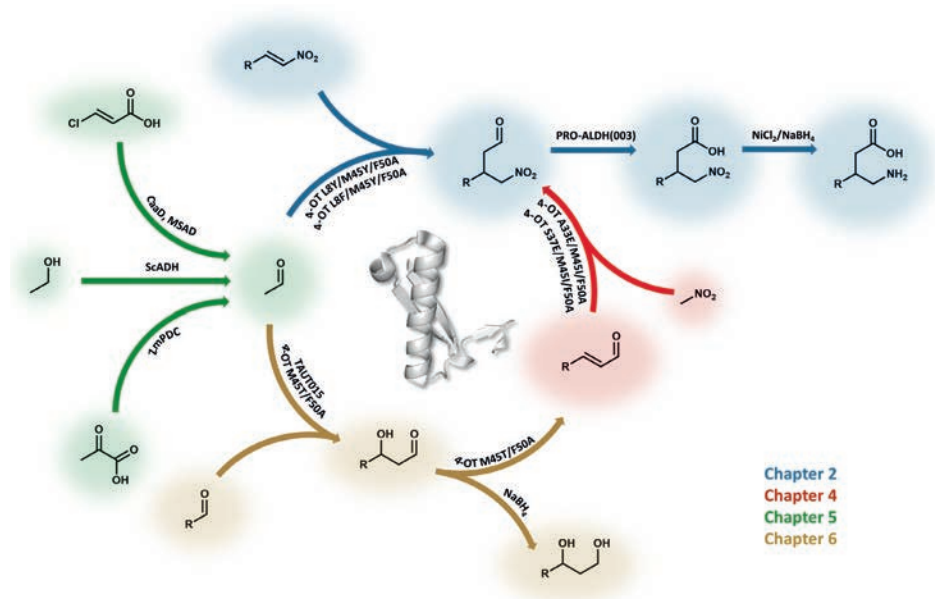
In **chapter 4**, we describe the engineering of 4-OT for a complementary enzymatic route towards  $\gamma$ -nitroaldehydes, via the Michael-type addition of nitromethane to cinnamaldehyde. The key catalytic step in this reaction is the iminium-activation of cinnamaldehyde by Pro-1 of 4-OT. We show that iminium activation of a mimic of cinnamaldehyde, 2-hydroxycinnamaldehyde, by Pro-1 of 4-OT results in the formation of a brightly colored species resembling a merocyanine dye. This iminium-activated colorimetric “turn-on” probe can be used as a pre-screening tool, as many inactive mutants do not form this brightly colored species. We developed a facile solid-phase pre-screening assay, termed as Activated Iminium Colony Staining (AICS), which reduced the screening burden up to 20-fold. After 2 rounds of directed evolution we identified two new mutants, 4-OT A33E/M45I/F50A and 4-OT S37E/M45I/F50A, with up to 39-fold increase in activity compared to the parental enzyme 4-OT F50A.

The highly reactive and toxic nature of acetaldehyde requires intricate handling, which can impede its usage in practical synthesis. In **chapter 5**, we describe our efforts to develop three enzymatic routes for *in situ* synthesis of acetaldehyde from a less toxic and less reactive precursor such as *trans*-3-chloroacrylic acid, ethanol or pyruvate. Two routes, using either *trans*-3-chloroacrylic acid or ethanol as starting substrate, afforded

effective concentrations of acetaldehyde and could be applied with 4-OT in one-pot cascade reactions to prepare pharmaceutical building blocks.

In **chapter 6**, we describe a study on the promiscuous aldolase activity of 4-OT. Earlier, a 4-OT mutant (4-OT M45T/F50A) with strongly enhanced aldolase activity for the condensation of acetaldehyde with benzaldehyde was identified. We show that this mutant can accept a wide range of aldehyde acceptors. Moreover, we identified four different aldehyde acceptors that undergo aldol coupling with acetaldehyde without noticeable dehydration of the resultant aldol products, resulting in the accumulation of chiral  $\beta$ -hydroxyaldehydes in the reaction mixtures. After reduction by  $\text{NaBH}_4$ , the corresponding 1,3-diols could be isolated in good yield and with excellent enantiopurity. Finally, we have screened a collection of 4-OT homologues and identified a synthetically useful carboligase, TAUT015, that was successfully applied as biocatalyst to expand the scope of accessible chiral 1,3-diols.

In **chapter 7**, we summarize the work described in this thesis and provide some future perspectives. An overview of the (chemo)enzymatic reactions reported in this thesis is presented in Figure 2.



**Figure 2. Overview of the (chemo)enzymatic reactions described in this thesis.** Abbreviations: CaaD: chloroacrylic acid dehalogenase, MSAD: malonate semialdehyde decarboxylase, ScADH: alcohol dehydrogenase, ZmPDC: pyruvate decarboxylase, TAUT015: 4-OT homologue, PRO-ALDH(003): aldehyde dehydrogenase.





# Chapter 1

## *New Trends in Enzyme Engineering: The Generation and Exploitation of Protein Mutability Landscapes for Enzyme engineering*

---

Jan-Ytzen van der Meer,<sup>[a]</sup> Lieuwe Biewenga<sup>[a]</sup> and Gerrit J. Poelarends<sup>\*,[a]</sup>

<sup>[a]</sup>Mr. J.-Y. Van der Meer, Mr. L. Biewenga, Prof. Dr. G.J. Poelarends

<sup>a</sup>Department of Chemical and Pharmaceutical Biology, Groningen Research Institute of Pharmacy,  
University of Groningen, Antonius Deusinglaan 1, 9713 AV Groningen, The Netherlands.

\*Corresponding author. Tel.: +31503633354; E-mail: [g.j.poelarends@rug.nl](mailto:g.j.poelarends@rug.nl);

Web: <http://www.rug.nl/staff/g.j.poelarends/>

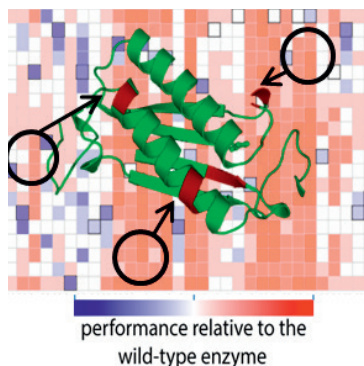
Published in *ChemBioChem* 10.1002/cbic.201600382



## Abstract

The increasing number of enzyme applications in chemical synthesis calls for new engineering methods to develop the biocatalysts of the future. An interesting concept in enzyme engineering is the generation of large-scale mutational data to chart protein mutability landscapes. These landscapes allow the important discrimination between beneficial mutations and those that are neutral or detrimental, providing detailed insight into sequence-function relationships. As such, mutability landscapes are a powerful tool to identify functional hotspots at any place in the amino acid sequence of an enzyme. These hotspots can be used as targets for combinatorial mutagenesis to yield superior enzymes with improved catalytic properties, stability or even new enzymatic activities. The generation of mutability landscapes for multiple properties of one enzyme provides the exciting opportunity to select mutations that are beneficial for either one or several of these properties. This review presents an overview of the recent advances in the construction of mutability landscapes and discusses their importance for enzyme engineering.

**Keywords:** Mutability landscapes; Enzyme engineering; biocatalysis, mutations, hotspots



## 1. Advantages of enzyme catalysis

Application of enzymes as catalysts in the production of chemicals has the potential of being a sustainable and efficient alternative to traditional catalysts used in organic synthesis. Enzymes are nature's catalysts and therefore generally function under mild reaction conditions (i.e., ambient temperatures in aqueous solvent systems). Furthermore, enzymes are biodegradable, non-toxic, readily available, and their production is not dependent on any rare elements. These features underline the sustainable potential of using enzymes as catalysts. Another important feature of enzymes is their excellent catalytic properties which can make them highly efficient catalysts. Enzymes are known for their high catalytic rates and excellent regio-, chemo- or stereoselectivity. Enantioselectivity is still a major challenge in traditional catalysis and is highly desirable for the production of pharmaceuticals. Finally, enzymes can be optimized for their application in industrial biocatalysis by means of protein engineering. Owing to these advantages, the number of applications for enzyme catalysts in the production of valuable chemicals, especially pharmaceuticals and agrochemicals, is increasing.<sup>[1-3]</sup>

## 2. Why is enzyme engineering required?

Typical goals of engineering projects in the field of biocatalysis can be divided into three topics. The first topic has a focus on the catalytic properties of enzymes and includes engineering projects which aim to improve catalytic activity, alter substrate scope or improve (enantio-)selectivity. As a result of engineering projects in this theme, there are now many examples of enzymes which carry out industrially relevant transformations, with practical turnover rates.<sup>[1-3]</sup> The second topic covers enzyme engineering projects which aim to improve enzyme stability. Enzymes can be unstable under process conditions, which may include high temperatures, extreme pH values, high substrate (and product) concentrations and/or the presence of organic solvents. Major improvements in enzyme stability can be achieved using enzyme engineering.<sup>[4]</sup> Alternatively, solvent engineering or enzyme immobilization can be used to address these stability issues. These methods have recently been reviewed elsewhere.<sup>[4-7]</sup> The third topic in enzyme engineering is the generation of enzymes which catalyze unnatural chemical transformations. Creating enzymes with new enzymatic activities is currently one of the frontiers in biocatalysis and there are two main approaches to achieve this. Firstly, the *de novo* computational design of enzymes, which involves the computational design of an active site and placing it in a suitable protein scaffold.<sup>[8-10]</sup> Enzyme engineering is required to improve the activity of the initial *de novo* designed protein to a practical

level. The second approach to create enzymes with new activities is to exploit the catalytic promiscuity of existing enzymes. Promiscuous activities are enzymatic activities other than the activity for which an enzyme has evolved and that are not part of the organism's physiology.<sup>[11]</sup> It has been long recognized that promiscuous activities can serve as a starting point for natural evolution of new enzymatic functions.<sup>[12, 13]</sup> Using nature's approach, enzyme engineering can be applied to improve promiscuous activities for the generation of novel biocatalysts for unnatural chemical transformations.<sup>[14]</sup>

### 3. Hotspot identification for enzyme engineering

Enzyme engineering can be viewed as an iterative procedure which starts with generating diversity in the wild-type (WT) enzyme and screening a collection of mutants for the desired properties. To efficiently engineer enzymes, researchers try to identify hotspot positions in an enzyme where mutations are likely to be beneficial.<sup>[15]</sup> Targeting these sites for combinatorial mutagenesis leads to relatively small libraries with a high percentage of positive hits. The identification of these hotspots requires extensive knowledge on the sequence-function relationships of an enzyme and the main ways to obtain this information is by analyzing the (crystal) structure of the enzyme, multi-sequence alignments (MSAs) of homologues proteins or empirical mutational data.

Hotspot identification based on the structure of an enzyme is the most commonly used method in enzyme engineering. Damborski and co-workers recently published an extensive review on *in silico* hotspot identification methods which are available as web tools.<sup>[16]</sup> The majority of these tools are structure-based and therefore require a crystal structure of the enzyme. The computational tools then identify hotspot positions, based on predicted protein-ligand interactions, binding pockets or residues present in access tunnels of enzymes with a buried active site. Computational tools for the identification of hotspots to improve enzyme stability are mainly based on crystallographic B-factors, although computational protein design and consensus methods are gaining momentum in this area.<sup>[4,16,17]</sup> Besides these *in silico* approaches, several experimental, semi-rational, structure-based enzyme engineering methods have been developed, which apply targeted site-saturation mutagenesis on active site residues. These methods include the highly successful CASTing method and derivatives thereof.<sup>[3,18,19]</sup>

Homology-based hotspot identification tools require a MSA of homologues proteins to identify the evolutionary conservation of specific amino acid residues in a protein. High conservation scores suggest that a specific residue is important for the structure

or function of the protein, whereas low conservation suggests that this residue may be mutated without the loss of function. Targeting positions with mutational robustness, therefore, increases the chance of obtaining viable mutant enzymes and thereby increases the quality of the library.<sup>[16]</sup>

The third basis on which hotspot identification can be conducted is empirical data. This data can be generated by screening libraries which were created using random mutagenesis methods such as error prone PCR. The hotspots identified in these libraries can be targeted by combinatorial site-saturation mutagenesis.<sup>[20]</sup> The main advantage of this approach is that it does not require extensive prior knowledge of the target enzyme.

Obviously, there are tools available which combine information from all three sources. A successful example of this is the PROSAR method. Here, a collection of enzyme variants which carry multiple mutations per sequence are generated and empirically tested for the desired activity. The initial pool of enzyme variants covers mutations which are selected based on a combination of structural information, analysis of MSAs and random mutagenesis.<sup>[21]</sup> By statistical analysis of the screening results, the PROSAR software tool then evaluates the contribution of each individual mutation in each enzyme variant with multiple mutations. The identified residue positions with beneficial mutations are used for the subsequent rounds of diversification and screening. This cycle is repeated until the engineering goal is met.

#### **4. Protein mutability landscapes**

An interesting concept in enzyme engineering is the generation and use of mutability landscapes. For this type of analysis, a large number of protein variants are analyzed to determine the effect of each single amino acid substitution on enzyme activity, selectivity or stability, providing detailed maps of beneficial, neutral and detrimental amino acids for each residue position and each enzyme property. The generation of mutability landscapes for multiple properties of one enzyme provides different landscapes, with the exciting opportunity to select mutations that are beneficial for either one or several of these properties and neutral or detrimental for others. Thus, in contrast to other systematic mutagenesis approaches such as gene site-saturation mutagenesis (GSSM), mutability landscapes do not only provide information on beneficial mutations but also on detrimental and neutral mutations. This gives valuable information on sequence-function relationships by revealing regions in the enzyme with mutational robustness as well as functionally important residues and hotspot positions.

The term ‘mutability landscape’ was first used by Rost and co-workers, who developed the screening for non-acceptable polymorphisms (SNAP) algorithm to predict the effect of single amino acid substitutions in disease related proteins.<sup>[22]</sup> The predictions of this SNAP algorithm are based on information from both a MSA and structural features of the protein of interest.<sup>[23]</sup> Alternatively, the sorting intolerant from tolerant (SIFT) algorithm can be used to make similar predictions based on residue conservation.<sup>[24]</sup> Both methods predict whether an amino-acid substitution will be neutral or lead to a functional effect but do not distinguish between detrimental or beneficial effects. This is sufficient when merely looking at pathogenicity because both gain-of-function and loss-of-function-mutations can lead to disease. However, it is of limited use when this mutability landscape is generated for enzyme engineering purposes.

Hecht *et al.* argue that the lack of comprehensive experimental mutagenesis data seems a crucial problem for the development of better computational tools and that the generation of such experimental data is constrained by the amount of required resources.<sup>[22]</sup> Indeed, available data from experimental protein mutability landscapes is scarce and the majority of these available studies cover protein-protein interactions or protein-DNA interactions.<sup>[25-27]</sup> In the last few years, however, there have been several reports on experimental mutability landscapes of enzymes. Here we present an overview of the recent advances in experimental mutability landscapes of enzymes to illustrate how these mutability landscapes were generated and used to gain insight in sequence-function relationships or exploited for enzyme engineering.

## **5. Generating mutability landscapes using a defined collection of single mutants**

There are two approaches for generating experimental protein mutability landscapes. The first approach involves the characterization of a defined collection of single mutants and the second approach is called deep mutational scanning (Fig. 1). To construct a defined collection of mutant enzymes, which covers (nearly) all possible single amino acid substitutions of an enzyme, requires significant effort and resources, but the characterization of the mutants can be relatively easy as it does not require any oversampling. Therefore, the screening methods are not limited to a high-throughput assay, giving more flexibility in the design of the assays and providing access to a broader range of analyses (e.g. HPLC, UV-spectroscopy). The following examples of mutability landscapes were generated using this approach.

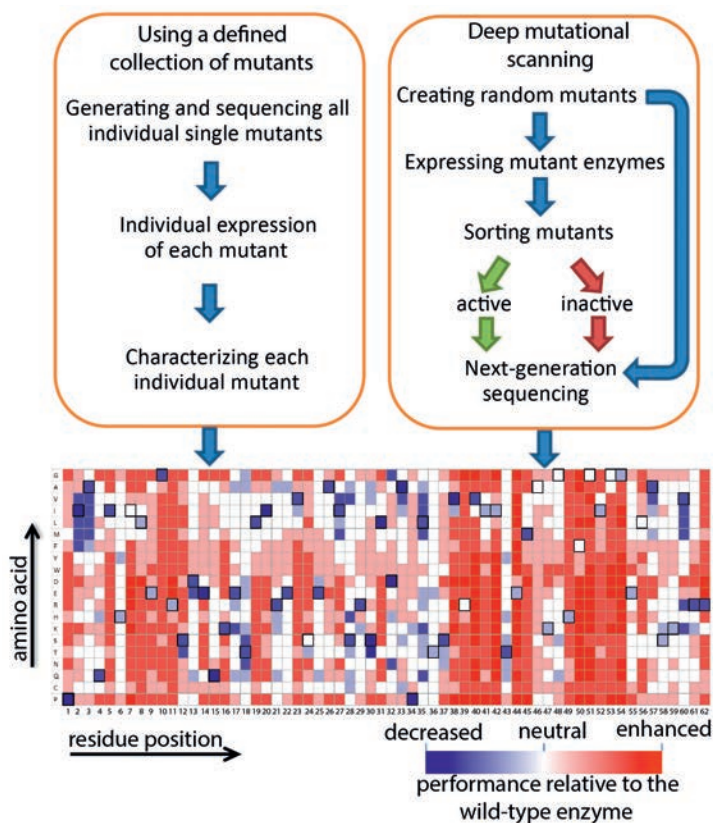


Figure 1. General methods for generating mutability landscapes.

### 5.1. Protease activity and stability

The usage of ‘site evaluation libraries’ described in a patent of Estell and Aehle was basically the first example where a mutability landscape of an enzyme was generated and applied in enzyme engineering.<sup>[28]</sup> The inventors used a defined collection of single mutants of a serine protease from *Cellulomonas* strain 69B4 (ASP), which covered at least 12 variants on each of its 189 residue positions. All members of this collection were screened for protease activity on three substrates (keratin, casein and succinyl-alanine-alanine-proline-phenylalanine-p-nitroanilide), for thermostability and for stability in the presence of 0.06% dodecylbenzenesulfonate (LAS). The performance of each mutant was scored as the apparent change of free energy in the process of interest, relative to WT ASP ( $\Delta\Delta G_{app}$ ). This value was calculated using the following formula:  $\Delta\Delta G_{app} = -RT\ln(P_{var}/P_{wt})$ , where  $P_{var}$  is the performance value of the variant and  $P_{wt}$  is the performance value of WT ASP. Therefore, negative  $\Delta\Delta G_{app}$  values indicate improved performance of the

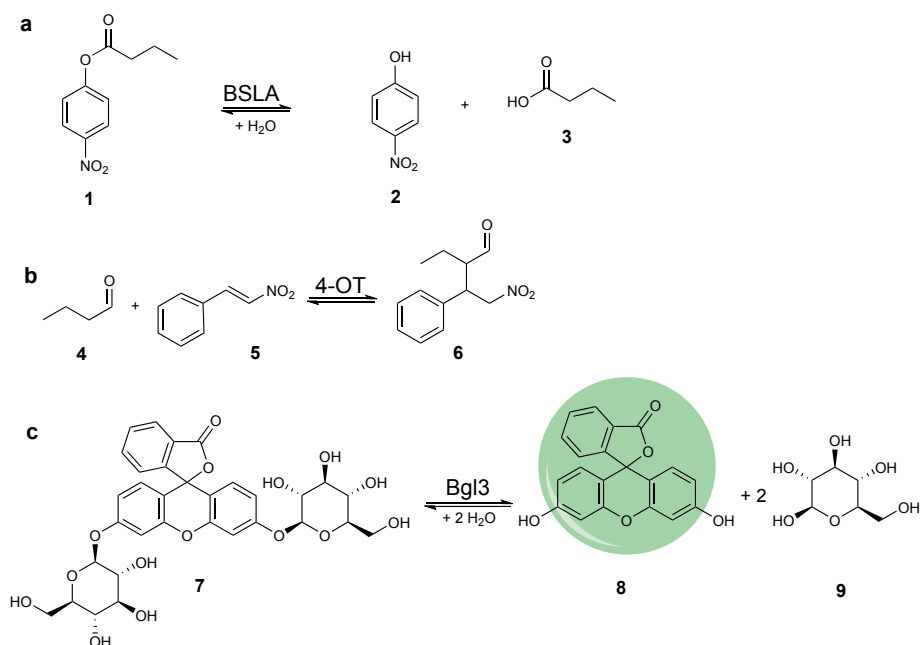
variant, relative to WT ASP. The majority (84% - 94%) of the 2851 analyzed single mutants performed worse compared to WT ASP based on activity or stability. Interestingly, 5-10% of the positions in ASP contained mutations that were deleterious for all analyzed properties. As the residues at these positions were also highly conserved among 20 non-redundant homologs of ASP, the authors concluded that these residues are required for the structural fold of the enzyme. Another remarkable finding was that most mutations which led to improved protease activity, where on positions located outside of the enzyme's active site. For example, the closest residue position at which mutations led to improved protease activity on keratin (Arg-14) was 13 Å away from the catalytic Ser-137. Therefore, targeted saturation mutagenesis on active site residues would most likely not have led to the identification of improved mutants for this reaction.

One unique advantage of this mutability landscape analysis is that it provides information on mutations which lead to the simultaneous improvement of multiple properties. For example, four positions were identified at which mutations led to both improved protease activity towards keratin and improved stability in the presence of LAS. These four positions were simultaneously randomized and the quality of the resulting library was determined based on the performance of 64 randomly picked mutants in both the activity and the LAS stability assay. The average observed performance of these mutants exceeded the expected average performance of the library members, which was calculated based on the assumption of additive effects of single mutations at the four sites. This indicated that information from the mutability landscape of an enzyme can provide valuable guidance for enzyme engineering.

## 5.2. Mutability landscapes for improved detergent stability

The large  $\alpha/\beta$ -hydrolase fold superfamily includes a broad range of synthetically useful enzymes.<sup>[29]</sup> Fulton *et al.* generated complete mutability landscapes of *Bacillus subtilis* lipase A (BSLA), which is an  $\alpha/\beta$ -hydrolase fold superfamily member, for stability in the presence of different detergents.<sup>[30]</sup> Therefore, the authors constructed a defined collection of single mutants, covering each amino acid substitution at each residue position of BSLA. This collection was constructed by performing site-saturation mutagenesis at each of the 181 residue positions in BSLA. The resulting 181 libraries were subsequently used to transform *E. coli* cells. From each library, plasmid DNA was isolated from 102 randomly picked colonies and sequenced to determine whether all 19 possible single mutants per residue position were present. Missing single mutants were separately constructed to ensure that the collection of mutants covered all possible 3439 single mutants of BSLA. Subsequently, the residual activity of each mutant was assessed after incubation with varying concentrations of four detergents with different physicochemical

properties (i.e. cationic, anionic, zwitterionic and non-ionic). The enzymatic activity of the BSLA mutants was measured using the screening substrate *p*-nitrophenyl butyrate (**1**), which after enzymatic hydrolysis yields *p*-nitrophenol (**2**) which can be detected using UV spectroscopy (Scheme 1a). By plotting the differences in residual activity of each mutant relative to WT BSLA, the authors could identify residue positions at which mutations led to increased tolerance or increased sensitivity towards detergents. By comparing this data to the crystallographic B-factors of BSLA, the authors observed that only two of the five regions in BSLA with high B-factors contained SDS tolerant variants, suggesting that B-factors are not a good predictor for hotspot positions which can be targeted to enhance detergent stability. Additionally, the authors observed that 84% of the hotspots for detergent tolerance were located on surface-exposed sites and that mainly substitutions to aromatic or charged residues, along with cysteine, improved detergent tolerance. This prompted the authors to suggest an optimized mutagenesis strategy using degenerate codons to introduce only those amino acids at solvent exposed sides, for efficiently improving the stability of other (BSLA)  $\alpha/\beta$ -hydrolase fold enzymes.



**Scheme 1.** Screening reactions used to generate mutability landscapes for enzymatic activity. a, The BSLA catalyzed hydrolysis of *p*-nitrophenyl butyrate (**1**) yielding *p*-nitrophenol (**2**) and butyric acid (**3**). b, The 4-OT catalyzed Michael-type addition of butanal (**4**) to *trans*- $\beta$ -nitrostyrene (**5**), yielding chiral  $\gamma$ -nitroaldehyde (**6**). c, The Bgl3 glycoside-bond cleavage of fluorescein di-( $\beta$ -D-glucopyranoside) (**7**), yielding fluorescein (**8**) and  $\beta$ -D-glucopyranoside (**9**).

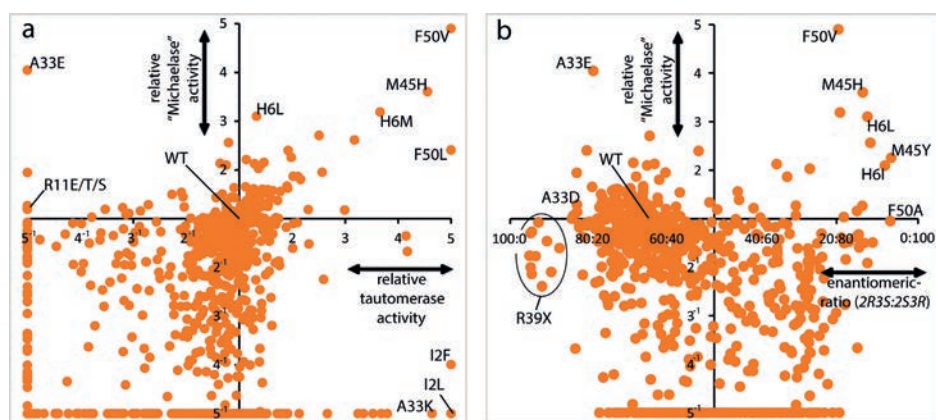


### 5.3. New catalytic functions and enantioselectivity

Poelarends and co-workers recently reported the use of mutability landscapes of the promiscuous enzyme 4-oxalocrotonate tautomerase (4-OT) to guide the engineering of novel biocatalysts for Michael-type additions.<sup>[31]</sup> The enzyme 4-OT is extremely promiscuous and its small monomer size of only 62 residues makes it an ideal template for mutability-landscape guided enzyme engineering.<sup>[32]</sup> One of 4-OT's promiscuous activities is the Michael-type addition of unmodified aldehydes to nitroalkenes yielding chiral  $\gamma$ -nitroaldehydes, which are valuable precursors for  $\gamma$ -aminobutyric acid (GABA)-based pharmaceuticals.<sup>[33-35]</sup> To generate the mutability landscapes, a defined collection of 4-OT genes was constructed which encoded at least 15 of the 19 possible variants at each residue position. Each member of this collection was individually characterized for the level of soluble protein expression, tautomerase and 'Michaelase' activities, and enantioselectivity. The level of soluble protein expression was determined for each mutant by using quantitative densitometry on SDS gels. After the 4-OT concentrations in the cell free extracts were quantified, the cell-free extracts were used in the activity and enantioselectivity assessments. All the activities were related to the amount of soluble 4-OT enzyme, yielding the specific activities of each mutant. An overview of the effect of each single mutant on both the tautomerase and 'Michaelase' activity (Fig. 2a) provides insight in the number of neutral amino acid substitutions, essential residues for one or both activities, and beneficial mutations. The positions where mutations lead to improved 'Michaelase' activity (His-6, Ala-33, Met-45 and Phe-50) were simultaneously varied in a focused library, which covered only those amino acid substitutions at each position that improved activity. This led to the identification of a triple mutant (H6M/A33E/F50V), which had an ~15-fold improvement in 'Michaelase' activity.

To screen for enantioselectivity, the authors assayed the enzymatic Michael-type addition of butanal (**4**) to *trans*- $\beta$ -nitrostyrene (**5**) (Scheme 1b). After following the progress of the reaction with UV-spectroscopy, the reaction mixtures were cleared by ultrafiltration and directly injected on a RP-HPLC system with a chiral stationary phase. Each single mutant was individually analyzed in this way, which allowed for the determination of both the 'Michaelase' activity and the enantiomeric ratio of the enzymatically produced 2-ethyl-3-phenyl-4-nitrobutanal (**6**) (Scheme 1b). When the activity data is plotted versus the enantioselectivity data (Fig. 2b) it becomes apparent that single amino acid substitutions can have significant effects on improving, inverting or losing the enantioselectivity. In the case of 4-OT, an inversion in enantioselectivity was required to produce precursors for the biologically most active enantiomer of the GABA-analogues. Therefore, the authors made combinations of the single mutants which had the most pronounced inversion in enantioselectivity (H6I, M45Y and F50A) leading to the identification of 4-OT

M45Y/F50A which produced the 2S3R-enantiomer of 2-ethyl-3-phenyl-4-nitrobutanal (**6**) with a e.r. of 96:4. This double mutant also had inverted enantioselectivity in the acetaldehyde addition to various nitroalkenes compared to WT 4-OT, producing the pharmaceutically relevant enantiomers of GABA precursors in enantiomeric ratios up to 97:3. The Michaelase activity of M45Y/F50A was also improved relative to WT 4-OT, which was not surprising because the mutability landscape already indicated that single mutations at these positions led to improved activity (Fig. 2b). Structural analysis of the M45Y/F50A mutant revealed the opening of a hydrophobic pocket in the active site of 4-OT which could accommodate the phenyl group of *trans*- $\beta$ -nitrostyrene (**5**). It seems likely that this new binding pocket is related to the inverted enantioselectivity of M45Y/F50A. The simultaneous improvement in activity and enantioselectivity underlines the usefulness of mutability landscapes in enzyme engineering.



**Figure 2.** Mutability landscape data derived from van der Meer *et al.*<sup>[31]</sup> **A**), mutational effects on 4-OT's tautomerase activity, plotted versus the mutational effects on 4-OT's promiscuous Michael-type addition reaction. **B**), mutational effects on 4-OT's enantioselectivity in the Michael-type addition reaction, plotted versus the mutational effects on 4-OT's activity in the Michael-type addition reaction.

## 6. Generating mutability landscapes using deep mutational scanning

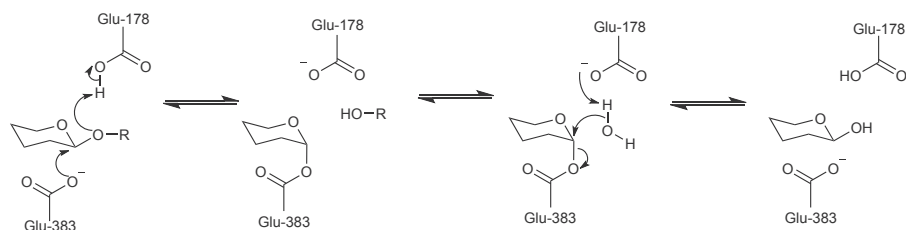
As mentioned above, it requires significant effort and resources to generate a defined gene collection encoding all single mutants of an enzyme. This bottleneck can be circumvented by using deep mutational scanning. For this, diversity in the WT enzyme is created followed by high-throughput sorting of active mutants from inactive mutants

(e.g. by flow cytometry, microfluidics, phage display or growth selection). This allows for the enrichment of active mutants. Conducting next-generation sequencing enables the comparison of the DNA read counts in the sorted library relative to the unsorted (or pre-selected) library (Fig. 1).<sup>[36,37]</sup> Using this approach, the enrichment factor (E-factor, which is given by the ratio of the DNA read counts of a specific variant in the sorted to the unsorted library) of each mutant can be determined and compared to the E-factor of the WT enzyme. A mutability landscape can be generated based on these E-factors, which maps the beneficial, neutral and detrimental effects of (nearly) all single amino acid substitutions of an enzyme. However, to obtain full coverage a large oversampling is required, demanding high throughputs for both the functional sorting and sequencing. Several examples of mutability landscapes using deep mutational scanning to investigate protein-DNA or protein-protein interactions can be found in the literature.<sup>[25-27,36,37]</sup> Recently, the first studies on the generation of mutability landscapes of enzymes by using deep mutational scanning have been published, which are discussed below.

### 6.1. Mutability landscape generation using microfluidics

$\beta$ -glucosidases are enzymes which cleave  $\beta$ -D-glucosidic bonds by hydrolysis, which can be an important step in the conversion of biomass into fermentable sugars.<sup>[38]</sup> Romero *et al.* have generated mutability landscapes of a  $\beta$ -glucosidase from *Streptomyces* sp. (Bgl3) using a deep mutational scanning approach in combination with a micro-fluidic based sorting system.<sup>[39]</sup> For this, they generated a random mutant library of Bgl3 using error-prone PCR with an average of 3.8 mutations per Bgl3 gene. This library was first analyzed using high-throughput sequencing to establish the DNA read counts in the unsorted library. After expressing this library in *E. coli* BL21(DE3), single *E. coli* cells were encapsulated in a micro droplet containing lysing agents and fluorescein di-( $\beta$ -D-glucopyranoside) (7), which is a fluorogenic substrate for Bgl3 (Scheme 1c). Micro droplets containing an active Bgl3 variant were sorted based on fluorescence, using a microfluidic device. This way the authors achieved a throughput of 100 s<sup>-1</sup>. DNA was retrieved from the sorted micro droplets and sequenced using Illumina sequencing. After analyzing 10<sup>7</sup> variants, the effects of the mutations were determined based on changes in the frequency of occurrence of each mutation before and after the functional sorting. Because of the disadvantage of working with an error-prone library, mainly those amino-acid substitutions which require one nucleotide mutation per codon were accessed in this study. Therefore, only 31% of all possible single amino acid substitutions were analyzed. Nevertheless, the generated mutability landscape gave important insights in sequence-function relationships of the enzyme. For example, two essential residues (Lys-461 and Asn-307), which are located outside of the enzyme's active site, were identified in this study. Crystal structure analysis of Bgl3 revealed that Lys-461 is part of a network of salt

bridges, which suggests that this residue plays a role in structural stability of the enzyme. Asn-307 is in hydrogen bonding distance with Glu-178, which is the catalytic acid/base in Bgl3 (Scheme 2). It was therefore suggested that Asn-307 induced a crucial shift in the  $pK_a$  of this catalytic residue. Single mutations that improve the thermostability of Bgl3 have been identified in a slightly modified microfluidic screening protocol, which included a heat challenge (65°C for 10 min). Again,  $10^7$  enzyme variants were analyzed, revealing several single mutants with improved thermostability including mutant S325C. Further characterization of this mutant revealed a 5.3°C increase in  $T_{50}$  relative to WT Bgl3.



**Scheme 2.** General mechanism of glucosidic-bond cleavage by Bgl3. Mechanism is derived from Zechel *et al.*<sup>[40]</sup>

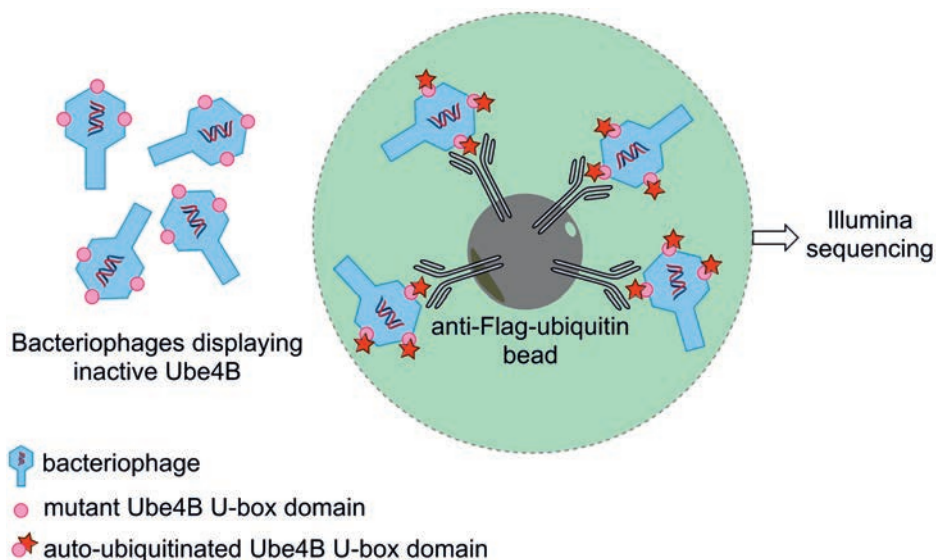
## 6.2. Mutability landscape generation using growth selection

Aminoglycoside-3'-phosphotransferase II (APH(3')II) is a kinase involved in antibiotic resistance that catalyzes the phosphorylation of aminoglycoside antibiotics leading to their inactivation. Melnikov *et al.* performed a single-substitution mutational scan on APH(3')II by analyzing the effect of these mutations on the enzyme's activity and substrate specificity, using kanamycin and five other aminoglycoside antibiotics.<sup>[41]</sup> For this, the genes coding for each single mutant were individually prepared using a microarray-based DNA synthesis (MITE) approach. All synthesized genes were pooled in equimolar amounts and used to transform *E. coli* cells. These cells were cultured in liquid medium in the presence of aminoglycoside antibiotics, thereby selecting for cells which express an active APH(3')II mutant. After this selection, DNA was isolated from the surviving cells and sequenced using an Illumina sequencing approach to determine the frequency of occurrence of each mutant. By determining the difference in abundance of each mutant before and after selection, the authors could map the effect of all single amino acid substitutions on activity on six aminoglycoside antibiotics. Based on these maps, amino acid substitutions were identified which led to a shift in substrate specificity either towards kanamycin or towards one of the five other tested aminoglycoside antibiotics. By making combinations of these specific amino acid substitutions, the authors engineered five pairs of APH(3')IIs which either favor or disfavor any of the tested antibiotics over kanamycin.

For example, Paro+ and Paro- is a pair of APH(3')IIs, engineered to either favor or disfavor paromomycin over kanamycin. Paro+ showed unaltered activity for paromomycin (MIC= 2000-4000 µg/ml) relative to WT APH(3')II but a decreased activity for kanamycin (MIC 31.3 µg/ml). Paro- had a decreased activity for paromomycin (MIC= 62.5 µg/ml) relative to WT APH(3')II but unaltered activity towards kanamycin (MIC= 2000 µg/ml) relative to WT APH(3')II. This remarkable shift in substrate specificity underlines the applicability of mutability landscapes to identify hotspots for enzyme engineering.

### 6.3. Mutability landscape generation using phage display

E3-ubiquitin ligases are enzymes which catalyze an ubiquitin transfer from E2-ubiquitin conjugating enzymes to lysines of substrate proteins. This ubiquitination promotes degradation of the substrate protein, which is a crucial process for homeostasis. Ube4b, for example, functions as an E3-ubiquitin ligase, which has been linked to cancer pathogenesis as it ubiquitinates the p53 tumor suppressor *in vivo*.<sup>[42]</sup> A mutability landscape for the activity of the Ube4b enzyme has been generated and analyzed in order to identify the molecular determinants that modulate the ligase activity of these E3 ligases.<sup>[43]</sup> A deep-mutational scanning approach was conducted on the U-box domain of Ube4b. This is the active domain of the enzyme which can perform an auto-ubiquitination. Libraries with on average two random nucleotide mutations per gene were generated, sequenced and subsequently displayed on bacteriophages. Bacteriophages which display active (auto-ubiquitinated) U-box domains were then enriched using antibodies against (FLAG)-ubiquitin. Because these antibodies were immobilized on agarose beads, unbound bacteriophages could be washed away (Fig. 3). DNA was isolated from enriched bacteriophages and subsequently sequenced using Illumina technology. By comparing the DNA-read counts of each mutation before and after the enrichment, an E-factor was calculated. In this way, 98,289 unique mutant enzymes were characterized, of which 932 single mutants. Mapping the E-factors of these single mutants revealed that some regions (*e.g.* loop 1,2 and helix 1) were less tolerant to mutations than other portions of the U-box domain. Interestingly, several single mutants could be identified from this mutability landscape with improved activity relative to WT. Combining these beneficial single mutations had a synergistic effect and resulted in two double mutants (M1124V/N1142T and D1139N/N1142T) which each have a 22-fold enhanced ubiquitin ligase activity relative to the WT U-box domain. Mechanistic studies on these improved single and double mutants revealed that all beneficial mutations either enhanced the ligase activity by improving binding of the U-box domain to the E2-ubiquitin complex or by improving allosteric activation of the E2-ubiquitin complex. This illustrates that beneficial mutations can be useful both for the generation of superior enzymes and to provide useful insight in enzyme mechanisms.



**Figure 3.** Enrichment procedure for bacteriophages displaying active (auto-ubiquitinated) U-box domains of Ube4b. Bacteriophages displaying inactive Ube4B U-box domains do not bind to anti-Flag-ubiquitin beads and are washed away. Only the bacteriophages which display active auto-ubiquitinated Ube4B U-box domains bind to the anti-Flag-ubiquitin beads and are sequenced.<sup>[43]</sup>

## 7. Summary and Outlook

Currently, most studies on enzyme mutability landscapes have focused on small enzymes (Table 1), which is related to the required costs and effort to generate a mutability landscape. When using a defined collection of single mutants, the bottleneck lays in the generation of this defined mutant gene collection. Currently, PCR-based site-directed mutagenesis techniques are mostly used for the generation of the mutants. Other more recently developed mutagenesis techniques include chemo-enzymatic methods (e.g. SeSaM),<sup>[44]</sup> micro-array based DNA synthesis (e.g. MIRE)<sup>[41]</sup> or nonsense suppressor t-RNA methods.<sup>[25]</sup> The development of these methods might reduce the required amount of effort and costs to generate a defined collection of single mutants. Moreover, because of the ever decreasing costs of commercially available synthetic DNA, the most economical way to obtain a defined collection of single mutants of an enzyme might be DNA synthesis.<sup>[45]</sup> In the case of deep-mutational scanning the bottleneck for generating mutability landscapes lays in the high-throughput sequencing and high-throughput screening. Both these techniques are rapidly evolving<sup>[36,37,46]</sup> which might facilitate the generation of mutability landscapes using deep mutational scanning.

**Table 1.** Available studies on experimental mutability landscape analyses of enzymes.

Type of enzyme	Defined mutant collection	Deep mutational scanning	Investigated enzymatic property <sup>a</sup>	Used for hotspot identification <sup>b</sup>	Size of enzyme	Ref.
Protease	X		A, S, SS	X	189	[28]
Lipase	X		A, S		181	[30]
Tautomerase / 'Michaelase'	X		A, E, ES, SS	X	62	[31]
Glucosidase		X	A, S		500	[39]
Kinase		X	A, SS	X	263	[41]
Ligase		X	A	X	102 <sup>c</sup>	[43]

<sup>a</sup> S, stability; A, activity; E, expression; ES, enantioselectivity; SS, substrate specificity. <sup>b</sup> The box is checked when combinatorial mutagenesis was conducted on hotspots which were identified in the mutability landscape. <sup>c</sup> Only the U-box domain of Ube4b was analyzed.

In conclusion, mutability landscapes are a powerful tool to identify “hotspots” at any place in the amino acid sequence of an enzyme. These “hotspots” can be used as targets for combinatorial mutagenesis to yield superior enzymes with improved catalytic properties, stability or even new enzymatic activities. The generation of mutability landscapes for several properties of one enzyme (for example, stability and activity or activity and enantioselectivity) provides the unique opportunity to select mutations, which are beneficial for either one or both these properties. Furthermore, mutability landscapes can be used to advance our understanding of sequence-function relationships in enzymes since they provide systematic information on neutral, beneficial and detrimental amino acid substitutions. Both detrimental and beneficial mutations can be extremely helpful to elucidate enzyme mechanisms. Neutral mutations are thought to have an important role in natural enzyme evolution, because they may result in ‘neutral drift’.<sup>[47,48]</sup> Owing to these advantages, combined with the technical advances in high-throughput screening and DNA sequencing, we expect that mutability landscape analysis will become accessible for larger enzymes, and more commonly used for enzyme engineering in the coming years.

### Acknowledgement

The authors acknowledge funding from the Division of Earth and Life Sciences of the Netherlands Organisation of Scientific Research (ALW grant 820.02.021), the European Research Council under the European Community’s Seventh Framework Programme (FP7/2007-2013)/ERC Grant agreement n° 242293, and the European Union’s Horizon 2020 research and innovation programme under grant agreement No 635595.

## References

1. U. T. Bornscheuer, G. W. Huisman, R. J. Kazlauskas, S. Lutz, J. C. Moore, K. Robins, *Nature* **2012**, *485*, 185-194.
2. B. M. Nestl, S. C. Hammer, B. A. Nebel, B. Hauer, *Angew. Chem. Int. Ed. Engl.* **2014**, *53*, 3070-3095.
3. M. T. Reetz, *Angew. Chem. Int. Ed.* **2011**, *50*, 138-174.
4. A. S. Bommarius, M. F. Paye, *Chem. Soc. Rev.* **2013**, *42*, 6534-6565.
5. A. A. Homaei, R. Sariri, F. Vianello, R. Stevanato, *J. Chem. Biol.* **2013**, *6*, 185-205.
6. A. Illanes, A. Cauerhff, L. Wilson, G. R. Castro, *Bioresour. Technol.* **2012**, *115*, 48-57.
7. R. A. Sheldon, S. van Pelt, *Chem. Soc. Rev.* **2013**, *42*, 6223-6235.
8. G. Kiss, N. Çelebi-Ölçüm, R. Moretti, D. Baker, K. N. Houk, *Angew. Chem. Int. Ed. Engl.* **2013**, *52*, 5700-5725.
9. H. Kries, R. Blomberg, D. Hilvert, *Curr. Opin. Chem. Biol.* **2013**, *17*, 221-228.
10. A. Zanghellini, *Curr. Opin. Biotechnol.* **2014**, *29*, 132-138.
11. O. Khersonsky, D. S. Tawfik, *Annu. Rev. Biochem.* **2010**, *79*, 471-505.
12. R. A. Jensen, *Ann. Rev. Microbiol.* **1976**, *30*, 409-425.
13. P. J. O'Brien, D. Herschlag, *Chem. Biol.* **1999**, *6*, R91-R105.
14. H. Renata, Z. J. Wang, F. H. Arnold, *Angew. Chem. Int. Ed. Engl.* **2015**, *54*, 3351-3367.
15. N. J. Turner, *Nat.Chem.Biol.* **2009**, *5*, 567-573.
16. E. Sebestova, J. Bendl, J. Brezovsky, J. Damborsky in *Methods in Molecular Biology - Directed Evolution Library Creation* (Eds.: E. M. J. Gillam, J. N. Copp, D. Ackerley), Springer, New York, **2014**, pp. 291-314.
17. H. J. Wijma, R. J. Floor, D. B. Janssen, *Curr. Opin. Struct. Biol.* **2013**, *23*, 588-594.
18. M. T. Reetz, L. Wang, M. Bocola, *Angew. Chem. Int. Ed. Engl.* **2006**, *45*, 1236-1241.
19. M. T. Reetz, D. Kahakeaw, J. Sanchis, *Molecular BioSystems* **2009**, *5*, 115-122.
20. R. A. Chica, N. Doucet, J. N. Pelletier, *Curr. Opin. Biotechnol.* **2005**, *16*, 378-384.
21. R. J. Fox, S. C. Davis, E. C. Mundorff, L. M. Newman, V. Gavrilovic, S. K. Ma, L. M. Chung, C. Ching, S. Tam, S. Muley, J. Grate, J. Gruber, J. C. Whitman, R. A. Sheldon, G. W. Huisman, *Nat Biotechnol.* **2007**, *25*, 338-344.
22. M. Hecht, Y. Bromberg, B. Rost, *J. Mol. Biol.* **2013**, *425*, 3937-3948.
23. M. Hecht, Y. Bromberg, B. Rost, *BMC Genomics* **2015**, *16 Suppl 8*, S1.
24. N. Sim, P. Kumar, J. Hu, S. Henikoff, G. Schneider, P. C. Ng, *Nucleic Acids Res.* **2012**, *40*, W452-W457.



## Chapter 1

25. P. Markiewicz, L. G. Kleina, C. Cruz, S. Ehret, J. H. Miller, *J. Mol. Biol.* **1994**, *240*, 421-433.
26. T. A. Whitehead, A. Chevalier, Y. Song, C. Dreyfus, S. J. Fleishman, C. De Mattos, C. A. Myers, H. Kamisetty, P. Blair, I. A. Wilson, D. Baker, *Nat. Biotechnol.* **2012**, *30*, 543-548.
27. R. N. McLaughlin, Jr., F. J. Poelwijk, A. Raman, W. S. Gosal, R. Ranganathan, *Nature* **2012**, *491*, 138-142.
28. D. A. Estell, W. Aehle, **2011**, U.S. patent 12/778,915.
29. H. Jochens, M. Hesseler, K. Stiba, S. K. Padhi, R. J. Kazlauskas, U. T. Bornscheuer, *ChemBioChem.* **2011**, *12*, 1508-1517.
30. A. Fulton, V. J. Frauenkron-Machedjou, P. Skoczinski, S. Wilhelm, L. Zhu, U. Schwaneberg, K. E. Jaeger, *ChemBioChem.* **2015**, *16*, 930-936.
31. J. Y. van der Meer, H. Poddar, B. J. Baas, Y. Miao, M. Rahimi, A. Kunzendorf, R. van Merkerk, P. G. Tepper, E. M. Geertsema, A. M. Thunnissen, W. J. Quax, G. J. Poelarends, *Nat. Commun.* **2016**, *7*:10911 doi: 10.1038/ncomms10911.
32. B. J. Baas, E. Zandvoort, E. M. Geertsema, G. J. Poelarends, *ChemBioChem.* **2013**, *14*, 917-926.
33. E. M. Geertsema, Y. Miao, P. G. Tepper, P. de Haan, E. Zandvoort, G. J. Poelarends, *Chem. Eur. J.* **2013**, *19*, 14407-14410.
34. Y. Miao, E. M. Geertsema, P. G. Tepper, E. Zandvoort, G. J. Poelarends, *ChemBioChem.* **2013**, *14*, 191-194.
35. E. Zandvoort, E. M. Geertsema, J. B. Baas, W. J. Quax, G. J. Poelarends, *Angew. Chem. Int. Ed. Engl.* **2012**, *51*, 1240-1243.
36. D. M. Fowler, S. Fields, *Nat. Meth.* **2014**, *11*, 801-807.
37. H. Shin, B. Cho, *Int. J. Mol. Sci.* **2015**, *16*, 23094-23110.
38. J. A. Perez-Pons, A. Cayetano, X. Rebordosa, J. Lloberas, A. Guasch, E. Querol, *FEBS J.* **1994**, *223*, 557-565.
39. P. A. Romero, T. M. Tran, A. R. Abate, *Proc. Natl. Acad. Sci.* **2015**, *112*, 7159-7164.
40. D. L. Zechel, S. G. Withers, *Acc. Chem. Res.* **2000**, *33*, 11-18.
41. A. Melnikov, P. Rogov, L. Wang, A. Gnirke, T. S. Mikkelsen, *Nucleic Acids Res.* **2014**, *42*, e112.
42. H. Wu, S. L. Pomeroy, M. Ferreira, N. Teider, J. Mariani, K. I. Nakayama, S. Hatakeyama, V. A. Tron, L. F. Saltibus, L. Spyropoulos, R. P. Leng, *Nat. Med.* **2011**, *17*, 347-355.
43. L. M. Starita, J. N. Pruneda, R. S. Lo, D. M. Fowler, H. J. Kim, J. B. Hiatt, J. Shendure, P. S. Brzovic, S. Fields, R. E. Klevit, *Proc. Natl. Acad. Sci.* **2013**, *110*, E1263-E1272.
44. A. J. Ruff, T. Kardashliev, A. Dennig, U. Schwaneberg in *Methods in Molecular Biology - Directed Evolution Library Creation* (Eds.: E. M. J. Gillam, J. N. Copp, D. Ackerley), Springer, New York, **2014**, pp. 45-68.
45. R. Carlson, *Nat. Biotechnol.* **2009**, *27*, 1091-1094.

46. M. Wójcik, A. Telzerow, W. J. Quax, Y. L. Boersma, *Int. J. Mol. Sci.* **2015**, *16*, 24918-24945.
47. J. D. Bloom, P. A. Romero, Z. Lu, F. H. Arnold, *Biol. Direct* **2007**, *2*, 17.
48. M. Soskine, D. S. Tawfik, *Nat. Rev. Genet.* **2010**, *11*, 572-582.



# Chapter 2

## *Enantioselective Synthesis of Pharmaceutically Active $\gamma$ -Aminobutyric Acids Using a Tailor-Made Artificial Michaelase in One-Pot Cascade Reactions*

---

Lieuwe Biewenga<sup>†a</sup>, Thangavelu Saravanan<sup>†a</sup>, Andreas Kunzendorf<sup>a</sup>,  
Jan-Ytzen van der Meer<sup>a</sup>, Tjaard Pijning<sup>b</sup>, Pieter G. Tepper<sup>a</sup>, Ronald van Merkerk<sup>a</sup>,  
Simon J. Charnock<sup>c</sup>, Andy-Mark W. H. Thunnissen<sup>d</sup>, and Gerrit J. Poelarends<sup>\*a</sup>

<sup>a</sup>*Department of Chemical and Pharmaceutical Biology, Groningen Research Institute of Pharmacy,  
University of Groningen, Antonius Deusinglaan 1, 9713 AV Groningen, The Netherlands.*

<sup>b</sup>*Structural Biology Group, Groningen Institute of Biomolecular Sciences and Biotechnology,  
University of Groningen, Nijenborgh 7, 9747 AG Groningen, The Netherlands.*

<sup>c</sup>*Prozomix Ltd., Station Court, Haltwhistle, Northumberland NE49 9HN, U.K.*

<sup>d</sup>*Molecular Enzymology Group, Groningen Institute of Biomolecular Sciences and Biotechnology,  
University of Groningen, Nijenborgh 7, 9747 AG Groningen, The Netherlands.*

<sup>†</sup>*These authors contributed equally: Lieuwe Biewenga and Thangavelu Saravanan.*

<sup>\*</sup>*Corresponding author. Tel.: +31503633354; E-mail: g.j.poelarends@rug.nl;*

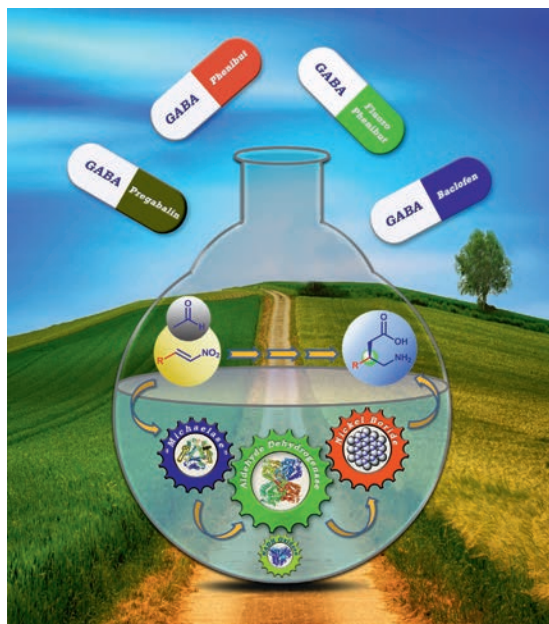
*Web: <http://www.rug.nl/staff/g.j.poelarends/>*

## Abstract

Chiral  $\gamma$ -aminobutyric acid (GABA) analogues represent abundantly prescribed drugs, which are broadly applied as anticonvulsants, antidepressants and for the treatment of neuropathic pain. Here we report a one-pot two-step biocatalytic cascade route for synthesis of the pharmaceutically relevant enantiomers of  $\gamma$ -nitrobutyric acids, starting from simple precursors (acetaldehyde and nitroalkenes), using a tailor-made highly enantioselective artificial 'Michaelase' (4-oxalocrotonate tautomerase mutant L8Y/M45Y/F50A), an aldehyde dehydrogenase with a broad non-natural substrate scope, and a cofactor recycling system. We also report a three-step chemoenzymatic cascade route for the efficient chemical reduction of enzymatically prepared  $\gamma$ -nitrobutyric acids into GABA analogues in one pot, achieving high enantiopurity (e.r. up to 99:1) and high overall yields (up to 70%). This chemoenzymatic methodology offers a step-economic alternative route to important pharmaceutically active GABA analogues, and highlights the exciting opportunities available for combining chemocatalysts, natural enzymes, and designed artificial biocatalysts in multistep syntheses.

### Keywords:

Systems biocatalysis, cascades, 'Michaelase',  $\gamma$ -aminobutyric acids,  $\gamma$ -nitrobutyric acids, enzyme engineering, pharmaceuticals



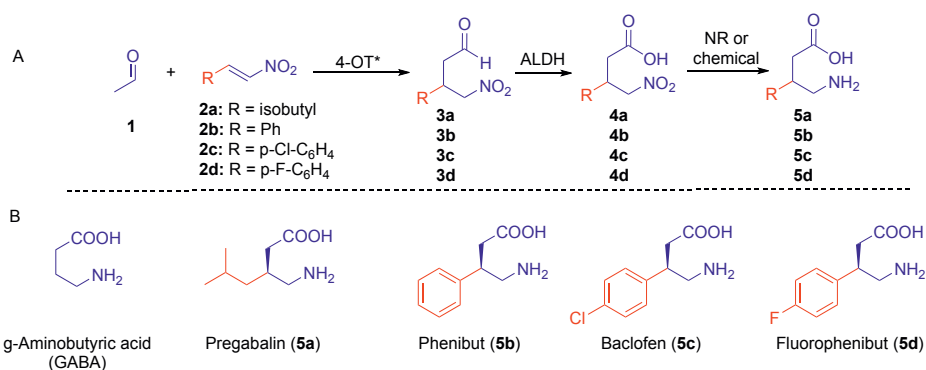
## Introduction

Analogues of  $\gamma$ -aminobutyric acid (GABA, Figure 1) represent abundantly prescribed drugs, which are broadly applied as anticonvulsants, antidepressants and for the treatment of neuropathic pain. With an increasing world population and life expectancy, the demand for GABA analogues is expected to even further increase. The efficient asymmetric synthesis of pharmaceutically active GABA analogues has therefore attracted enormous attention. With current synthesis routes often involving kinetic resolutions<sup>1-3</sup>, there is a need to investigate alternative asymmetric synthesis routes that are potentially greener, more sustainable, and more step-economic. In this regard, the use of a systems (bio)catalysis approach<sup>4-8</sup> in which different catalysts are combined to construct reaction cascades for efficient synthesis of GABA analogues is an attractive idea. This approach aims to minimize the number of reaction steps and improve the 'pot-economy' of the process<sup>9</sup>.

We envisioned that pharmaceutically active GABA analogues, such as pregabalin, phenibut, baclofen and fluorophenibut, could be prepared via one-pot three-step (chemo)enzymatic cascade reactions, using simple and inexpensive starting materials and avoiding (de-)protecting steps and intermediate purifications (Figure 1). For establishing the required C-C bond stereochemistry, the asymmetric Michael-type addition of acetaldehyde (**1**) to nitroalkenes **2a-d** is of high interest. This would give convenient access to chiral  $\gamma$ -nitroaldehydes **3a-d**, which in two steps (oxidation of **3a-d** into **4a-d** followed by reduction of **4a-d** into **5a-d**) can be converted into the desired GABA analogues.

The asymmetric Michael-type addition of **1** to **2a-d** is certainly not trivial. Multiple organocatalytic approaches to obtain enantioenriched  $\gamma$ -nitroaldehydes have been reported, mainly using small peptide- and proline-based organocatalysts<sup>10-13</sup>. However, examples including acetaldehyde as donor substrate are scarce and a high catalyst loading in organic solvent is typically applied. Therefore, there is great interest in the development of biocatalytic procedures for the enantioselective synthesis of  $\gamma$ -nitroaldehydes. However, enzymes that naturally catalyze these required carbon-carbon bond-forming Michael-type additions are not known to be present in nature. In fact, only a few enzymes are known to be able to catalyze any type of C-C bond-forming Michael-type addition<sup>14,15</sup>. Interestingly, Hilvert and co-workers published the elegant enzymatic synthesis of  $\gamma$ -nitroketones, but not  $\gamma$ -nitroaldehydes, by both acetone addition to nitroalkenes and nitroalkane addition to conjugated ketones using a highly engineered computationally designed retroaldolase<sup>16</sup>. We have previously reported

that 4-oxalocrotonate tautomerase (4-OT) can promiscuously catalyze the addition of small aldehydes, most notably the highly reactive acetaldehyde, to various aliphatic and aromatic nitroalkenes<sup>17-20</sup>. Analogous to proline-based organocatalysts, the 4-OT enzyme utilizes an N-terminal proline (Pro-1) as key catalytic residue in promiscuous aldol condensations<sup>21,22</sup> and Michael-type additions<sup>19</sup>, most likely via enamine catalysis<sup>22,23</sup>. By using mutability-landscape guided enzyme engineering<sup>24</sup>, a mutant of 4-OT (4-OT M45Y/F50A) was generated that showed inverted enantioselectivity in acetaldehyde additions to nitroalkenes<sup>25</sup>, allowing the enzymatic synthesis of the pharmaceutically relevant enantiomers of  $\gamma$ -nitroaldehydes **3a-d**. However, the enantioselectivity of 4-OT M45Y/F50A is too low for biocatalytic application, providing the desired  $\gamma$ -nitroaldehyde products with only modest enantiomeric excess<sup>25</sup>.



**Figure 1. Chemoenzymatic cascade synthesis of pharmaceutically active GABA analogues.** A) Envisioned (chemo)enzymatic cascade synthesis of pharmaceutically active GABA analogues. Abbreviations: 4-OT\*, newly engineered 4-OT variant that functions as a highly enantioselective artificial ‘Michaelase’; ALDH, aldehyde dehydrogenase; NR, nitroreductase. B) Structures of GABA and its analogues pregabalin, phenibut, baclofen and fluorophenibut.

Here, we report the development of a tailor-made artificial ‘Michaelase’, which exhibits improved enantioselectivity, activity and cosolvent stability compared to the parental enzyme 4-OT M45Y/F50A, for additions of **1** to nitroalkenes **2a-d** yielding  $\gamma$ -nitroaldehydes **3a-d** with outstanding enantiopurity. This artificial ‘Michaelase’ was combined with a natural aldehyde dehydrogenase and a cofactor recycling NADH-oxidase to give a one-pot two-step reaction cascade for the synthesis of  $\gamma$ -nitrobutyric acids **4a-d** in high yields and with excellent enantiomeric excess. Finally, the reaction cascade was further extended by the inclusion of nickel boride, promoting the conversion of enzymatically prepared **4a-d** into the desired GABA analogues **5a-d**, which resulted in a one-pot three-step chemoenzymatic reaction cascade (Figure 1A). Given that all steps

were performed under aqueous conditions with high conversions, the desired GABA analogues **5a-d** were obtained in good isolated yields of up to 70% and with excellent enantiomer ratio (e.r.) values of up to 99:1. This new methodology offers a step-economic alternative route to important pharmaceutically active GABA analogues, and highlights the exciting opportunities available for combining chemocatalysts, natural enzymes, and designed artificial biocatalysts in multistep syntheses of valuable chemical products.

## Experimental Methods

### Library Screening

After transformation with 4-OT I2X/L8X/M45Y/F50A library DNA, individual *E. coli* BL21(DE3) colonies were used to inoculate 2x 1.25 ml LB supplemented with 100  $\mu$ g/ml ampicillin and 100  $\mu$ M isopropyl  $\beta$ -D-1-thiogalactopyranoside (IPTG) in 96-deep well plates (Greiner Bio-one, 96-well Masterblock). The plates were sealed with sterile gas-permeable seals (Greiner Bio-one, BREATHseal) and incubated at 37  $^{\circ}$ C, overnight shaking at 250 rpm. After the incubation, the plates were centrifuged (2232 x g, 8 min). The supernatant was discarded and the individual pellets were lysed by resuspension in 350  $\mu$ l BugBuster (Novagen) supplemented with 25 U/ml benzonase (Novagen). The lysis was continued for 20 min at room temperature under vigorous shaking. The lysates were cleared by centrifugation (2232 x g, 55 min, 4  $^{\circ}$ C) after which the Cell Free Extract (CFE) was obtained. The final reaction mixture for monitoring the addition of **1** to **2a** consisted of the following: CFE (40% v/v), 150 mM **1**, 5 mM **2a**, DMSO (5% v/v) in 20 mM sodium phosphate buffer pH 6.5, 500  $\mu$ l final volume. The reactions were performed in 96-deep well plates sealed by ultraviolet transparent plate seals (VIEWseal, Greiner Bio-One) at room temperature. After 50 min the reaction was stopped by extraction with 300  $\mu$ l toluene, which caused the proteins to precipitate. The organic layer was separated from the water layer by centrifugation (2232 x g, 20 min). The plates containing the water and organic layer were incubated at -80  $^{\circ}$ C for 30 min to freeze the water layer and hence preventing accidental uptake of part of the water layer. An aliquot of 50  $\mu$ l from the organic layer was transferred to a GC vial by a robotic pipetting station. For analysis of the amount and enantiopurity of the enzymatic product **3a**, 8  $\mu$ l of the organic layer was injected on a gas chromatograph using an Astec CHIRALDEX G-TA column, isocratic 125  $^{\circ}$ C (carrier gas He, r1.69 ml/min). Flame ionization detection; retention time *S*-**3a**: 25.6 min, retention time *R*-**3a** 26.9 min. The assignment of the absolute configuration of product **3a** was based on earlier reported data<sup>17,25</sup>.



### Semi-Preparative Scale Experiments

All semi-preparative scale experiments were based on previously reported reaction conditions<sup>17,25</sup>. Instead of DMSO, ethanol was used as a co-solvent, since the desired products **3a-d** were obtained at the highest enantiopurity using 4-OT L8Y/M45Y/F50A in combination with ethanol as the co-solvent. For the enzymatic synthesis of **3a**, the reaction mixture consisted of the following: 5.0 mg **2a** (3 mM), 6.5 mg 4-OT L8Y/M45Y/F50A (75  $\mu$ M, 2.5 mol% compared to **2a**), 2.56 ml ethanol (20% v/v) in 20 mM sodium phosphate buffer pH 6.5 to a final volume of 12.8 mL. The reaction was performed in a 25 ml Erlenmeyer flask sealed with a rubber stopper. The reaction was initiated by the addition of 108  $\mu$ l of **1** (150 mM) and incubated at room temperature. At timely intervals a sample was withdrawn from the reaction mixture and the reaction progress was monitored by following the depletion in absorbance at 249 nm corresponding to the concentration of substrate **2a**<sup>17</sup>. After the measurement, the sample was combined with the reaction mixture. After 50 min the reaction was finished and the reaction mixture was extracted 3x with 10 ml toluene. The organic layers were combined, washed with brine and dried with anhydrous Na<sub>2</sub>SO<sub>4</sub>. The dried organic layer was concentrated in vacuo, yielding product **3a** without any further purification (4.2 mg, 63% isolated yield). The enantiopurity and absolute configuration was determined by GC with a chiral stationary phase (see library screening).

For the enzymatic synthesis of **3b-d**, the reaction mixture consisted of the following: nitroalkene (17.9 mg **2b**, 2 mM; 14.3 mg **2c**, 1.3 mM; or 20.1 mg **2d**, 2 mM), 11.5 mg 4-OT L8Y/M45Y/F50A (28  $\mu$ M, 1.4 mol% compared to **2b** and **2d**, 2.15 mol% compared to **2c**), ethanol (12 ml, 20% (v/v) for **2b** and **2d** and 18 ml, 30% (v/v) for **2c**) in 20 mM sodium phosphate buffer pH 6.5 to a final volume of 60 ml. The reaction was performed in a 100 ml Erlenmeyer flask sealed with a rubber stopper. The reaction was initiated by the addition of 169  $\mu$ l **1** (50 mM) and incubated at room temperature. At timely intervals a sample was withdrawn from the reaction mixture and the reaction progress was monitored by following the depletion in absorbance ( $\lambda_{\text{max}}$  of **2b** and **2c** at 320 nm,  $\lambda_{\text{max}}$  of **2d** at 322 nm). After the measurement, the sample was combined with the reaction mixture. When the reaction was finished, the reaction mixture was extracted with 3x 40 ml ethyl acetate. The organic layers were combined, washed with brine and dried with anhydrous Na<sub>2</sub>SO<sub>4</sub>. The dried organic layer was concentrated in vacuo, yielding **3b-d** without any further purification (**3b** 20.1 mg, 87% yield; **3c** 17.2 mg, 97% yield; and **3d** 24.0 mg, 94% yield). The aldehyde functionality of **3b** and **3c** was derivatized to a cyclic acetal (see supporting methods). The enantiopurity of **3b-d** was determined by reverse phase HPLC using a chiralpak AD-RH column (150 mm x 4.6 mm, Daicel) (MeCN/water 70:30 for **3b**, MeCN/water 62:38 for **3c** and MeCN/water 30:70 for **3d**, 25 °C, 0.5

ml/min flow rate). Detection at 220 nm, retention time *R*-**3b**: 8.3 min, *S*-**3b** 10.8 min; *R*-**3c**: 29 min, *S*-**3c** 37 min; *R*-**3d**: 39 min and *S*-**3d** 41 min. The absolute configuration was determined by literature comparison<sup>17</sup>.

### Two-Step Enzymatic Cascade Synthesis

The reaction mixture contained **2a** (22 mg, 3 mM), 4-OT L8Y/M45Y/F50A (5 mol% compared to **2a**), 5 ml ethanol (10% v/v) in 0.1 M sodium phosphate buffer (pH 7.3) to a final volume of 50 mL. The reaction was initiated by the addition of 50 mM of **1** (2.5 mL from a freshly prepared 1 M stock solution) and the reaction mixture was incubated at room temperature. At timely intervals a sample was withdrawn from the reaction mixture and the reaction progress was monitored by following the depletion in absorbance at 249 nm corresponding to the concentration of substrate **2a**<sup>17</sup>. After the measurement the sample was combined with the reaction mixture. After 20 min, the reaction was finished (full conversion of **2a** into **3a**) and then to this PRO-ALDH(003) (0.5 mg/ml), PRO-NOX(009) (1 mg/ml) and NAD<sup>+</sup> (8 mM) were added. We used this step-wise cascade approach because PRO-ALDH(003) also showed activity towards starting substrate **1**. Favorably, the enzyme showed ~3-fold higher activity towards substrate **3a** (at 3 mM) than towards **1** (at 50 mM). Using this methodology, complete oxidation of **3a** into **4a** was achieved with the formation of less than 10 mM acetate as by-product from **1**. After the PRO-ALDH(003)-catalyzed oxidation reaction was completed, the reaction was stopped by acidifying the reaction mixture to pH 5. The reaction mixture was extracted with 3x 50 ml ethyl acetate. The organic layers were combined, washed with brine and dried with anhydrous Na<sub>2</sub>SO<sub>4</sub>. The dried organic layer was concentrated in vacuo and the product was purified by silica gel column chromatography (hexane/ethylacetate 4:1) to obtain **4a** (18 mg, 64% yield). The acid functionality of **4a** was derivatized to the corresponding methyl ester (see supporting methods). The same reaction procedures were followed for the synthesis of **4b-4d**. The enantiopurity of derivatized **4a** was determined by GC using a chiral Astec CHIRALDEX G-TA column, 100 °C to 140 °C at a rate of 4 °C/min, then 140 °C to 160 °C at a rate of 2 °C/min and finally held at 160 °C. Flame ionization detection, retention time *R*-**4a**: 16 min; *S*-**3c**: 17 min. The enantiopurity of derivatized **4b-d** was analyzed by reverse phase HPLC using a chiral column (Chiralpak-ID, 150 mm x 4.6 mm, Daicel®) (MeCN/water 30:70, 25 °C, 1 ml/min flow rate). Detection at 210 nm, retention time *R*-**4b**: 22.6 min, *S*-**4b**: 24.5 min; *R*-**4c**: 19.4 min, *S*-**4c**: 24.4 min; *R*-**4d**: 23.5 min, *S*-**4d**: 26.6 min.

### Three-Step Chemoenzymatic Cascade Synthesis

The reaction mixture contained nitoalkene (3 mM of **2a** and 4 mM of **2b-d**), 4-OT L8Y/M45Y/F50A (28-32 mg; 5 mol% compared to **2a-2d**), ethanol (10% v/v) in 0.1 M NaH<sub>2</sub>PO<sub>4</sub>

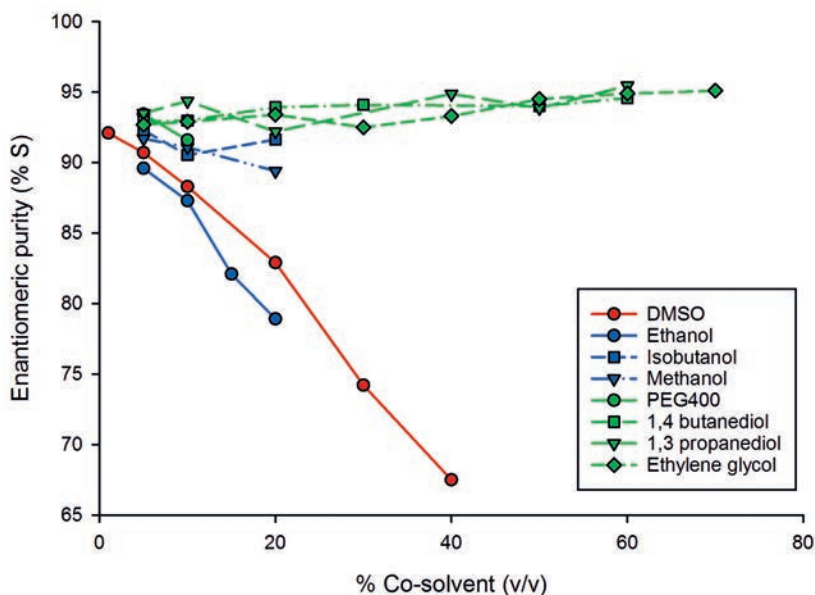
buffer (pH 7.3) to a final volume of 25 ml. The reaction was initiated by the addition of 50 mM of **1** (1.25 ml from a freshly prepared 1 M stock solution in 0.1 M NaH<sub>2</sub>PO<sub>4</sub> buffer pH 7.3) and the reaction mixture was incubated at room temperature. At timely intervals, a sample was withdrawn from the reaction mixture and the reaction progress was monitored by following the depletion in absorbance corresponding to the concentration of nitroalkene (see above). After the measurement the sample was combined with the reaction mixture. After the completion of the reaction, PRO-ALD(003) (0.5 mg/ml), PRO-NOX(009) (1 mg/ml) and NAD<sup>+</sup> (0.5 mM) were added to the reaction mixture. After 30 min, the reaction was quenched by acidifying the reaction mixture using 5 M HCl until the pH dropped to 3. To this mixture, NiCl<sub>2</sub>·6H<sub>2</sub>O (40 mM) and NaBH<sub>4</sub> (40 mM) were added at 0 °C, and stirring was continued for 24 h at room temperature. The reaction mixture was filtered through a celite pad and the filtrate concentrated in vacuo, after which the resulting concentrated mixture was acidified to pH 3-4. The acidified reaction mixture was loaded on a column packed with cation exchange resin (5g of Dowex® 50WX8 hydrogen form). After washing with deionized water (4 column volumes), the product was eluted out with 0.5 - 1 M ammonia solution (4 column volumes). The ninhydrin-positive fractions were collected, combined and lyophilized to yield the desired products **5a-5d**. Products **5a-d** were derivatized to diastereomers using sodium 2,4-dinitro-5-fluorophenyl-L-valine amide (see supporting methods) and their enantiopurity was determined using reverse phase HPLC on a C18 column (Kinetex 5u EVO C18 100A, 150 mm x 4.6 mm, Phenomenex®) at 25 °C (1 ml/min flow rate). The mobile phase was a 60:40 (v/v) mixture of aqueous buffer (0.2% triethylamine, pH adjusted to 3.5 with dilute orthophosphoric acid) and MeCN. Detection at 340 nm, retention time *S-5a*: 16.1 min, *R-5a*: 21.4 min; *R-5b*: 10.9 min, *S-5b*: 12.1 min; *R-5c*: 17.2 min, *S-5c*: 20.2 min; *R-5d*: 12.3 min, *S-5d*: 13.6 min.

## Results

### Effect of Cosolvents on the Enzymatic Michael-Type Addition Reaction

The previously reported 4-OT mutant M45Y/F50A<sup>25</sup> exhibits inverted enantioselectivity when compared to wild-type 4-OT, allowing the enzymatic synthesis of the pharmaceutically relevant enantiomers of  $\gamma$ -nitroaldehydes **3a-d** (Figure 1). However, the enantioselectivity of 4-OT M45Y/F50A is too low for biocatalytic application, with for instance the production of **3a** with an e.r. of 90:10 and **3c** with an e.r. of only 62:38<sup>25</sup>. To improve the enantioselectivity of the 4-OT M45Y/F50A catalyzed Michael-type additions, we initially focused our efforts on solvent engineering. Indeed, the use of different solvent systems has been reported to have a substantial effect on the

enantioselectivity of different enzyme catalyzed reactions<sup>26,27</sup>. We selected the 4-OT catalyzed addition of **1** to **2a** to test the effect of co-solvents on enantioselectivity, because the optical purity of product **3a** can be straightforwardly analyzed by gas chromatography. Although the enzyme was still active at concentrations of up to 40% (v/v) DMSO, increasing the DMSO concentration had a detrimental effect on the enantioselectivity of the 4-OT M45Y/F50A-catalyzed reaction (Figure 2). A similar loss of enantioselectivity was observed when short-chain alcohols such as ethanol or methanol were used as co-solvent. Notably, ethanol and methanol concentrations of >20% (v/v) led to enzyme precipitation, which indicates a destabilizing effect of the co-solvent on the biocatalyst. Using aqueous solutions of glyceline and ethaline (deep-eutectic solvents) led to inactivation of the enzyme at all tested concentrations (10–60%, v/v). Diols, however, could be used in high concentrations without observing inactivation and with a positive effect on the enantioselectivity (Figure 2). For example, using up to 60% (v/v) ethylene glycol, 1,3-propanediol or 1,4-butanediol did not affect the enzymatic activity and the e.r. of product **3a** increased up to 95:5.



**Figure 2.** Effect of cosolvents on the enantioselectivity of the enzymatic Michael-type addition reaction. Enantiopurity of **3a** (as the percentage of the *S*-enantiomer) in the 4-OT M45Y/F50A catalyzed addition of **1** to **2a** performed in the presence of increasing concentrations of different cosolvents.

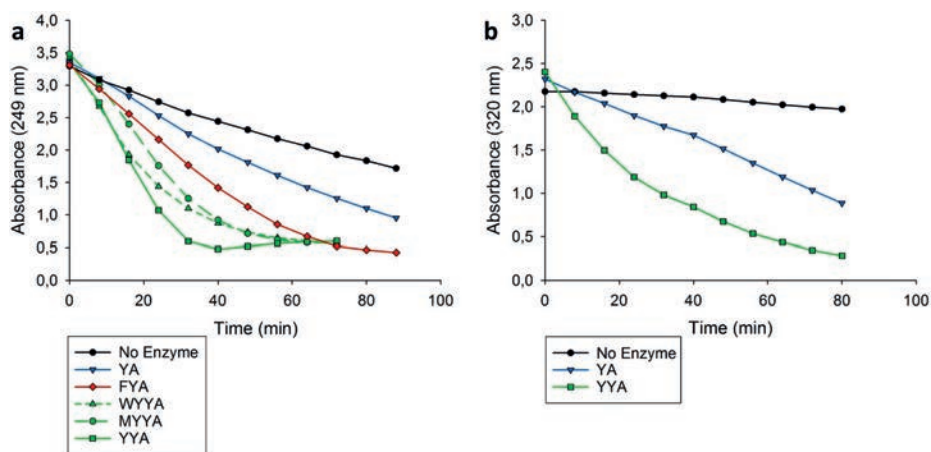
### Engineering of a Highly Enantioselective Artificial ‘Michaelase’

Although solvent engineering slightly improved the enantioselectivity of the addition of **1** to **2a** catalyzed by 4-OT M45Y/F50A, at best a product e.r. of 95:5 was achieved. Therefore, protein engineering was required to further enhance the enantioselectivity of this reaction. To guide our engineering efforts, we used the available crystal structure of 4-OT M45Y/F50A in complex with nitrostyrene **2b** (PDB: 5CLO)<sup>25</sup>. This structure suggests that the substrate may bind in two distinct orientations relative to Pro-1, differing by a rotation of ~180 degrees around the longitudinal axis of the substrate molecule. One substrate orientation is consistent with the formation of the *R* enantiomer of the product, whereas the other is consistent with formation of the *S* enantiomer. We reasoned that by reducing the rotational freedom of the substrate in the active site and/or repositioning of Pro-1 relative to the substrate, the enantioselectivity of M45Y/F50A may be improved. Therefore, we selected position Leu-8, the side chain of which is involved in formation of the binding pocket for the substrate’s R-group (Figure 1), and position Ile-2, which is located next to Pro-1, for site-saturation mutagenesis<sup>28</sup>. Accordingly, a focused library was constructed by randomizing positions Ile-2 and Leu-8 using NNK degeneracy, yielding library I2X/L8X/M45Y/F50A. As the model reaction for activity and enantioselectivity screening, the addition of **1** to **2a** was chosen because the quantity and optical purity of product **3a** can be easily analyzed by gas chromatography. We reasoned that if mutants displaying enhanced activity and enantioselectivity were to be found, then they could be tested for their ability to facilitate the enantioselective addition of acetaldehyde to aromatic nitroalkenes as well.

The library was transformed into *Escherichia coli* cells, and screened by evaluating 960 transformants. This resulted in the identification of two quadruple mutants (I2W/L8Y/M45Y/F50A and I2M/L8Y/M45Y/F50A) with significantly improved activity and enantioselectivity. These mutant enzymes were purified to homogeneity and assayed for their ability to catalyze the addition of **1** to **2a**. Mutants I2W/L8Y/M45Y/F50A and I2M/L8Y/M45Y/F50A both displayed excellent enantioselectivity, providing product *S*-**3a** with an e.r. of 99:1 (Table S1), and enhanced activity (4- and 5-fold, respectively) when compared to that of M45Y/F50A (Figure 3a). Since both quadruple mutants contain the L8Y mutation, the triple mutant L8Y/M45Y/F50A was constructed to test the importance of the mutation at position Ile-2. In addition, the triple mutant L8F/M45Y/F50A was constructed to test if a structurally similar residue at position Leu-8 might result in better catalytic properties. Interestingly, mutants L8Y/M45Y/F50A and L8F/M45Y/F50A exhibited excellent enantioselectivity in the addition reaction yielding *S*-**3a** with an e.r. of 99:1 (Table S1). To our delight, the catalytic activity of mutant L8Y/M45Y/F50A for the addition of **1** to **2a** was considerably improved (6-fold) compared to that of the starting

mutant M45Y/F50A and higher than those of the quadruple mutants I2M/L8Y/M45Y/F50A and I2W/L8Y/M45Y/F50A (Figure 3a). The activity of mutant L8F/M45Y/F50A was slightly lower than that of L8Y/M45Y/F50A. Satisfyingly, mutant L8Y/M45Y/F50A also displayed higher activity (3.5-fold) for the addition of **1** to the aromatic nitroalkene **2c** when compared to that of mutant M45Y/F50A (Figure 3b).

Coincidentally, when compared to mutant M45Y/F50A, mutant L8Y/M45Y/F50A also proved to be much more stable in the presence of the desirable cosolvent ethanol, tolerating concentrations of up to 30% (v/v) without losing activity or enantioselectivity. In fact, the enantioselectivity of the L8Y/M45Y/F50A catalyzed addition of **1** to **2a** in the presence of 25% (v/v) ethanol was exceptional, yielding optically pure *S*-**3a** (the pregabalin precursor) with an e.r. of >99:1 (Table S1). On the other hand, mutant L8F/M45Y/F50A was rather unstable in the presence of ethanol, with significantly reduced activity already at an ethanol concentration of 10% (v/v). Because the L8Y/M45Y/F50A mutant showed the highest catalytic activity and enantioselectivity, and because it tolerates relatively high concentrations of the cosolvent ethanol without loss of performance, it was selected for further study.



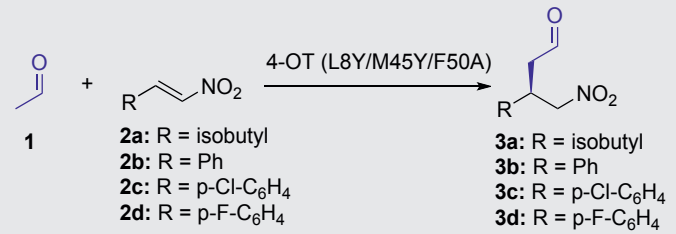
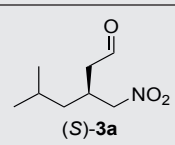
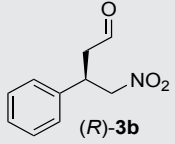
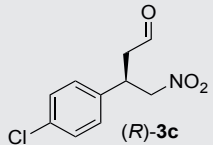
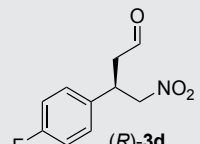
**Figure 3. Progress curves of Michael-type additions catalyzed by different 4-OT variants.** (a) Addition of **1** (150 mM) to **2a** (5 mM) in 20 mM sodium phosphate buffer (pH 6.5) and 5% (v/v) DMSO. Different 4-OT mutants were used as a catalyst at a concentration of 100  $\mu$ M. YA: 4-OT M45Y/F50A, FYA: 4-OT L8F/M45Y/F50A, WYYA: 4-OT I2W/L8Y/M45Y/F50A, MYYA: 4-OT I2M/L8Y/M45Y/F50A, YYA: 4-OT L8Y/M45Y/F50A. (b) Addition of **1** (65 mM) to **2c** (1.3 mM) in 20 mM sodium phosphate buffer (pH 6.5) and 45% (v/v) DMSO. Different 4-OT mutants were used as a catalyst at a concentration of 18  $\mu$ M. YA: 4-OT M45Y/F50A, YYA: 4-OT L8Y/M45Y/F50A.

The synthetic usefulness of mutant L8Y/M45Y/F50A was further investigated in semi-preparative scale experiments with nitroalkene acceptors **2a-d** (Table 1). Ethanol was used as a cosolvent in all reactions because in 20-30% (v/v) ethanol excellent enantioselectivity was observed for the enzymatic additions. Under optimized conditions, products **3a-d** could be synthesized with good isolated yields (63-97 %) and excellent enantioselectivity (e.r. up to >99:1). In all cases, the pharmaceutically relevant enantiomer of the  $\gamma$ -nitroaldehyde was synthesized [(*S*)-**3a** and (*R*)-**3b-d**]. These results demonstrate the potential of this newly engineered artificial 'Michaelase', which displays enhanced activity, improved enantioselectivity and better ethanol stability compared to the parent enzyme, for the asymmetric synthesis of optically pure  $\gamma$ -nitroaldehydes, which are important precursors for the valuable pharmaceuticals pregabalin, phenibut, baclofen and fluorophenibut.

### Structural Basis for the Enhanced Enantioselectivity

To obtain insight into how the active site of 4-OT L8Y/M45Y/F50A is remodelled to facilitate the highly enantioselective Michael-type addition reactions, we determined its crystal structure at 1.9 Å resolution (Table S2). The overall hexameric structure of 4-OT L8Y/M45Y/F50A is very similar to that of wild-type 4-OT (PDB entry: 4X19)<sup>23</sup> and of mutant M45Y/F50A (PDB entry: 5CLN)<sup>25</sup>, with C $\alpha$ -backbone root-mean-square-deviations of 0.38 Å and 0.69 Å, respectively. A more detailed structural comparison of 4-OT L8Y/M45Y/F50A and the parental enzyme 4-OT M45Y/F50A with bound nitrostyrene (PDB entry: 5CLO)<sup>25</sup> shows that the extra mutation (L8Y) in the triple mutant results in a slight shift of residues 5-17, comprising the last residues of strand  $\beta$ 1, the following loop, and the N-terminal residues of helix  $\alpha$ 1 (maximum shift is 1.4 Å). In addition, residues 46-48 comprising the loop/ $\beta$ <sub>10</sub>-helix after strand  $\beta$ 2 show a maximum shift of 2.2 Å. Residues Tyr-45 and Ala-50 line the active site pocket, their side chains having a conformation very similar to that observed in the M45Y/F50A mutant. Notably, two conformations of Tyr-45 were observed. In one conformation, the side chain of Tyr-45 is rotated towards the substrate binding pocket. Overlaying this conformation with the structure of M45Y/F50A in complex with nitrostyrene (**2b**) would result in a clash between the aromatic ring of **2b** and Tyr-45, suggesting that this is not the active conformation. In the second conformation, the side chain of Tyr-45 is flipped to the back of the active site, making H-bond interactions with His-6, similar as is observed in the structure of M45Y/F50A.

**Table 1.** 4-OT L8Y/M45Y/F50A catalyzed Michael-type additions of **1** to **2a-d** yielding **3a-d**<sup>a</sup>.

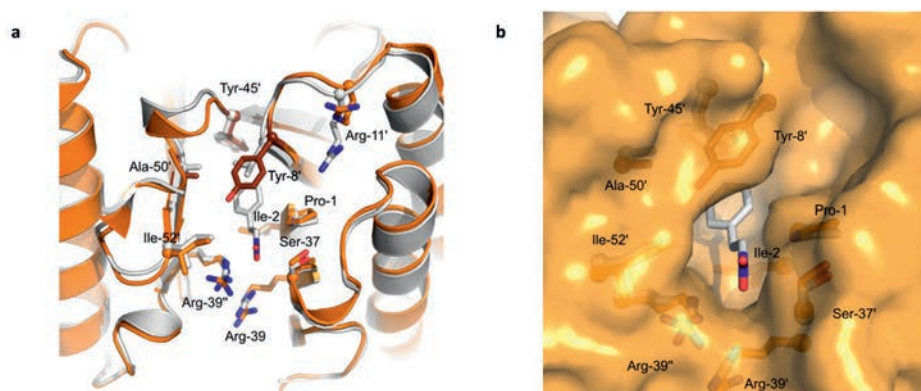
Entry	Product	Reaction time (min)	Co-solvent	Isolated yield (%)	e.r. <sup>b</sup>
	 <p><b>1</b> + <b>2a-d</b> <math>\xrightarrow{4\text{-OT (L8Y/M45Y/F50A)}}</math> <b>3a-d</b></p> <p><b>2a:</b> R = isobutyl  <b>2b:</b> R = Ph  <b>2c:</b> R = p-Cl-C<sub>6</sub>H<sub>4</sub>  <b>2d:</b> R = p-F-C<sub>6</sub>H<sub>4</sub></p> <p><b>3a:</b> R = isobutyl  <b>3b:</b> R = Ph  <b>3c:</b> R = p-Cl-C<sub>6</sub>H<sub>4</sub>  <b>3d:</b> R = p-F-C<sub>6</sub>H<sub>4</sub></p>				
1	 <p>(S)-<b>3a</b></p>	50	20% EtOH	63	>99:1
2	 <p>(R)-<b>3b</b></p>	80	20% EtOH	87	98:2
3	 <p>(R)-<b>3c</b></p>	55	30% EtOH	97	>99:1
4	 <p>(R)-<b>3d</b></p>	140	20% EtOH	94	>99:1

<sup>a</sup> Assay conditions: reaction mixtures consisted of 150 mM **1** (synthesis of **3a**), or 50 mM **1** (synthesis of **3b-d**), 3 mM **2a**, 1.3 mM **2c** or 2 mM **2b** and **2d** in 20 mM sodium phosphate buffer, pH 6.5. The reaction volume was 12.8 mL (synthesis of **3a**) or 60 mL (synthesis of **3b-d**). 1.4 mol% of 4-OT (compared to concentration nitroalkene) was used, except for entry 1 (2.5 mol%) and entry 3 (2.15 mol%).

<sup>b</sup> For **3a**, the e. r. was determined by GC with a chiral stationary phase. For **3b** and **3c**, the enzymatic product was first converted into the corresponding ethylene glycol acetal, of which the e. r. was determined by HPLC with a chiral stationary phase. For **3d**, the e.r. was directly determined by HPLC with a chiral stationary phase. The absolute configuration was determined by comparison to literature data<sup>17,25</sup>.



Most importantly, at the interface of each dimer of the L8Y/M45Y/F50A mutant, the Tyr-8 side chain lies “above” the hydrophobic pocket that binds the aromatic ring of **2b** in the M45Y/F50A mutant (Figure 4a). While the pocket has become narrower due to the extra L8Y mutation, the comparison with the M45Y/F50A structure indicates that it remains capable to accommodate **2b**, with the aromatic rings of Tyr-8 and the substrate forming a parallel displaced pi-pi stacking interaction (Figure 4b). The narrower substrate binding pocket is expected to be less tolerant towards accommodating the alternate binding mode of **2b**, which would give the opposite enantiomer of the corresponding product, because the substrate has little freedom to adjust its conformation and orientation to minimize unfavorable close contacts. These results indicate that the enhanced enantioselectivity of the L8Y/M45Y/F50A mutant is not due to major structural changes but is related to the replacement of Leu-8 by a larger tyrosine residue, resulting in a narrower substrate binding pocket that prevents the substrate to bind in two distinct orientations relative to Pro-1. Thus, the reduction in size of the substrate binding pocket explains the excellent enantioselectivity of the 4-OT L8Y/M45Y/F50A mutant.



**Figure 4. Crystal structure of the newly engineered 4-OT L8Y/M45Y/F50A mutant.** (a) Overlay of the crystal structures of mutant M45Y/F50A with bound nitrostyrene (**2b**, in light grey color, PDB entry: 5CLO) and mutant L8Y/M45Y/F50A (in orange color, this work). The extra L8Y mutation of the triple mutant has the tyrosine side chain lying above a hydrophobic pocket where nitrostyrene (**2b**) is bound in the double mutant, close to the catalytic proline residue (Pro-1). (b) A close-up view at a similar orientation, with the surface representation of the L8Y/M45Y/F50A mutant added, showing that nitrostyrene (**2b**) would snugly fit in the hydrophobic pocket.

### One-Pot Two-Step Enzymatic Cascade Synthesis of $\gamma$ -Nitrobutyric Acids.

Having constructed an artificial ‘Michaelase’ that is highly enantioselective for the addition of acetaldehyde (**1**) to nitroalkenes **2a-d**, yielding  $\gamma$ -nitroaldehydes (*S*)-**3a** and (*R*)-**3b-d** (Table 1), we focused our attention on the development of a two-step enzymatic

cascade process for the synthesis of  $\gamma$ -nitrobutyric acids **4a-d**, which are more stable and practical GABA analogue precursors than the corresponding  $\gamma$ -nitroaldehydes. The synthesis of these compounds requires the coupling of the newly developed 'Michaelase' with an appropriate aldehyde oxidase or aldehyde dehydrogenase (Table 2). Notably, using an oxidase would result in the formation of hydrogen peroxide as a by-product, which could unfavourably react with the aldehyde substrates and enzyme catalysts, necessitating an additional step for its removal. Therefore, an aldehyde dehydrogenase with a broad non-natural substrate scope, PRO-ALDH(003) available from Prozomix Ltd., was used. Initial experiments showed that PRO-ALDH(003) was dependent on either NAD<sup>+</sup> or NADP<sup>+</sup> and could accept both (*R*)-**3a** and (*S*)-**3a** as substrates. For cofactor regeneration, we used an NADH oxidase, PRO-NOX(009) available from Prozomix Ltd., which converts NADH to NAD<sup>+</sup> using O<sub>2</sub> and producing H<sub>2</sub>O. A water-forming NOX enzyme is particularly suited to include in the envisioned cascade synthesis, as it does not employ substrates or generate products that complicate subsequent work-up, resulting in increased purity and yield of **4a-d**. Using our newly engineered artificial 'Michaelase' 4-OT L8Y/M45Y/F50A in combination with PRO-ALDH(003) and PRO-NOX(009), we developed a one-pot two-step biocatalytic cascade to synthesize the desired product (*S*)-**4a** in 64% isolated yield and with an excellent e.r. of 99:1 (Table 2, entry 1). Using similar conditions, also (*R*)-**4b-d** could be synthesized in good isolated yield (67-70%) and with excellent e.r. values of up to 99:1 (Table 2, entries 2-4).

Interestingly, the modularity of this enzymatic cascade approach and the ability of PRO-ALDH(003) to accept both enantiomers of **3a-d** as substrates, also allowed for the synthesis of the opposite enantiomers of  $\gamma$ -nitrobutyric acids **4a-d**. To this end, we replaced 4-OT L8Y/M45Y/F50A by the previously engineered enantiocomplementary 4-OT variant A33D, which produces the opposite enantiomers of **3a-d**<sup>25</sup>. Under optimized conditions and using the appropriate combination of enzymes in one pot, the desired products (*R*)-**4a** and (*S*)-**4b-d** were obtained with excellent e.r. values (99:1) and in 62-74% isolated yield (Table 2, entries 5-8).

**Table 2.** One-pot two-step biocatalytic enantioselective synthesis of  $\gamma$ -nitrobutyric acids **4a-d**<sup>a</sup>.

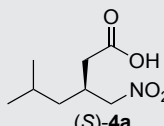
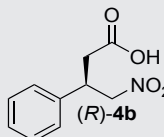
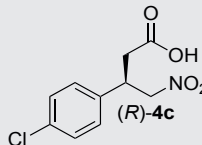
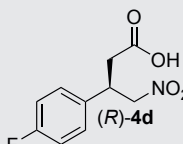
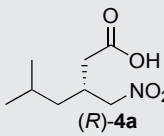
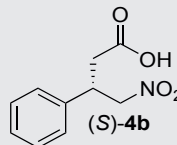
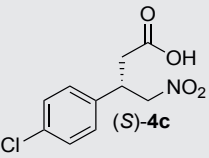
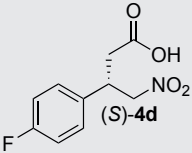
Entry	Product	4-OT variant	Reaction time		Isolated yield (%)	e.r. <sup>d</sup>
			Step I (min) <sup>b</sup>	Step II (min) <sup>c</sup>		
1	 (S)- <b>4a</b>	L8Y/M45Y/ F50A	30	10	64	99:1
2	 (R)- <b>4b</b>	L8Y/M45Y/ F50A	20	5	68	98:2
3	 (R)- <b>4c</b>	L8Y/M45Y/ F50A	30	5	70	99:1
4	 (R)- <b>4d</b>	L8Y/M45Y/ F50A	25	5	67	99:1
5	 (R)- <b>4a</b>	A33D	70	15	71	99:1
6	 (S)- <b>4b</b>	A33D	40	5	67	99:1

Table 2. Continued.

Entry	Product	4-OT variant	Reaction time		Isolated yield (%)	e.r. <sup>d</sup>
			Step I (min) <sup>b</sup>	Step II (min) <sup>c</sup>		
7	 (S)- <b>4c</b>	A33D	70	5	62	99:1
8	 (S)- <b>4d</b>	A33D	30	5	74	99:1

<sup>a</sup> Michael-type addition of **1** to **2a-d** catalyzed by 4-OT L8Y/M45Y/F50A or 4-OT A33D to form either *S*-**3a** and *R*-**3b-d** or *R*-**3a** and *S*-**3b-d** respectively, followed by aldehyde oxidation catalyzed by PRO-ALDH(003), using PRO-NOX(009) for cofactor recycling, to form either *S*-**4a** and *R*-**4b-d** or *R*-**4a** and *S*-**4b-d**. The reaction mixtures consisted of 50 mM **1**, 3 mM **2a** or 4 mM **2b-d** in 100 mM sodium phosphate buffer pH 7.3 and 10% (v/v) ethanol. 5 mol% of 4-OT (compared to concentration nitroalkene) was used, except for entry 6 (1.5 mol%) and entry 7 (3 mol%). PRO-ALDH(003) was added to a final concentration of 0.5 mg/mL, PRO-NOX(009) was added to a final concentration of 1 mg/mL. The concentration NAD<sup>+</sup> was 8 mM (**4a**) or 10 mM (**4b-d**). The amounts of applied cofactor were adjusted such that short reaction times and high conversions were achieved. <sup>b</sup> Monitored by UV spectroscopy; <sup>c</sup> Monitored by HPLC. <sup>d</sup> Products **4a-d** were esterified and the e.r. was determined by GC or HPLC with a chiral stationary phase.

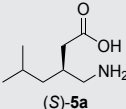
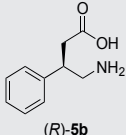
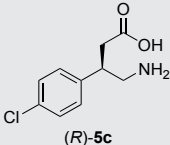
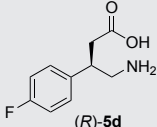
### One-Pot Three-Step Chemoenzymatic Cascade Synthesis of $\gamma$ -Aminobutyric Acids.

Having developed a modular one-pot two-step enzymatic cascade for the production of both enantiomers of  $\gamma$ -nitrobutyric acids **4a-d**, we focused our attention on the extension of this cascade by including a third step, converting **4a-d** into the desired pharmaceutically active  $\gamma$ -aminobutyric acids **5a-d**. To the best of our knowledge, no enzymes have been reported that can catalyze the full reduction of aliphatic nitro groups, such as those present in **4a-d**, to the corresponding amines<sup>29</sup>. While several

procedures for the chemical reduction of nitro groups have been reported, most of these use organic solvents and are not suitable to be executed in aqueous solutions. To prevent the necessity of intermediate purification of **4a-d**, we selected nickel boride as catalyst ( $\text{NiCl}_2 \cdot 6\text{H}_2\text{O}$  in combination with  $\text{NaBH}_4$ ), because this system has been reported to catalyze the reduction of nitro groups in aqueous solutions<sup>30,31</sup>. Pleasingly, when using 10 equivalents of nickel boride ( $\text{NaBH}_4/\text{NiCl}_2 \cdot 6\text{H}_2\text{O}$ ) the complete reduction of **4a-d** into **5a-d** in aqueous buffer at pH 3 was achieved.

Having established that nickel boride can be used to set up a three-step chemoenzymatic cascade reaction, 4-OT L8Y/M45Y/F50A, PRO-ALDH(003), PRO-NOX(009) and  $\text{NaBH}_4/\text{NiCl}_2 \cdot 6\text{H}_2\text{O}$  were combined in one pot. This chemoenzymatic cascade proved to be efficient, converting the simple starting materials **1** and **2a** into the desired product (*S*)-**5a** with an isolated yield of 55% and an excellent e.r. of 98:2 (Table 3, entry 1). Similarly, this one-pot three-step reaction cascade could also be used to efficiently synthesize (*R*)-**5b-d** in good isolated yields (65-70%) and with excellent e.r. values of 99:1 (Table 3, entries 2-4). Thus, access to the pharmaceuticals pregabalin, phenibut, baclofen and fluorophenibut has been achieved using this newly developed one-pot, three-step chemoenzymatic cascade process.

**Table 3.** One-pot three-step chemoenzymatic cascade synthesis of pharmaceuticals **5a-d**<sup>a</sup>.

Entry	Product	Reaction time			Isolated yield (%) <sup>e</sup>	e.r. <sup>f</sup>
		Step I (min) <sup>b</sup>	Step II (min) <sup>c</sup>	Step III (h) <sup>d</sup>		
1	 (S)- <b>5a</b>	30	60	24	55	98:2
2	 (R)- <b>5b</b>	20	30	24	70	99:1
3	 (R)- <b>5c</b>	30	30	24	65	99:1
4	 (R)- <b>5d</b>	25	30	24	68	99:1

<sup>a</sup> Michael-type addition of **1** to **2a-d** catalyzed by 4-OT L8Y/M45Y/F50A to form **S-3a** and **R-3b-d**, followed by aldehyde oxidation catalyzed by PRO-ALDH(003), using PRO-NOX(009) for cofactor recycling, to form **S-4a** and **R-4b-d**, followed by nitro reduction catalyzed by nickel boride to form **S-5a** and **R-5b-d**. The reaction mixtures consisted of 50 mM **1**, 3 mM **2a** or 4 mM **2b-d** in 100 mM sodium phosphate buffer pH 7.3 and 10% (v/v) ethanol. 5 mol% of 4-OT (compared to concentration of nitroalkene) was used; PRO-ALDH(003) was added to a final concentration of 0.5 mg/mL; PRO-NOX(009) was added to a final concentration of 1 mg/mL; 0.5 mM of NAD<sup>+</sup> was added; 40 mM of NiCl<sub>2</sub> and 40 mM of NaBH<sub>4</sub> were used. Because of the poor stability of nickel boride in aqueous buffer, a 10-fold excess of this reagent (40 mM versus 3-4 mM **2a-d**) is required to achieve high conversion.

<sup>b</sup> Monitored by UV spectroscopy.

<sup>c</sup> Monitored by HPLC.

<sup>d</sup> Monitored by TLC.

<sup>e</sup> Purified by cation exchange chromatography.

<sup>f</sup> Products **5a** - **5d** were derivatized using N<sub>α</sub>-(2,4-dinitro-5-fluorophenyl)-L-valinamide and the e.r. of the corresponding diastereomers was determined by HPLC with an achiral stationary phase. Notably, the e.r. of product **5a** is likely to be 99:1; however, because a minor contaminant from the derivatizing agent has the same retention time as the *R* enantiomer of product **5a**, the e.r. for product **5a** is cautiously reported as 98:2 consistent with the observed peaks in the HPLC chromatogram.

## Discussion

The use of different enzymes to build reaction cascades for efficient synthesis of valuable chemical products is an emerging concept in biocatalysis<sup>4-7</sup>. This approach aims to minimize the number of reaction steps, prevent the necessity of isolation of intermediate products and improve the ‘pot-economy’ of the process<sup>9</sup>. Using such a systems biocatalysis strategy, we have successfully demonstrated a one-pot two-step biocatalytic cascade reaction for the synthesis of the pharmaceutically relevant enantiomers of  $\gamma$ -nitrobutyric acids starting from simple achiral precursors (acetaldehyde and nitroalkenes). This synthetic route highlights a highly enantioselective carbon–carbon bond-forming step catalyzed by a newly engineered artificial ‘Michaelase’, 4-OT L8Y/M45Y/F50A, which has no known biological counterpart. This tailor-made enzyme shows non-natural substrate promiscuity, accepting acetaldehyde for enantioselective addition to various nitroalkenes. The modularity of this two-step cascade reaction was demonstrated by using a previously engineered enantiocomplementary 4-OT mutant, 4-OT A33D<sup>25</sup>, to synthesize the opposite enantiomers of the  $\gamma$ -nitrobutyric acids. We also developed a three-step chemoenzymatic cascade route for the conversion of enzymatically prepared  $\gamma$ -nitrobutyric acids by chemical reduction into the corresponding  $\gamma$ -aminobutyric acids in one pot. Given that the efficient large-scale process for the manufacture of pregabalin (**5a**) by Pfizer is based on a lipase-catalysed kinetic resolution with ex-situ enantiomer recycling<sup>32</sup>, our one-pot asymmetric chemoenzymatic cascade synthesis offers an alternative route to pharmaceutically active  $\gamma$ -aminobutyric acid products, as exemplified by the enantioselective synthesis of pregabalin (3 steps, 55% yield), phenibut (3 steps, 70% yield), baclofen (3 steps, 65% yield) and fluorophenibut (3 steps, 68% yield). This shows the exciting opportunities available for combining chemocatalysts, natural enzymes, and designed artificial biocatalysts in multistep syntheses<sup>5</sup>. It is anticipated that such a rapid synthesis strategy, which avoids (de-)protecting steps and intermediate purifications, could be flexibly extended to various other targets, for example by using aldehyde donors with different chain length or by starting from structurally distinct nitroalkenes.

The key step of the developed cascade synthesis is the enantioselective Michael addition of acetaldehyde to a nitroalkene catalyzed by an engineered 4-OT variant (mutant L8Y/M45Y/F50A). The development of this highly enantioselective 4-OT variant was achieved by using an engineering strategy that involves simultaneous randomization at two sites (Ile-2 and Leu-8, Figure 4) lining the substrate binding pocket of 4-OT M45Y/F50A, a double mutant previously designed for inverting the stereoselectivity to that required for GABA precursor synthesis.<sup>25</sup> The major benefit of this structure-guided engineering strategy lies in the reduction of the sequence space that has to be covered in order to identify valuable mutants.<sup>28</sup> Screening of 960 transformants, which is only a relatively small fraction of clones with respect to statistical requirements for significant coverage of possible combinations in a double-site SSM library with NNK codons, yielded two quadruple mutants (I2W/L8Y/M45Y/F50A and I2M/L8Y/M45Y/F50A) with significantly improved activity and the desired high enantioselectivity. The best performing triple mutant, L8Y/M45Y/F50A, was deductively constructed from these quadruple mutants. Notably, this triple mutant must have been included in the double-site SSM library, but apparently was not identified due to insufficient screening. From this result it can be concluded that a single-site SSM strategy at both positions would have been an even more efficient engineering strategy, with only the need for screening of ~200 transformants with acceptable statistical coverage of possible mutants. In summary, we provide evidence that three amino acid substitutions (L8Y/M45Y/F50A) are sufficient to invert and dramatically enhance the enantioselectivity of 4-OT for the Michael addition of acetaldehyde to various nitroalkenes.

A structural comparison of 4-OT M45Y/F50A complexed with a nitroalkene substrate<sup>25</sup> and unliganded 4-OT L8Y/M45Y/F50A suggested that the new triple mutant has a narrower binding pocket for the nitroalkene substrate (Figure 4). This reduction in size of the substrate binding pocket may prevent the nitroalkene substrate to bind in two distinct orientations relative to Pro-1, potentially explaining the enhanced enantioselectivity. However, the possibility that the observed nitroalkene substrate orientation is not relevant for the catalytic bond-forming event can not be excluded, because the covalent enzyme-bound enamine intermediate, which is formed between Pro-1 and acetaldehyde,<sup>23</sup> may influence the binding orientation of the nitroalkene substrate and the trajectory of the nucleophilic addition.

The currently employed amounts (1.4–5.0 mol%) of 4-OT L8Y/M45Y/F50A are lower and reactions times of  $\leq 30$  min are generally shorter than in the few conventional organocatalytic methodologies for identical type of Michael reactions.<sup>12,33,34</sup> Moreover, we take advantage of the water solubility of the enzyme to achieve the reactions in



environmentally friendly aqueous buffer and not in organic solvent. Nevertheless, the relatively high amounts of the 4-OT L8Y/M45Y/F50A enzyme employed, required to ensure short reaction times to outcompete non-enzymatic water addition to nitroalkene substrates **2a-d**, might pose a possible constraint on the use of our developed chemoenzymatic cascade synthesis for large-scale transformations. To address this issue, we currently run protein engineering studies with the aim to further enhance the unnatural ‘Michaelase’ activities of 4-OT L8Y/M45Y/F50A, which would allow lower catalyst loadings.

While all three steps of the cascade reaction were performed under aqueous conditions with high conversions, the chemical reduction step requires 10 equivalents of nickel boride and has to be performed at acidic pH, which is not directly compatible with the enzymatic steps that require neutral pH. It will be interesting to investigate whether the chemical reduction step can be replaced by an enzymatic step, converting the  $\gamma$ -nitrobutyric acids **4a-d** into GABA analogues **5a-d**. However, such an entirely enzymatic cascade route for the straightforward synthesis of pharmaceutically active GABA analogues awaits the discovery or engineering of an appropriate nitroreductase capable of completely reducing the aliphatic nitro group of **4a-d** to the amine group [at the expense of NAD(P)H] in a controlled and efficient fashion, without formation of unstable, reactive intermediates that lower the yield of the desired amino acid product. Although increasingly investigated, reported nitroreductases prefer nitroaromatic substrates and often catalyze only partial reduction of nitro groups to the reactive nitroso or hydroxylamino products<sup>29</sup>. We have initiated studies aimed at exploring a large collection of putative nitroreductases, as well as engineered variants thereof, for their ability to efficiently reduce  $\gamma$ -nitrobutyric acids **4a-d** into  $\gamma$ -aminobutyric acids **5a-d**. As such, our work sets the stage for further development of ‘artificial metabolisms’ for the greener, more sustainable and step-economic synthesis of an important class of pharmaceuticals.

## Supporting Information

Additional experimental procedures and characterizations of compounds. This information is available free of charge on the ACS Publications website.

### Data availability

All data are available from the corresponding author upon reasonable request.

### **Author contributions**

A.K. and J.Y.v.d.M. performed the solvent engineering. L.B. and R.v.M. performed the enzyme engineering. T.S., L.B. and S.J.C. developed the one-pot cascade reactions. L.B., T.S., and P.G.T. performed the preparative biotransformations and product analysis, including synthesis of reference compounds and the chiral high-performance liquid and gas chromatography experiments. T.P. and A.M.W.H.T. performed the X-ray crystallography experiments. G.J.P. supervised the scientific work. L.B., T.S., T.P., and G.J.P. wrote the paper.

### **Competing interests**

The authors declare no competing interests.

### **Acknowledgement**

We acknowledge financial support from the Netherlands Organization of Scientific Research (VICI grant 724.016.002), the European Research Council (PoC grant 713483), and the European Union's Horizon 2020 Research and Innovation Programme under grant agreement No 635595 (CarbaZymes). We thank M.H. de Vries and M. Jeronimus-Stratingh for their expert assistance in acquiring ESI-MS data, and Yan Ni for insightful discussions. Parts of this research were carried out at PETRA III at DESY, a member of the Helmholtz Association (HGF). We would like to thank E. Reddem and A. Burkhardt for assistance in using the P11 beamline.

## References

1. Winkler, C. K.; Clay, D.; Davies, S.; O'Neill, P.; McDaid, P.; Debarge, S.; Steflík, J.; Karmilowicz, M.; Wong, J. W.; Faber, K. Chemoenzymatic Asymmetric Synthesis of Pregabalin Precursors via Asymmetric Bioreduction of  $\beta$ -Cyanoacrylate Esters Using Ene-Reductases. *J. Org. Chem.* **2013**, *78*, 1525-1533.
2. Debarge, S.; McDaid, P.; O'Neill, P.; Frahill, J.; Wong, J. W.; Carr, D.; Burrell, A.; Davies, S.; Karmilowicz, M.; Steflík, J. Evaluation of Several Routes to Advanced Pregabalin Intermediates: Synthesis and Enantioselective Enzymatic Reduction Using Ene-Reductases. *Org. Process Res. Dev.* **2014**, *18*, 109-121.
3. Hoekstra, M. S.; Sobieray, D. M.; Schwindt, M. A.; Mulhern, T. A.; Grote, T. M.; Huckabee, B. K.; Hendrickson, V. S.; Franklin, L. C.; Granger, E. J.; Karrick, G. L. Chemical Development of CI-1008, an Enantiomerically Pure Anticonvulsant. *Org. Process Res. Dev.* **1997**, *1*, 26-38.
4. Muschiol, J.; Peters, C.; Oberleitner, N.; Mihovilovic, M. D.; Bornscheuer, U. T.; Rudroff, F. Cascade Catalysis – Strategies and Challenges En Route to Preparative Synthetic Biology. *Chem. Commun.* **2015**, *51*, 5798-5811.
5. Rudroff, F.; Mihovilovic, M. D.; Gröger, H.; Snajdrova, R.; Iding, H.; Bornscheuer, U. T. Opportunities and Challenges for Combining Chemo- and Biocatalysis. *Nat. Catal.* **2018**, *1*, 12.
6. Fessner, W. Systems Biocatalysis: Development and Engineering of Cell-Free “Artificial Metabolisms” for Preparative Multi-Enzymatic Synthesis. *New Biotechnol.* **2015**, *32*, 658-664.
7. Faber, K.; Fessner, W.; Turner, N. J. Engineering Biocatalysts for Synthesis Including Cascade Processes. *Adv. Synth. Catal.* **2015**, *357*, 1565-1566.
8. Payer, S. E.; Pollak, H.; Schmidbauer, B.; Hamm, F.; Juričić, F.; Faber, K.; Glueck, S. M. Multienzyme One-Pot Cascade for the Stereoselective Hydroxyethyl Functionalization of Substituted Phenols. *Org. Lett.* **2018**, *20*, 5139-5143.
9. Hayashi, Y. Pot Economy and One-Pot Synthesis. *Chem. Sci.* **2016**, *7*, 866-880.
10. Wiesner, M.; Upert, G.; Angelici, G.; Wennemers, H. Enamine Catalysis with Low Catalyst Loadings - High Efficiency via Kinetic Studies. *J. Am. Chem. Soc.* **2009**, *132*, 6-7.
11. Wiesner, M.; Revell, J. D.; Tonazzi, S.; Wennemers, H. Peptide Catalyzed Asymmetric Conjugate Addition Reactions of Aldehydes to Nitroethylene - A Convenient Entry into  $\gamma$ -Amino Acids. *J. Am. Chem. Soc.* **2008**, *130*, 5610-5611.
12. Wiesner, M.; Revell, J. D.; Wennemers, H. Tripeptides as Efficient Asymmetric Catalysts for 1,4-Addition Reactions of Aldehydes to Nitroolefins—A Rational Approach. *Angew. Chem., Int. Ed.* **2008**, *120*, 1897-1900.
13. Wiesner, M.; Neuburger, M.; Wennemers, H. Tripeptides of the Type H-D-Pro-Pro-Xaa-NH<sub>2</sub> as Catalysts for Asymmetric 1,4-Addition Reactions: Structural Requirements for High Catalytic Efficiency. *Chem. Eur. J.* **2009**, *15*, 10103-10109.

14. Miao, Y.; Rahimi, M.; Geertsema, E. M.; Poelarends, G. J. Recent Developments in Enzyme Promiscuity for Carbon–Carbon Bond-Forming Reactions. *Curr. Opin. Chem. Biol.* **2015**, *25*, 115-123.
15. Geertsema, E. M.; Poelarends, G. J. Enzymatic Carbon–Carbon Bond-Forming Michael-Type Additions. In *Science of Synthesis: Biocatalysis in Organic Synthesis 2*; Faber, K., Fessner, W. D. and Turner, N., Eds.; Thieme: Stuttgart, Germany, **2014**; Vol. 2, p 313-333.
16. Garrabou, X.; Verez, R.; Hilvert, D. Enantiocomplementary Synthesis of  $\gamma$ -Nitroketones Using Designed and Evolved Carboligases. *J. Am. Chem. Soc.* **2017**, *139*, 103-106.
17. Geertsema, E. M.; Miao, Y.; Tepper, P. G.; de Haan, P.; Zandvoort, E.; Poelarends, G. J. Biocatalytic Michael-Type Additions of Acetaldehyde to Nitroolefins with the Proline-Based Enzyme 4-Oxalocrotonate Tautomerase Yielding Enantioenriched  $\gamma$ -Nitroaldehydes. *Chem. Eur. J.* **2013**, *19*, 14407-14410.
18. Miao, Y.; Tepper, P. G.; Geertsema, E. M.; Poelarends, G. J. Stereochemical Control of Enzymatic Carbon–Carbon Bond-Forming Michael-Type Additions by “Substrate Engineering”. *Eur. J. Org. Chem.* **2016**, *32*, 5350-5354.
19. Zandvoort, E.; Geertsema, E. M.; Baas, B.; Quax, W. J.; Poelarends, G. J. Bridging Between Organocatalysis and Biocatalysis: Asymmetric Addition of Acetaldehyde to  $\beta$ -Nitrostyrenes Catalyzed by a Promiscuous Proline-Based Tautomerase. *Angew. Chem., Int. Ed.* **2012**, *124*, 1266-1269.
20. Miao, Y.; Geertsema, E. M.; Tepper, P. G.; Zandvoort, E.; Poelarends, G. J. Promiscuous Catalysis of Asymmetric Michael-Type Additions of Linear Aldehydes to  $\beta$ -Nitrostyrene by the Proline-Based Enzyme 4-Oxalocrotonate Tautomerase. *ChemBioChem* **2013**, *14*, 191-194.
21. Zandvoort, E.; Geertsema, E. M.; Quax, W. J.; Poelarends, G. J. Enhancement of the Promiscuous Aldolase and Dehydration Activities of 4-Oxalocrotonate Tautomerase by Protein Engineering. *ChemBioChem* **2012**, *13*, 1274-1277.
22. Zandvoort, E.; Baas, B.; Quax, W. J.; Poelarends, G. J. Systematic Screening for Catalytic Promiscuity in 4-Oxalocrotonate Tautomerase: Enamine Formation and Aldolase Activity. *ChemBioChem* **2011**, *12*, 602-609.
23. Poddar, H.; Rahimi, M.; Geertsema, E. M.; Thunnissen, A.M.W.H.; Poelarends, G. J. Evidence for the Formation of an Enamine Species During Aldol and Michael-type Addition Reactions Promiscuously Catalyzed by 4-Oxalocrotonate Tautomerase. *ChemBioChem* **2015**, *16*, 738-741.
24. van der Meer, J.; Biewenga, L.; Poelarends, G. J. The Generation and Exploitation of Protein Mutability Landscapes for Enzyme Engineering. *ChemBioChem* **2016**, *17*, 1792-1799.
25. van der Meer, J.; Poddar, H.; Baas, B.; Miao, Y.; Rahimi, M.; Kundendorf, A.; van Merkerk, R.; Tepper, P. G.; Geertsema, E. M.; Thunnissen, A.M.W.H.; Quax, W.J.; Poelarends, G.J. Using Mutability Landscapes of a Promiscuous Tautomerase to Guide the Engineering of Enantioselective Michaelases. *Nat. Commun.* **2016**, *7*, 10911.
26. Carrea, G.; Ottolina, G.; Riva, S. Role of Solvents in the Control of Enzyme Selectivity in Organic Media. *Trends Biotechnol.* **1995**, *13*, 63-70.

## Chapter 2

27. Klibanov, A. M. Enzymatic Catalysis in Anhydrous Organic Solvents. *Trends Biochem. Sci.* **1989**, *14*, 141-144.
28. Reetz, M. T.; Bocola, M.; Carballeira, J. D.; Zha, D.; Vogel, A. Expanding the Range of Substrate Acceptance of Enzymes: Combinatorial Active-Site Saturation Test. *Angew. Chem., Int. Ed.* **2005**, *44*, 4192-4196.
29. Durchschein, K.; Hall, M.; Faber, K. Unusual Reactions Mediated by FMN-Dependent Ene- and Nitro-Reductases. *Green Chem.* **2013**, *15*, 1764-1772.
30. Khurana, J. M.; Gogia, A. Synthetically Useful Reactions with Nickel Boride. A Review. *Org. Prep. Proced. Int.* **1997**, *29*, 1-32.
31. Osby, J. O.; Ganem, B. Rapid and Efficient Reduction of Aliphatic Nitro Compounds to Amines. *Tetrahedron Lett.* **1985**, *26*, 6413-6416.
32. Martinez, C. A.; Hu, S.; Dumond, Y.; Tao, J.; Kelleher, P.; Tully, L. Development of a Chemoenzymatic Manufacturing Process for Pregabalin. *Org. Process Res. Dev.*, **2008**, *12*, 392-398.
33. García-García, P.; Ladépêche, A.; Halder, R.; List, B. Catalytic Asymmetric Michael Reactions of acetaldehyde. *Angew. Chem., Int. Ed.* **2008**, *47*, 4719-4721.
34. Hayashi, Y.; Itoh, T.; Ohkubo, M.; Ishikawa, H. Asymmetric Michael Reaction of Acetaldehyde Catalyzed by Diphenylprolinol Silyl Ether. *Angew. Chem., Int. Ed.* **2008**, *47*, 4722-4724.

## Supporting Information

### Supporting methods

#### Mutant library construction

The 4-OT I2X/L8X/M45Y/F50A library was constructed by Quikchange mutagenesis. The following primers were used: Fw 5'-GGA GAT ATA CAT ATG CCT **NNK** GCC CAG ATC CAC ATC **NNK** GAA GGC CGC AGC G-3' and Rv 5'-C GCT GCG GCC TTC **MNN** GAT GTG GAT CTG GGC **MNN** AGG CAT ATG TAT ATC TCC-3'. The mutated codons are indicated in bold. An amount of 40 ng of pJexpress 414 plasmid DNA containing the 4-OT M45Y/F50A gene was used as template<sup>1</sup>. The PCR reaction was conducted using Phusion polymerase (New England Biolabs) in 1x HF buffer (New England Biolabs) in 50  $\mu$ l reaction volume. The following PCR program was used: 98 °C, 10 min (initial denaturation), followed by 18 cycles of 95 °C for 30 s, 55 °C for 1 min, 68 °C for 4 min and a final elongation step of 68 °C for 10 min. The reaction mixtures were 2x diluted with water followed by the addition of 1  $\mu$ l FastDigest DpnI (ThermoFisher Scientific) to digest the template DNA. Reaction mixtures were placed at 37 °C for 2 h, followed by a DpnI denaturing step of 80 °C for 10 min. An initial transformation in *Escherichia coli* Dh5 $\alpha$  cells was performed and the plasmid DNA of individual colonies was isolated and sequenced to confirm the library quality. To collect the library DNA, four parallel transformations were performed using 50  $\mu$ l of 10-beta competent *E. coli* cells (High efficiency, New England Biolabs) and 5  $\mu$ l of the PCR reaction mixture. Transformants were selected on LB-agar plates supplemented with 100  $\mu$ g/ml ampicillin. About 8000 colonies were resuspended in 10 mM sodium phosphate buffer and plasmid DNA was isolated and subsequently used to transform *E. coli* BL21(DE3) cells for expression of the 4-OT mutants.

#### Production and purification of enzymes

The enzyme 4-OT M45Y/F50A was purified according to an earlier described purification procedure<sup>2</sup>. Mutant enzymes 4-OT I2M/L8Y/M45Y/F50A, 4-OT I2W/L8Y/M45Y/F50A, 4-OT L8F/M45Y/F50A and 4-OT L8Y/M45Y/F50A were purified according to a modified protocol. Accordingly, the cell pellets were resuspended in 10 mM TRIS buffer pH 8 (instead of 10 mM sodium phosphate buffer). After loading on the DEAE-sepharose column, the column was washed 2x with 8 ml 10 mM TRIS buffer pH 8. The 4-OT enzyme was eluted with 3x 8 ml of 10 mM sodium phosphate buffer pH 7.3. Typically, the second and third elution fraction would contain pure 4-OT (>95% purity as determined by SDS-PAGE). Fractions containing pure 4-OT enzyme were combined and concentrated using a vivaspin centrifugal concentrator (Sartorius, 5000

MWCO). After concentration to ~3 ml, the sample was two times diafiltrated with 20 ml of 10 mM sodium phosphate buffer pH 7.3 to remove any traces of the TRIS buffer. The concentration of purified 4-OT was determined by the Waddell method<sup>3</sup>. The purified 4-OT was aliquoted, flash frozen in liquid nitrogen and stored at -80 °C until further use. All purified 4-OT mutants were analyzed by electron spray ionization mass spectrometry to confirm the correct mass of the protein. PRO-ALDH(003) and PRO-NOX(009) were provided as crude cell-free extracts by Prozomix Ltd, and used without any further purification.

### Progress curves of the enzyme-catalyzed reactions

To monitor the progress of the enzymatic addition of **1** to **2a**, the reaction mixtures consisted of the following: 5 mM **2a**, DMSO (5% v/v), 100 μM 4-OT (2 mol% compared to **2a**) in 20 mM sodium phosphate buffer pH 6.5, 500 μl final volume. The reaction was initiated by the addition of **1** to a final concentration of 150 mM (from a freshly prepared 1.5 M stock solution in 20 mM sodium phosphate buffer pH 6.5). Every 8 min a sample was withdrawn from the reaction mixture and a UV-VIS spectrum was measured from 200 to 500 nm. After the measurement, the sample was returned to the reaction mixture. The progress curve was constructed based on the depletion in absorbance at 249 nm, which corresponds to the  $\lambda_{\max}$  of substrate **2a**<sup>4</sup>. After 88 minutes, the reaction mixtures were extracted with 500 μl toluene. The organic layer was separated from the aqueous layer by centrifugation. A sample from the organic layer was transferred to a GC vial and analyzed by GC with a chiral stationary phase (see library screening). No product **3a** could be identified for the control reaction without enzyme.

To monitor the progress of the enzymatic addition of **1** to **2c**, the reaction mixtures consisted of the following: 1.3 mM **2c**, DMSO (45% v/v), 18 μM 4-OT (1.4 mol% compared to **2c**) in 20 mM sodium phosphate buffer pH 6.5, 500 μl final volume. The reaction was initiated by the addition of **1** to a final concentration of 65 mM (from a freshly prepared 650 mM stock solution in sodium phosphate buffer pH 6.5). Every 8 min a sample was withdrawn from the reaction mixture and a UV-VIS spectrum was measured from 200 to 500 nm. After the measurement, the sample was returned to the reaction mixture. The progress curve was constructed based on the depletion in absorbance at 320 nm, which corresponds to the  $\lambda_{\max}$  of substrate **2c**<sup>4</sup>.

### Crystallization and data collection

The newly engineered artificial Michaelase (4-OT L8Y/M45Y/F50A) was crystallized by hanging drop vapour diffusion at 293 K. Drops consisted of equal volumes of protein solution (14.5 mg/ml) in 10 mM sodium phosphate buffer, pH 7.3 and 18% (w/v) PEG

3350, 0.2 M sodium formate, 0.1 M Bis-Tris propane-HCl, pH 7.0. Rod-like crystals appeared within a few days and were cryoprotected using 24% (w/v) PEG 3350, 16% (v/v) glycerol, 50 mM sodium formate, 50 mM Bis-Tris propane-HCl, pH 6.5. Diffraction data were collected at beamline P11 of DESY (Hamburg, Germany) and indexed, integrated and scaled using XDS<sup>5</sup>. Statistics are given in Table S2.

### **Structure determination and refinement**

The structure of 4-OT L8Y/M45Y/F50A, containing three hexamers (chains A-F, G-L and M-R) in the asymmetric unit of the  $P 2_12_12_1$  cell, was determined by molecular replacement using PHASER<sup>6</sup> with the wild-type 4-OT structure (PDB entry: 4X19)<sup>7</sup> as a search model. Crystallographic refinement was carried out with Refmac<sup>8</sup> using non-crystallographic symmetry restraints, and alternated with map inspection and manual rebuilding in Coot<sup>9</sup>. Refinement statistics and model quality are given in Table S2. Structural figures were prepared with PyMOL (The PyMOL Molecular Graphics System, Version 2.0 Schrödinger, LLC). Atomic coordinates and structure factors have been deposited at the Protein Data Bank with entry 6FPS.

### **Solvent screening**

The effect of increasing concentrations of different co-solvents on the enantioselectivity of 4-OT M45Y/F50A for the addition of **1** to **2a** was investigated. Small scale reactions were set up consisting of the following: 325  $\mu$ g 4-OT M45Y/F50A (160  $\mu$ M), 3 mM **2a**, 50 mM **1**, and co-solvent in 20 mM sodium phosphate buffer pH 5.5, 300  $\mu$ l final volume. The reactions were performed in a 1 mm quartz cuvette. The following co-solvents were used: DMSO, ethanol, methanol, isobutanol, PEG400, 1,4-butanediol, 1,3-propanediol, ethylene glycol, glyceline, and ethaline in concentrations ranging from 1% to 70%. The decrease in absorbance at 249 nm, corresponding to the concentration of **2a**, was followed over time. In case 4-OT M45Y/F50A precipitated during the course of the reaction, the reaction mixture was discarded. After the reaction was finished, the reaction mixture was extracted with 700  $\mu$ l toluene. The organic layer was collected and concentrated by applying a  $N_2$  flow. A sample of the remaining organic layer was analyzed by GC to determine the enantiopurity of the enzymatic product **3a**. The following GC program was used: gradual temperature increase: 40 °C to 120 °C at 10 °C per min. Gradual temperature increase 120 °C to 140 °C at 0.5 °C per min. Gradual temperature decrease 140 °C to 40 °C at 10 °C per min. Flame ionization detection: Retention time *S*-enantiomer: 30.4 min, retention time *R*-enantiomer 31.3 min. The assignment of the absolute configuration was based on earlier reported data<sup>1,4</sup>.



## Synthesis of racemic references

### Synthesis of racemic $\gamma$ -nitrobutyraldehydes

The racemic  $\gamma$ -nitrobutyraldehydes **3a-d**, which served as reference compounds for analyses by HPLC and GC on a chiral stationary phase, were synthesized according to literature procedures<sup>4</sup>.

### Synthesis of racemic $\gamma$ -nitrobutyric acids (**4a-d**)

The racemic  $\gamma$ -nitrobutyric acids **4a-d** were synthesized from the corresponding racemic  $\gamma$ -nitrobutyraldehydes **3a-d** according to literature procedures<sup>10</sup>. The racemic  $\gamma$ -nitrobutyraldehyde **3a** (0.5 mmol) was dissolved in 3 ml of *tert*-butanol:water (3:1) and cooled to 0°C. To this  $\text{NaH}_2\text{PO}_4 \cdot 2\text{H}_2\text{O}$  (1 mmol), 2-methyl-2-butene (2 mmol) and  $\text{NaClO}_2$  (6.96 g, 1.5 mmol) were added and the mixture was stirred for 30 min. The reaction mixture was quenched with a saturated  $\text{NH}_4\text{Cl}$  solution and extracted with EtOAc. The organic layer was dried over anhydrous  $\text{MgSO}_4$ , concentrated in vacuo and the crude reaction product was purified by silica gel column chromatography using petroleum ether:EtOAc (8:2) to get racemic **4a**. The same procedure was followed for the synthesis of **4b-4d**.

### Synthesis of racemic $\gamma$ -aminobutyric acids (**5a-d**)

Racemic  $\gamma$ -aminobutyric acids **5a** and **5c** were commercially available. Racemic **5b** and **5d** were synthesized from the corresponding racemic  $\gamma$ -nitrobutyric acids (**4b** and **4d**) according to literature procedures<sup>11</sup>. Racemic **4b** (0.1 mmol) was dissolved in 10% ethanol (3 ml) and cooled to 0°C. To this  $\text{NiCl}_2 \cdot 6\text{H}_2\text{O}$  (1 mmol), followed by  $\text{NaBH}_4$  (1 mmol) was added carefully. The reaction mixture was stirred at room temperature for 24 h. After complete conversion, the reaction mixture was filtered through celite pad and the filtrate was concentrated in vacuo. The resulting concentrated mixture was acidified to pH 3-4 and then loaded on a column packed with cation exchange resin (5 g of Dowex® 50WX8 hydrogen form). After washing with deionized water (4 column volumes), the product was eluted out with 0.5 M (4 column volumes) - 1 M ammonia solution (4 column volumes). The ninhydrin-positive fractions were collected, combined and lyophilized to yield racemic product **5b**. Racemic **5d** was synthesized by following the same procedure.

## Product derivatization for determination of enantiomeric excess and reaction progress

### Derivatization of **3b** and **3c** for enantiomeric excess determination

The aldehyde functionality of **3b** was derivatized to a cyclic acetal according to a literature procedure<sup>4,12</sup>. The aldehyde functionality of **3c** was also derivatized to the cyclic acetal. To a solution of 100 mg (0.44 mmol) racemic **3c** in 16 ml DCM (anhydrous) and 0.51 g (8.2 mmol) ethylene glycol (anhydrous), 15 mg of paratoluene sulfonic acid was added. The reaction was stirred overnight at room temperature. After silica column purification the cyclic acetal was obtained (91 mg, 76% yield).

<sup>1</sup>H NMR (500 MHz, Chloroform-*d*)  $\delta$  7.30 (d,  $J = 8.4$  Hz, 2H), 7.17 (d,  $J = 8.5$  Hz, 2H), 4.76 (dd,  $J = 12.7, 6.1$  Hz, 1H), 4.73 (dd,  $J = 6.2, 3.2$  Hz, 0H), 4.55 (dd,  $J = 12.6, 9.4$  Hz, 1H), 3.99 – 3.93 (m, 1H), 3.85 – 3.79 (m, 1H), 3.79 – 3.72 (m, 1H), 2.08 (ddd,  $J = 14.2, 7.7, 3.2$  Hz, 1H), 2.00 – 1.94 (m, 1H).

<sup>13</sup>C NMR (126 MHz, CDCl<sub>3</sub>)  $\delta$  137.97, 133.68, 129.30, 128.93, 102.21, 80.26, 65.12, 65.05, 39.43, 37.27.

### Derivatization for HPLC analysis of reaction progress<sup>13</sup>

For the cascade reactions, a sample (50  $\mu$ l) from the reaction mixture was taken at constant time intervals and mixed with 50  $\mu$ l of derivatizing agent, which contains O-benzylhydroxylamine hydrochloride 40 mg mL<sup>-1</sup>; 0.25 mmol mL<sup>-1</sup> in pyridine:methanol:water (9:9:2). After incubation for 30 min at room temperature, the sample was centrifuged at 13,000 rpm for 10 min and then the supernatant was analyzed by reverse phase HPLC using a C18 column (Kinetex 5u EVO C18 100A, 150 mm x 4.6 mm, Phenomenex®). Mobile phase A was water with 0.1% TFA and mobile phase B was ACN with 0.1% TFA. The detailed mobile phase gradient was as follows: 0 min, 5% B; 0.01–6 min, 8–35% B; 6–14 min, 35–85% B; 14–15.5 min, 85–90% B; 15.5–16.5 min 90–8% B and 16.5–20 min, 5% B, while the flow rate was 1 ml/min.

### Derivatization of **4a-d** for enantiomeric excess determination

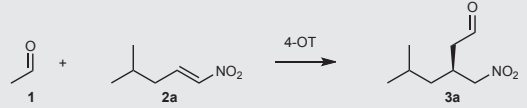
The compound **4a** (0.1 mmol) was dissolved in dichloromethane (1 ml) and methanol (0.2 ml), and cooled down to 0 °C. To this cooled solution, EDC.HCl (0.15 mmol) was added, followed by DMAP (0.01 mmol). After 1 h, the reaction mixture was quenched with saturated NH<sub>4</sub>Cl, extracted with diethyl ether and the organic layer was dried over anhydrous Na<sub>2</sub>SO<sub>4</sub>. The dried organic layer was concentrated in vacuo and the crude product was purified by silica gel column chromatography (hexane/ethylacetate

9:1) to yield the methyl ester of **4a** (75% yield). The same procedure was followed for derivatization of **4b-d**.

**Derivatization of 5a-d for enantiomeric excess determination**

$\gamma$ -Aminobutyric acids **5a-d** were derivatized to diastereomers using sodium 2,4-dinitro-5-fluorophenyl-L-valine amide<sup>14</sup>. To a solution of **5a** in water (0.1 ml, 1 mg/ml), a freshly prepared solution of sodium 2,4-dinitro-5-fluorophenyl-L-valine amide in MeCN/water 10:90 (0.1 ml, 10 mg/ml) was added. To this triethylamine (2.5  $\mu$ l) was added and the sample was vortexed at high speed. After 10 min incubation at room temperature, the sample was directly analyzed by HPLC. The same procedure was followed for derivatization of **5b-d**.

## Supporting Tables

**Table S1.** Enantioselectivity of 4-OT mutants for the addition of **1** to **2a** to yield **3a**<sup>a</sup>.


Entry	4-OT mutant	Co-solvent	Product e.r. <sup>b</sup>	Abs. Conf. <sup>c</sup>
1	M45Y/F50A	5 % DMSO	90:10	S
2	I2W/L8Y/M45Y/F50A	5 % DMSO	99:1	S
3	I2M/L8Y/M45Y/F50A	5 % DMSO	99:1	S
4	L8F/M45Y/F50A	5 % DMSO	99:1	S
5	L8Y/M45Y/F50A	5 % DMSO	99:1	S
6	L8Y/M45Y/F50A	25 % EtOH	>99:1	S

<sup>a</sup> Assay conditions: The reaction mixture consisted of 150 mM **1**, 5 mM **2a**, and 100  $\mu$ M of biocatalyst in 20 mM sodium phosphate buffer pH 6.5, 300  $\mu$ l reaction volume.

<sup>b</sup> Determined by GC with a chiral stationary phase.

<sup>c</sup> Determined by literature comparison<sup>1,4</sup>.

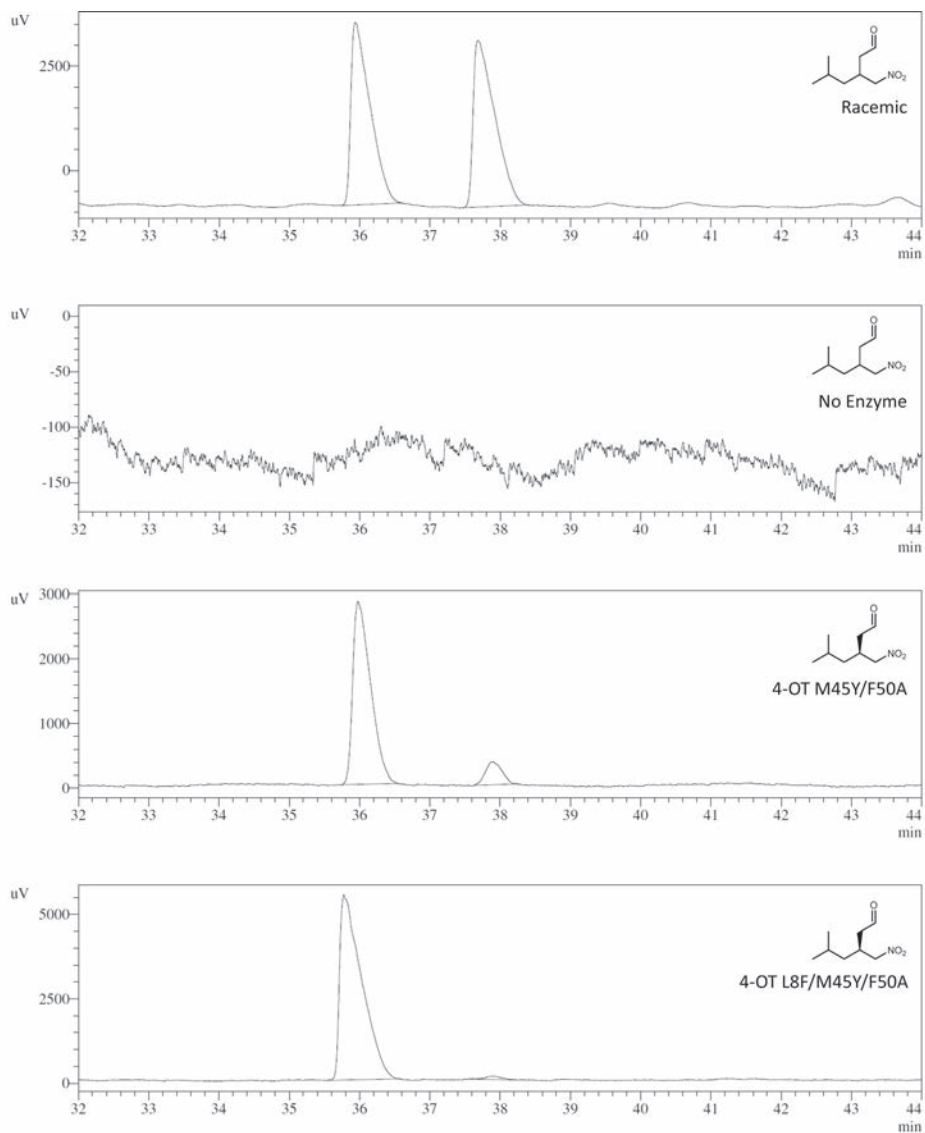
**Table S2.** Crystallographic data.

PDB entry	6FPS
<b>Data collection</b>	
Space group	$P 2_1 2_1 2_1$
Cell dimensions $a, b, c$ (Å)	90.0, 99.5, 118.3
Resolution (Å)	46-1.90 (1.93-1.90)
$R_{merge}$	0.081 (1.43)
$R_{pim}$	0.044 (0.820)
$CC_{1/2}$	0.999 (0.742)
$\langle I/\sigma \rangle$	14.2 (1.5)
Completeness (%) *	100.0 (100.0)
Redundancy *	8.1(7.7)
<b>Refinement</b>	
Resolution (Å)	46-1.90 (1.93-1.90)
Unique observations *	80195 (4096)
$R / R_{free}$	0.201 / 0.222
Number of atoms	
protein	8283
waters	341
Ligand molecules	phosphate ion (1), glycerol (1)
Average B-factor protein (Å <sup>2</sup> )	39.3
<b>Root mean square deviations</b>	
Bond lengths (Å)	0.015
Bond angles (°)	1.64
<b>Ramachandran</b>	
Favored (%)	99.9
Allowed (%)	0.1
Outliers (%)	0.0

\* values in brackets refer to the highest resolution shell

## Supporting Figures

### GC chromatograms of products from small scale reactions



**Figure S1.** GC chromatograms of racemic or enzymatically obtained **3a** in small scale reactions. The experimental conditions and product e.r. values are listed in Table S1.

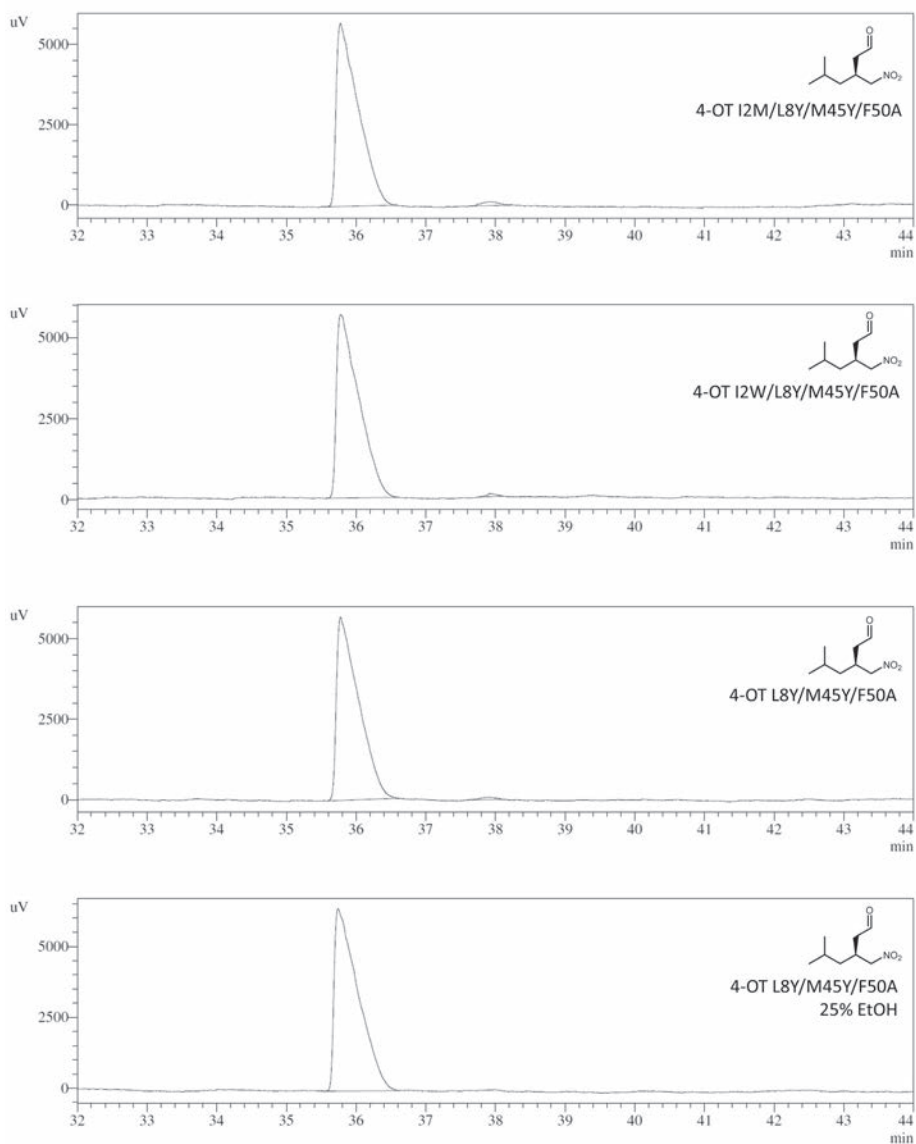
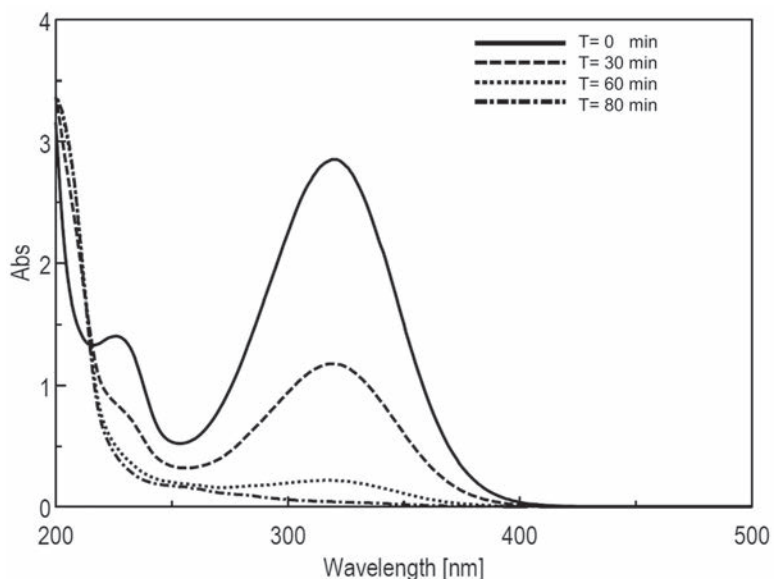
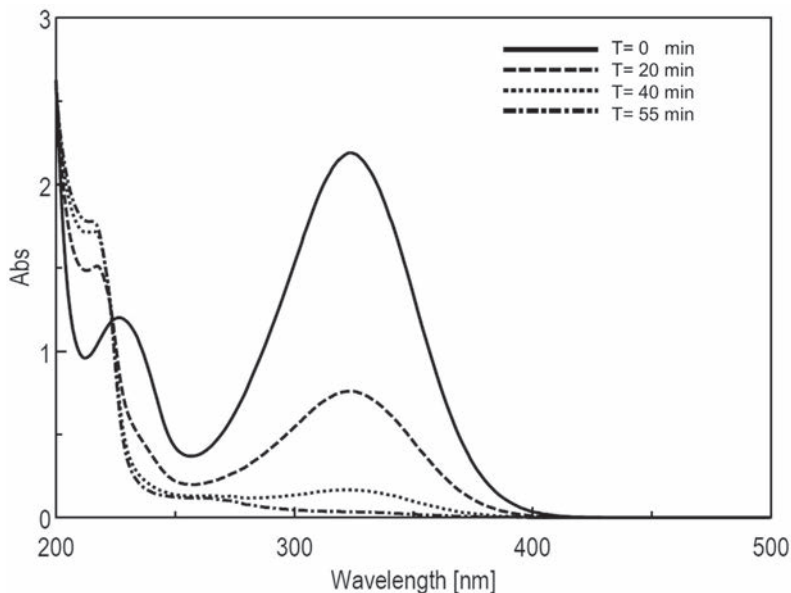


Figure S1. Continued.

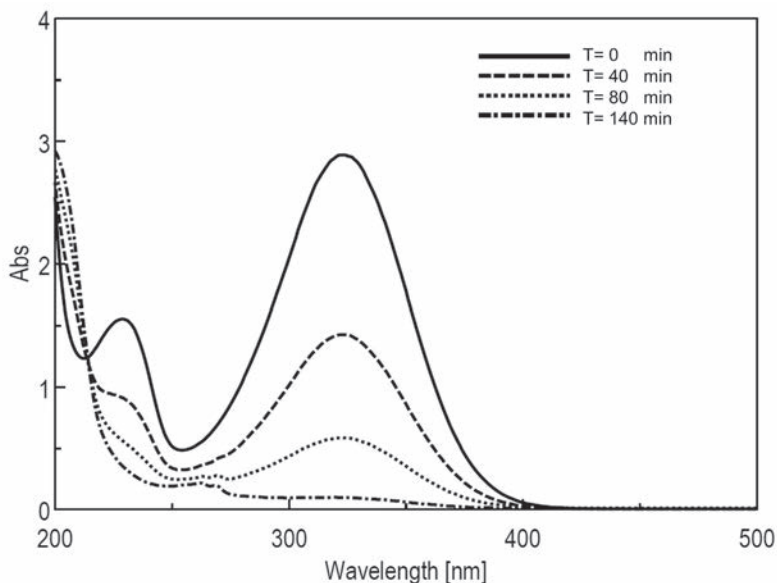
UV absorbance spectra monitoring the 4-OT L8Y/M45Y/F50A catalyzed addition of **1** to **2a-d**



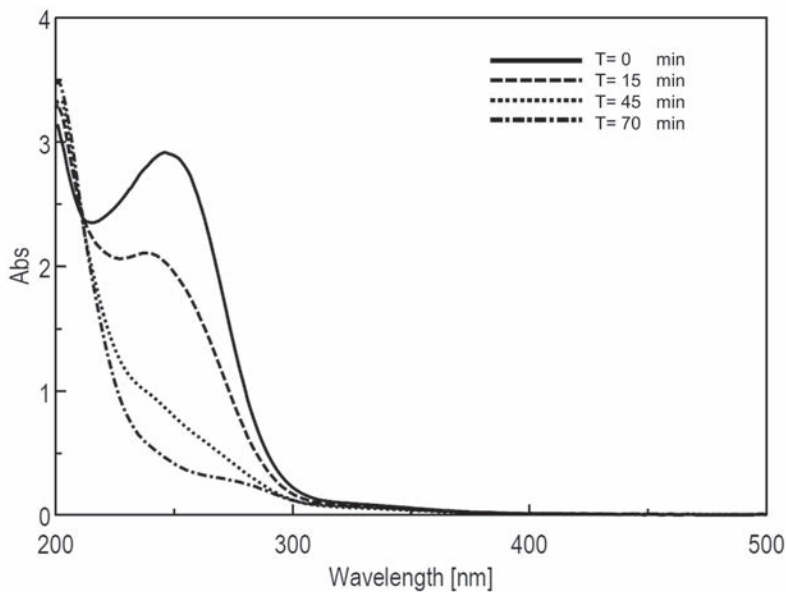
**Figure S2.** UV spectra showing the conversion of **2a** catalyzed by 4-OT L8Y/M45Y/F50A in 20 mM sodium phosphate buffer, pH 6.5, 20% ethanol.



**Figure S3.** UV spectra showing the conversion of **2b** catalyzed by 4-OT L8Y/M45Y/F50A in 20 mM sodium phosphate buffer, pH 6.5, 20% ethanol.

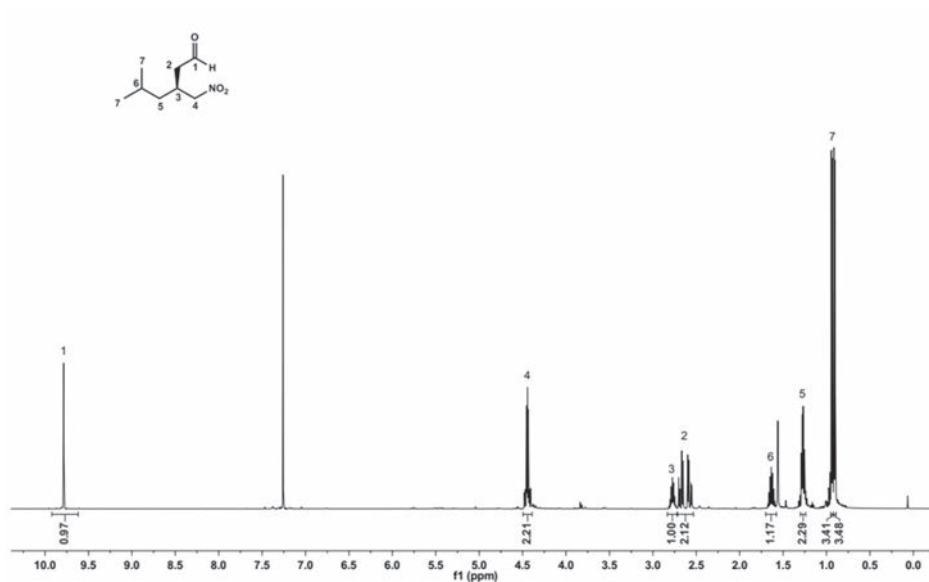


**Figure S4.** UV spectra showing the conversion of **2c** catalyzed by 4-OT L8Y/M45Y/F50A in 20 mM sodium phosphate buffer, pH 6.5, 30% ethanol.

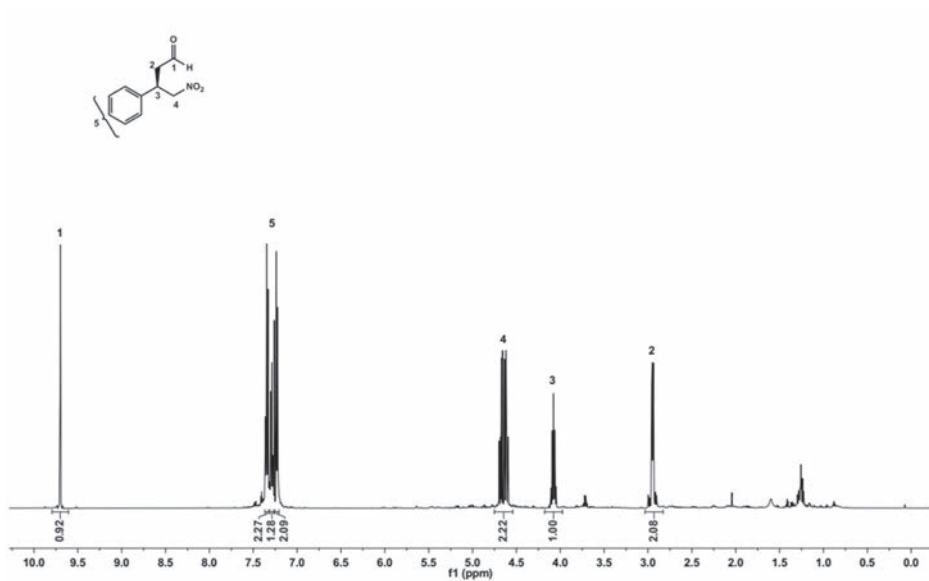


**Figure S5.** UV spectra showing the conversion of **2d** catalyzed by 4-OT L8Y/M45Y/F50A in 20 mM sodium phosphate buffer, pH 6.5, 20% ethanol.

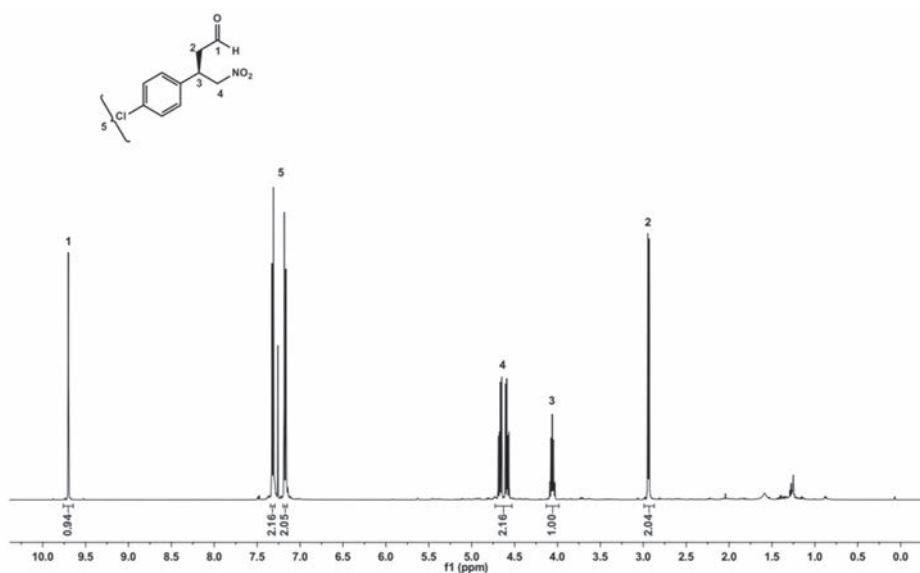


**<sup>1</sup>H NMR spectra of enzymatically obtained 3a-d**

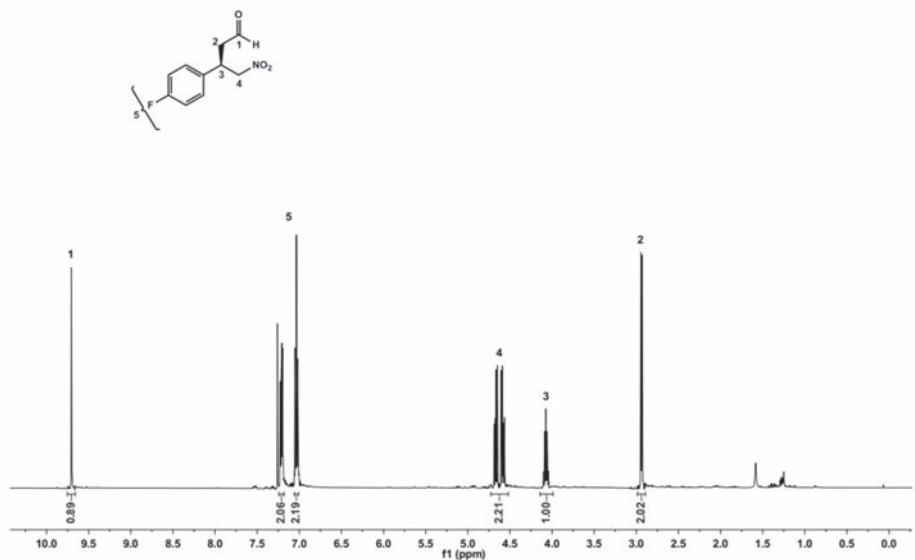
**Figure S6.** <sup>1</sup>H NMR spectrum of enzymatically prepared (*S*)-5-methyl-3-(nitromethyl)hexanal (**3a**) (400 MHz, CDCl<sub>3</sub>).



**Figure S7.** <sup>1</sup>H NMR spectrum of enzymatically prepared (*R*)-4-nitro-3-phenylbutanal (**3b**) (400 MHz, CDCl<sub>3</sub>).



**Figure S8.**  $^1\text{H}$  NMR spectrum of enzymatically prepared (*R*)-3-(4-chlorophenyl)-4-nitrobutanal (**3c**) (400 MHz,  $\text{CDCl}_3$ ).



**Figure S9.**  $^1\text{H}$  NMR spectrum of enzymatically prepared (*R*)-3-(4-fluorophenyl)-4-nitrobutanal (**3d**) (400 MHz,  $\text{CDCl}_3$ ).

### HPLC and GC chromatograms for enantiomeric excess determination of 3a-d

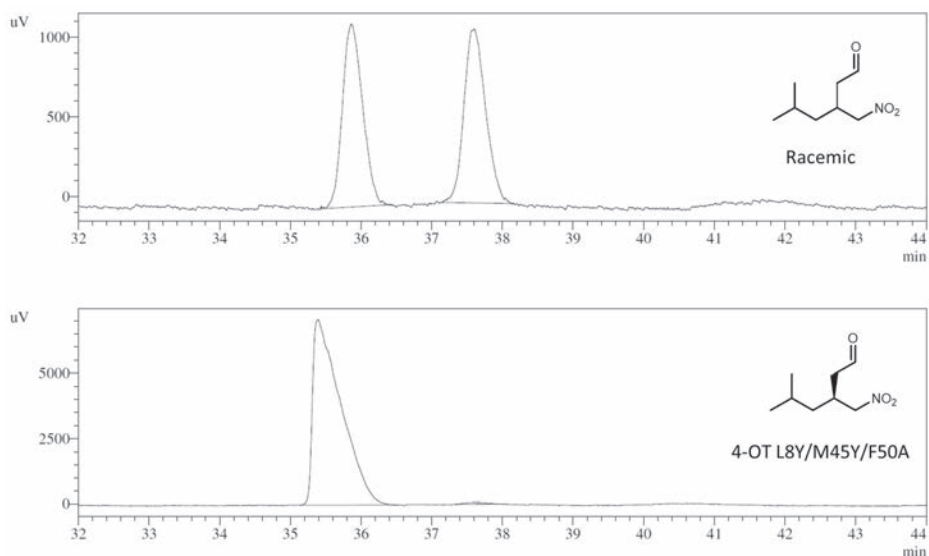


Figure S10. GC chromatograms of racemic and enzymatically obtained 3a.

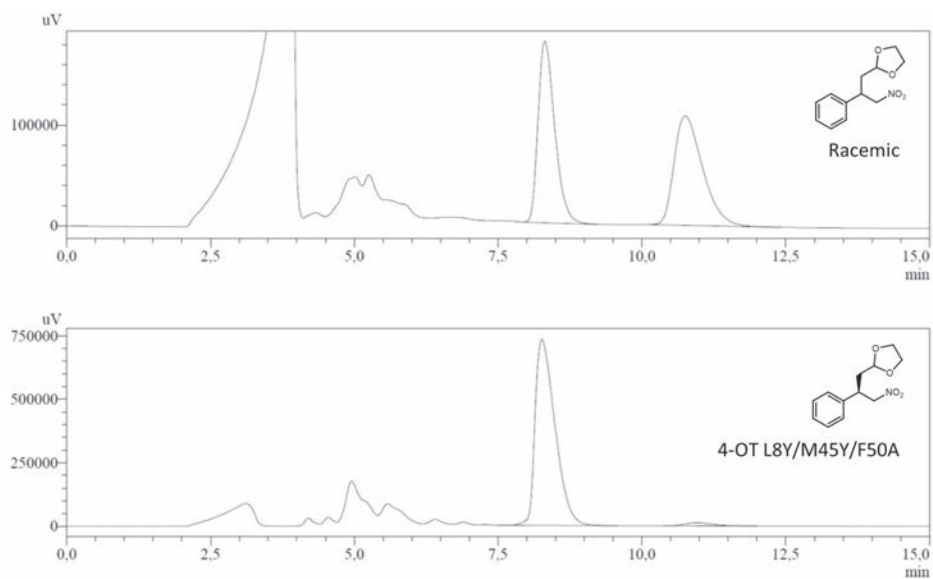


Figure S11. HPLC chromatograms of racemic and enzymatically obtained derivatized 3b.

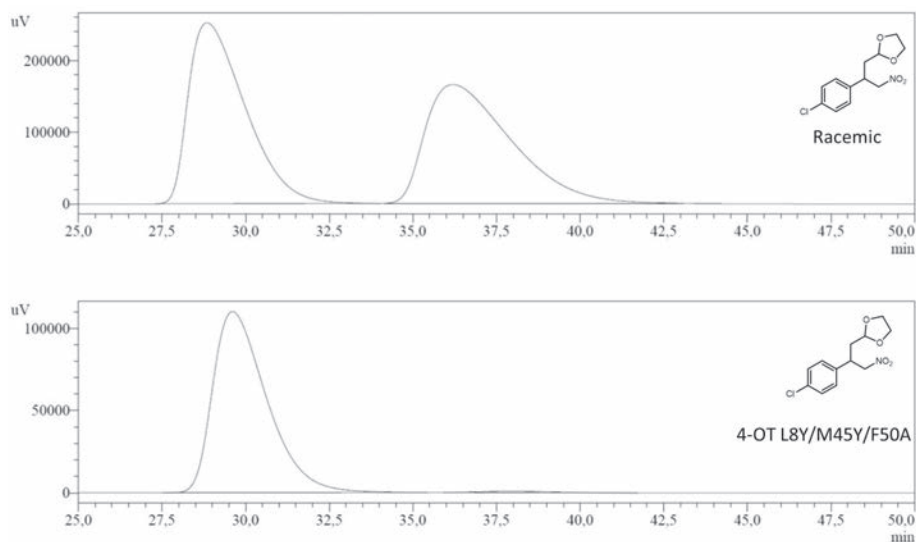


Figure S12. HPLC chromatograms of racemic and enzymatically obtained derivatized 3c.

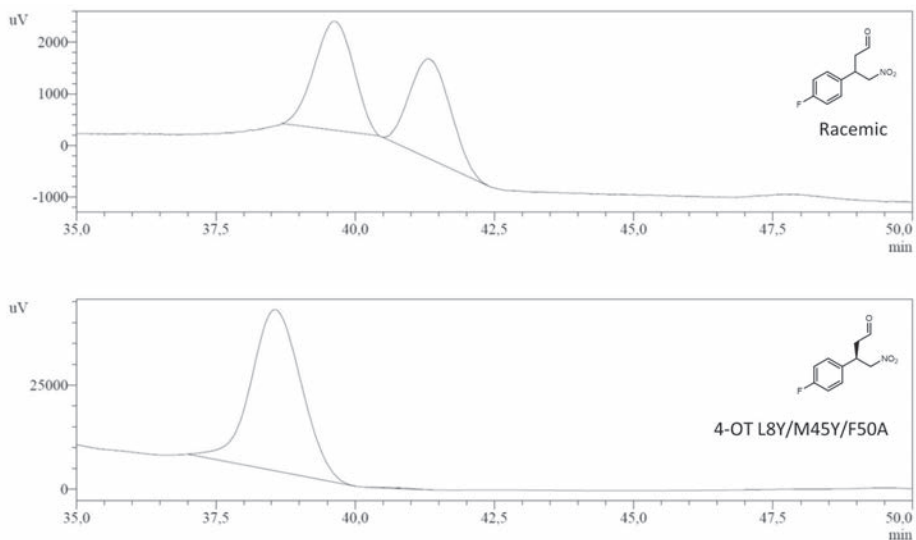
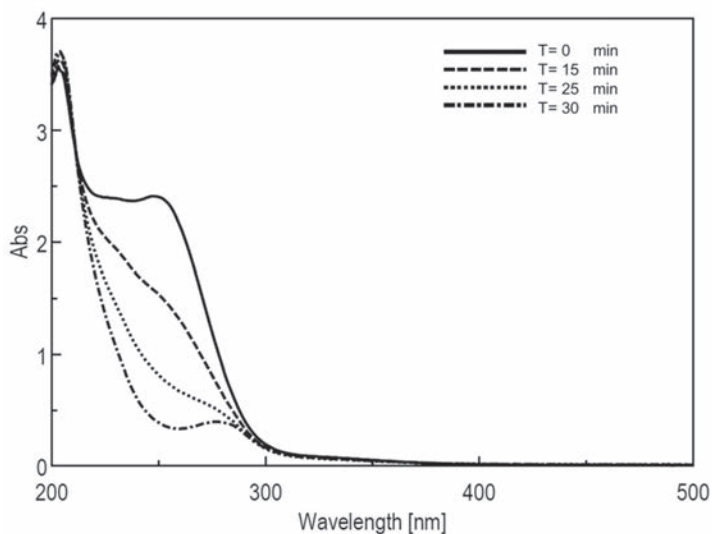
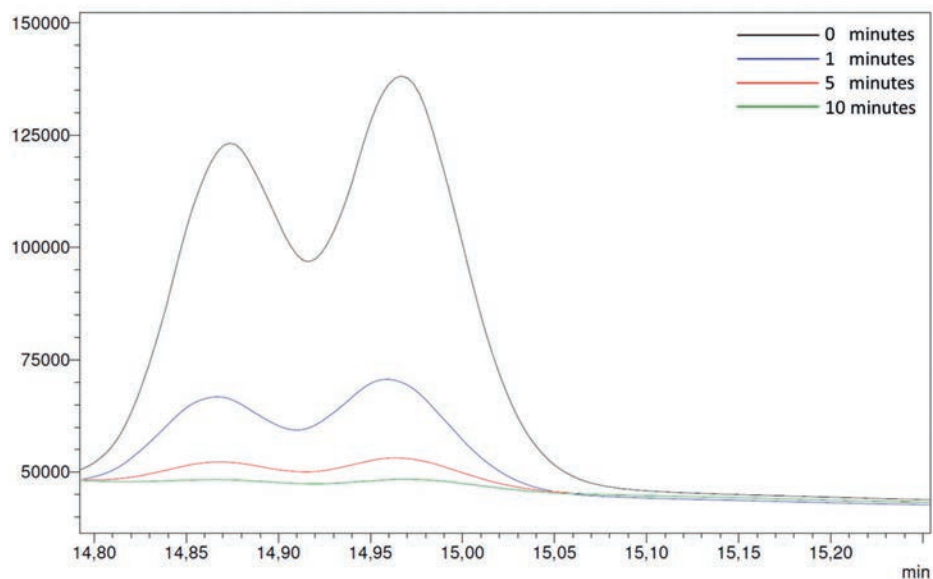


Figure S13. HPLC chromatograms of racemic and enzymatically obtained 3d.

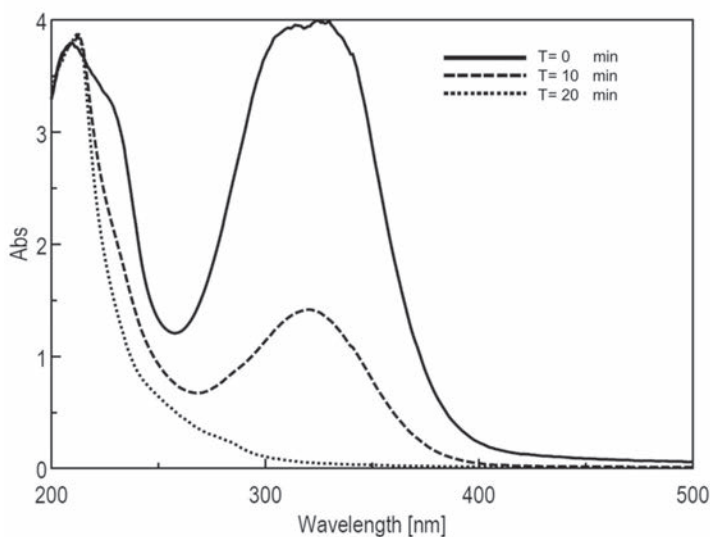
## UV spectra and HPLC chromatograms of the one-pot synthesis of 4a-d



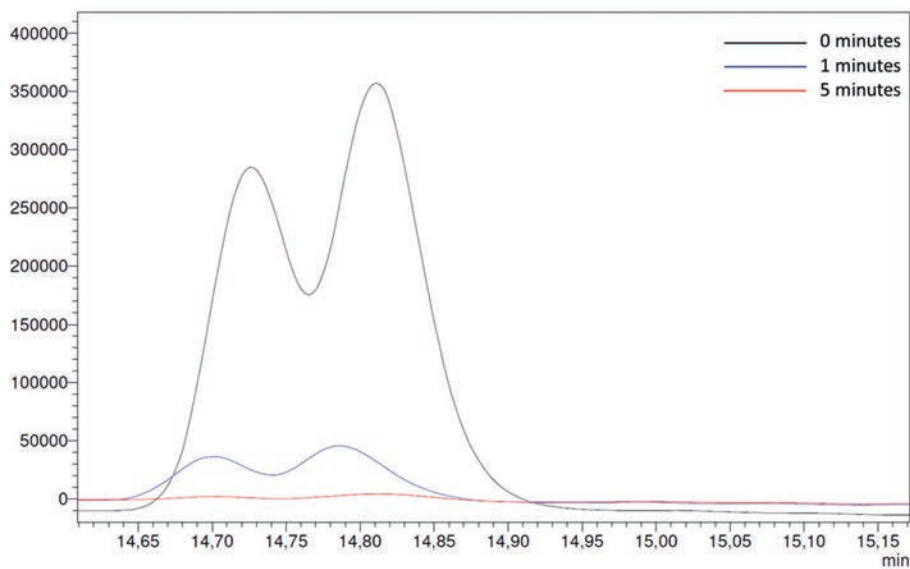
**Figure S14.** UV spectra showing the conversion of **2a** by 4-OT L8Y/M45Y/F50A in the first step of the enzymatic cascade.



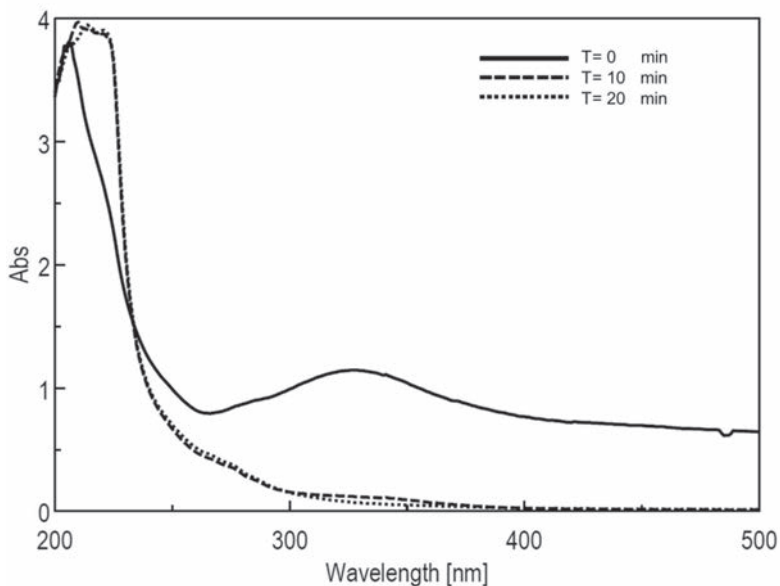
**Figure S15.** HPLC chromatograms showing the conversion of 4-OT L8Y/M45Y/F50A synthesized **3a** by PRO-ALDH(003) in the second step of the enzymatic cascade. The aldehyde functionality was derivatized to the corresponding (O)-benzyl oxime.



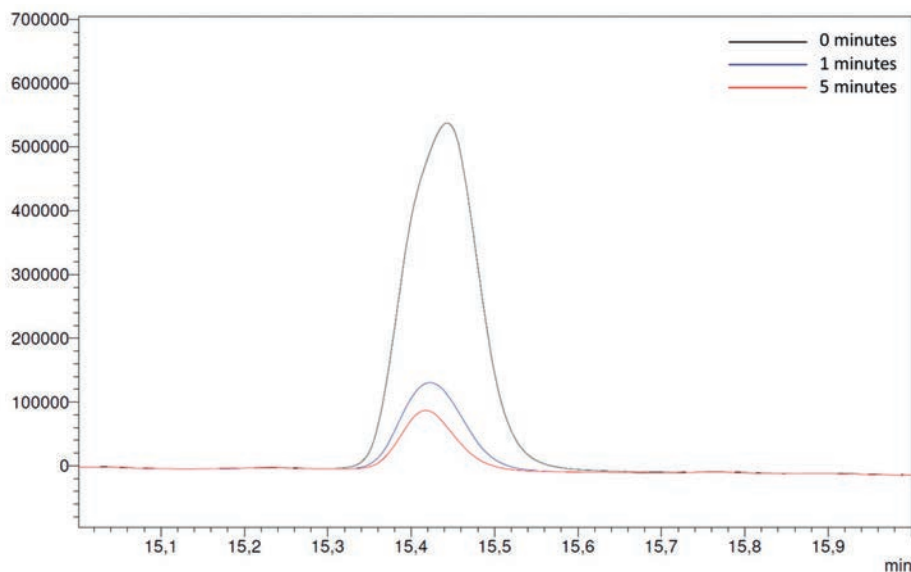
**Figure S16.** UV spectra showing the conversion of **2b** by 4-OT L8Y/M45Y/F50A in the first step of the enzymatic cascade.



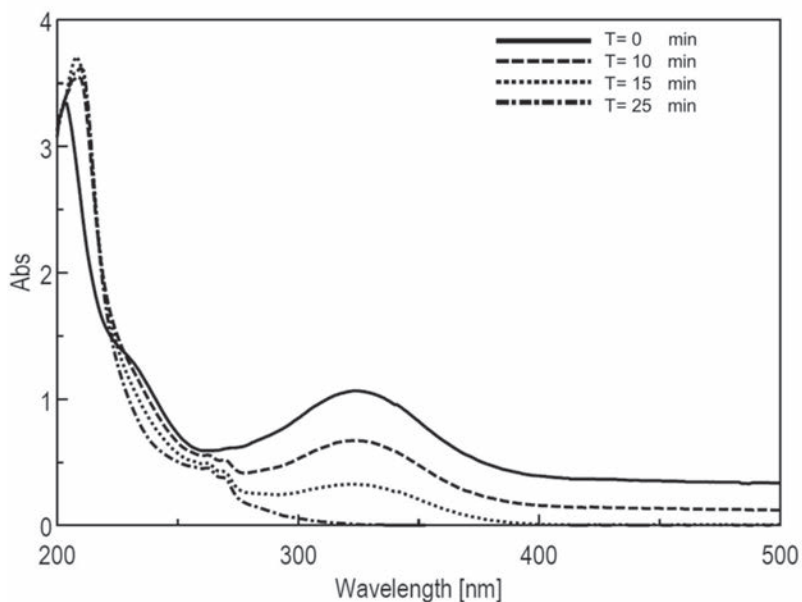
**Figure S17.** HPLC chromatograms showing the conversion of 4-OT L8Y/M45Y/F50A synthesized **3b** by PRO-ALDH(003) in the second step of the enzymatic cascade. The aldehyde functionality was derivatized to the corresponding (O)-benzyl oxime.



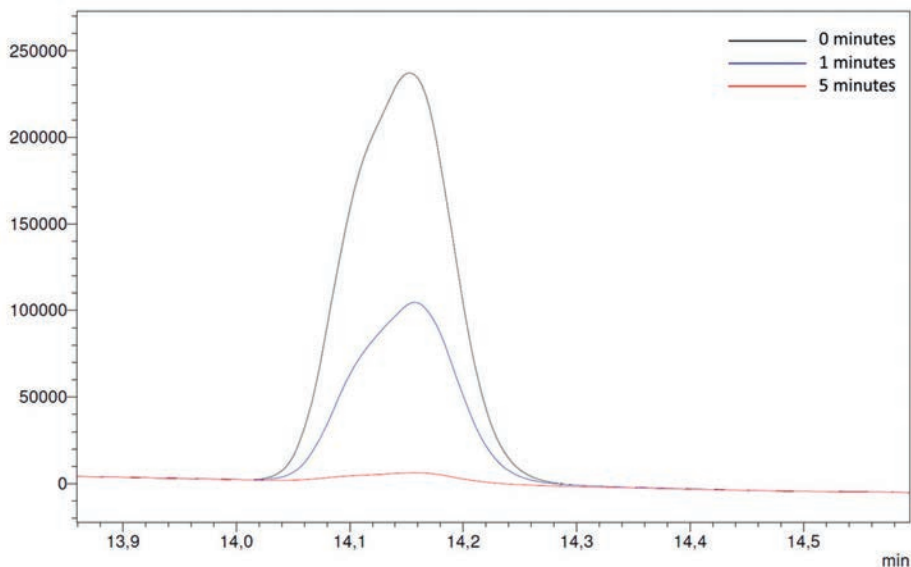
**Figure S18.** UV spectra showing the conversion of **2c** by 4-OT L8Y/M45Y/F50A in the first step of the enzymatic cascade.



**Figure S19.** HPLC chromatograms showing the conversion of 4-OT L8Y/M45Y/F50A synthesized **3c** by PRO-ALDH(003) in the second step of the enzymatic cascade. The aldehyde functionality was derivatized to the corresponding (O)-benzyl oxime.

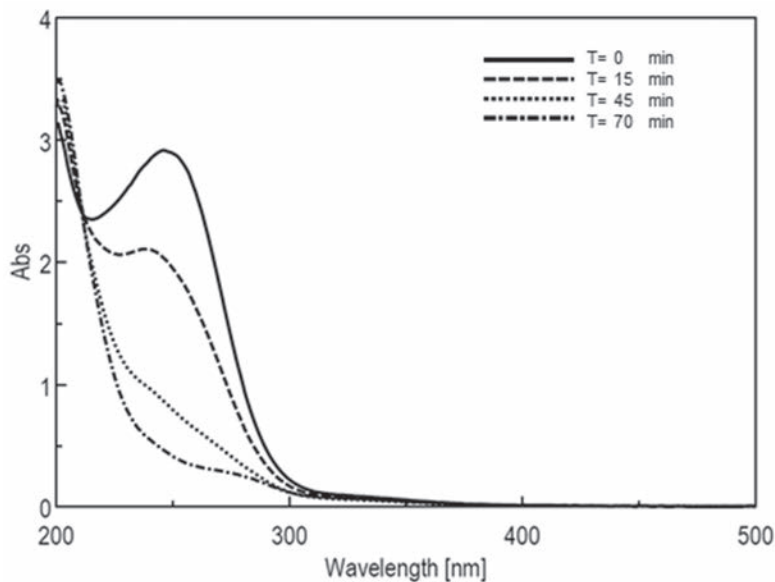


**Figure S20.** UV spectra showing the conversion of **2d** by 4-OT L8Y/M45Y/F50A in the first step of the enzymatic cascade.

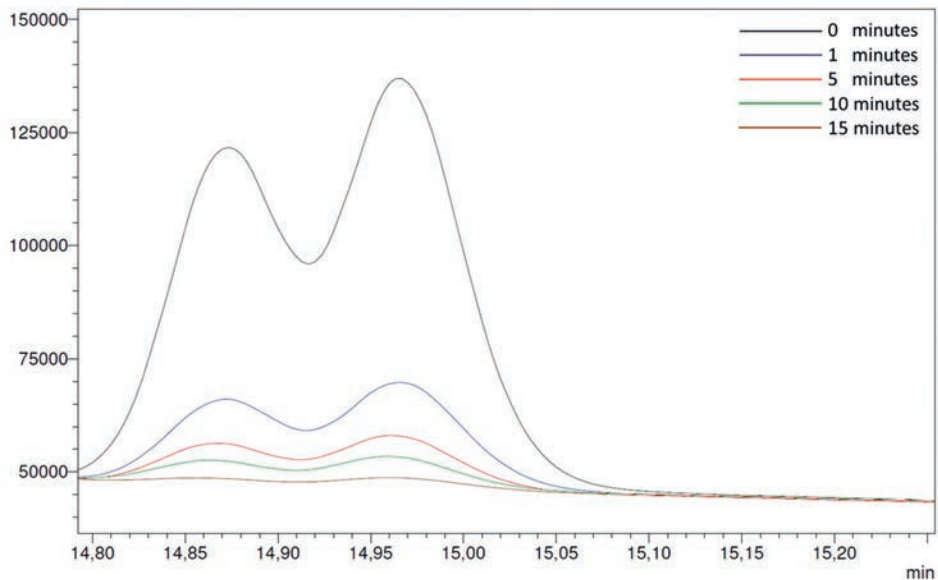


**Figure S21.** HPLC chromatograms showing the conversion of 4-OT L8Y/M45Y/F50A synthesized **3d** by PRO-ALDH(003) in the second step of the enzymatic cascade. The aldehyde functionality was derivatized to the corresponding (O)-benzyl oxime.

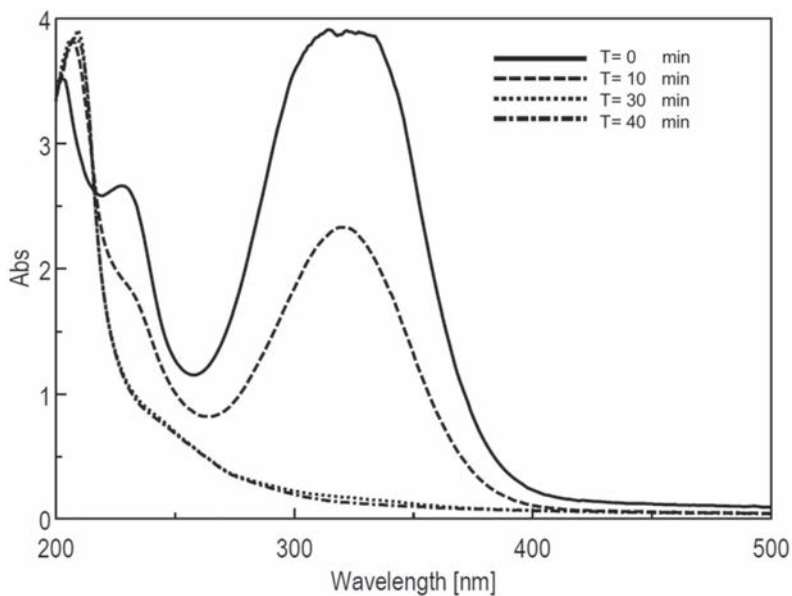




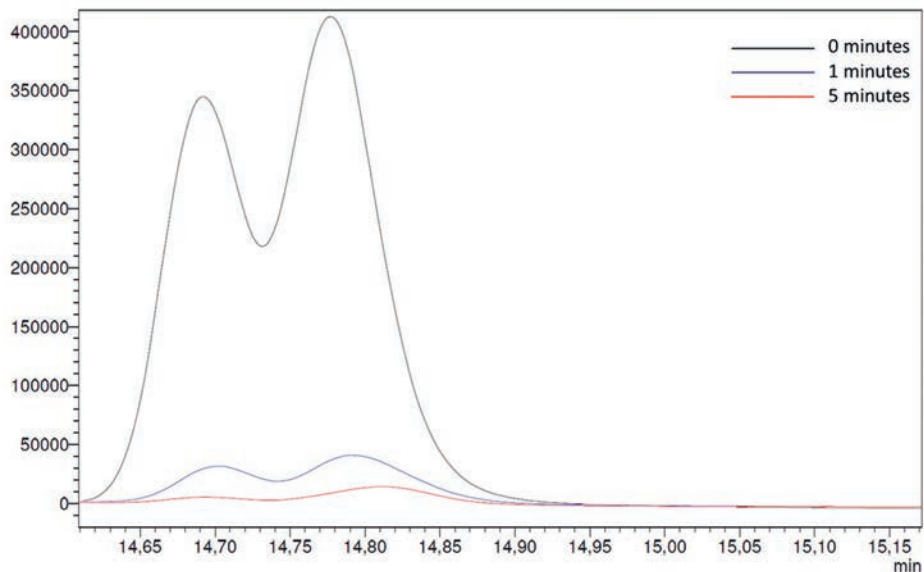
**Figure S22.** UV spectra showing the conversion of **2a** by 4-OT A33D in the first step of the enzymatic cascade.



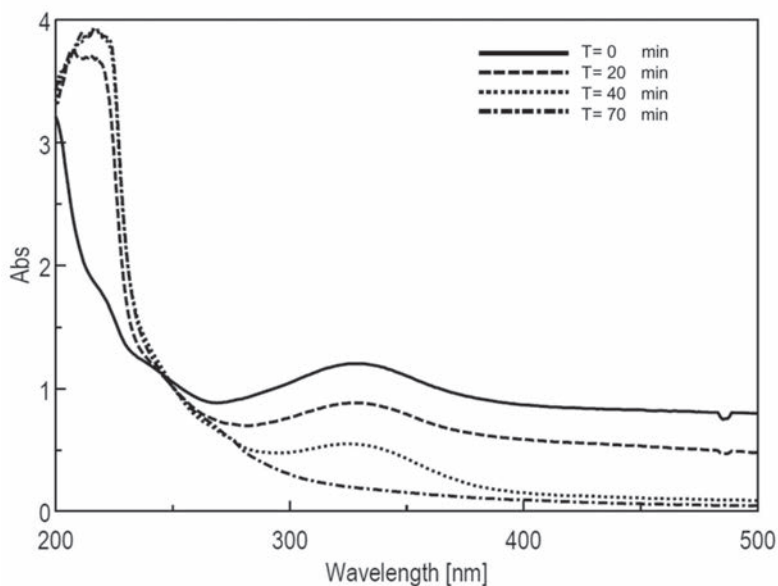
**Figure S23.** HPLC chromatograms showing the conversion of 4-OT A33D synthesized **3a** by PRO-ALDH(003) in the second step of the enzymatic cascade. The aldehyde functionality was derivatized to the corresponding (O)-benzyl oxime.



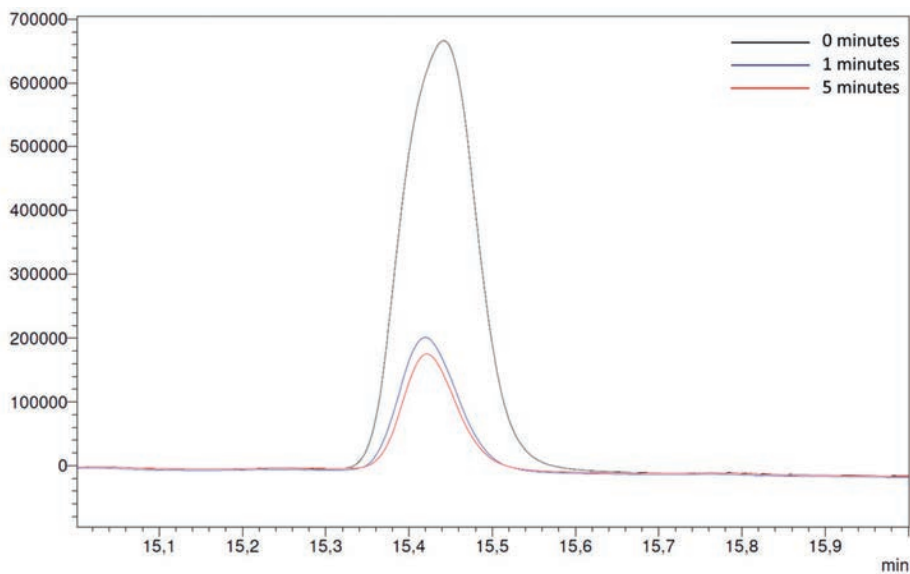
**Figure S24.** UV spectra showing the conversion of **2b** by 4-OT A33D in the first step of the enzymatic cascade.



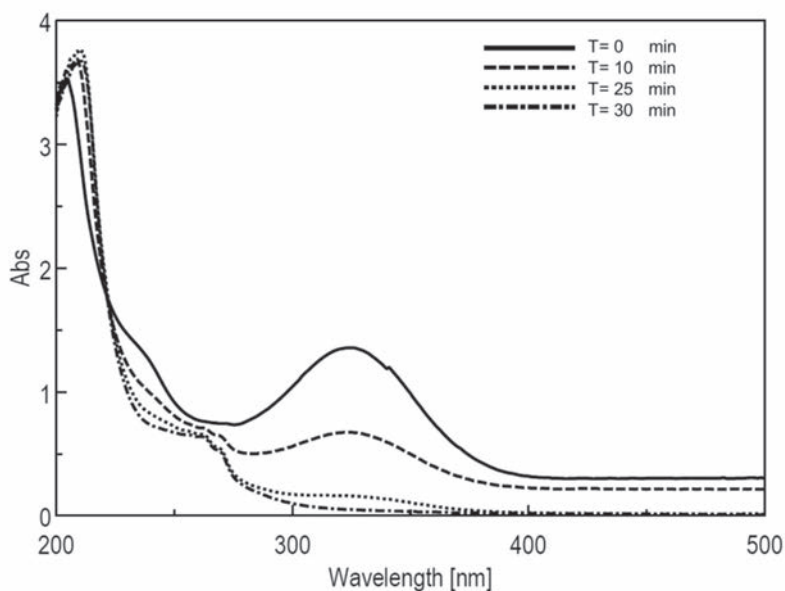
**Figure S25.** HPLC chromatograms showing the conversion of 4-OT A33D synthesized **3b** by PRO-ALDH(003) in the second step of the enzymatic cascade. The aldehyde functionality was derivatized to the corresponding (O)-benzyl oxime.



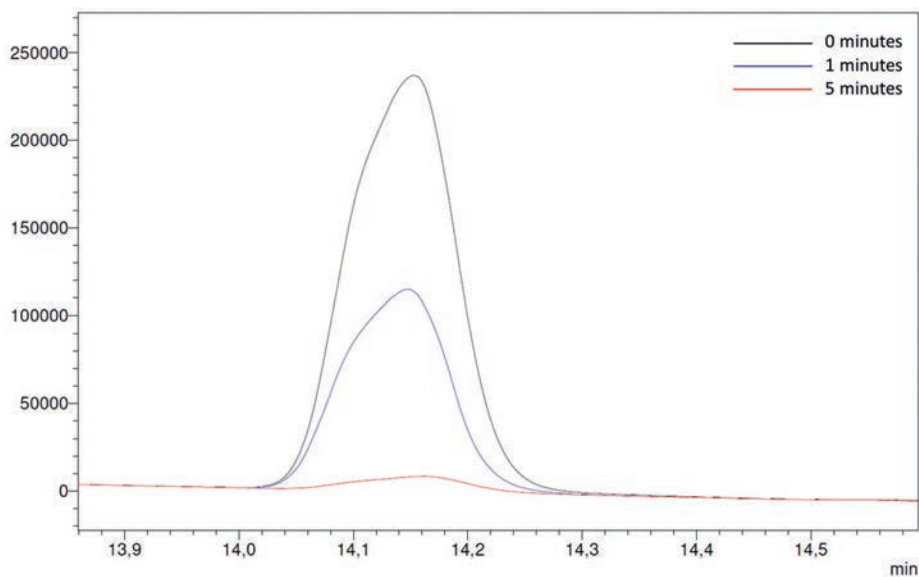
**Figure S26.** UV spectra showing the conversion of **2c** by 4-OT A33D in the first step of the enzymatic cascade.



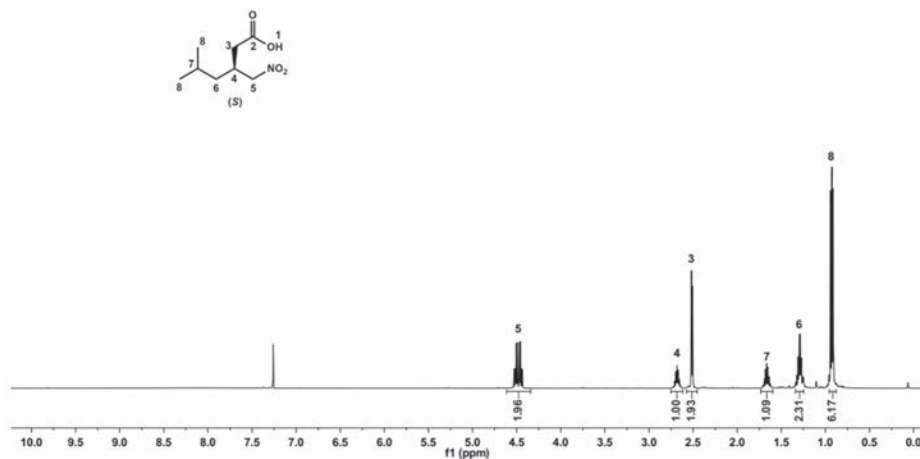
**Figure S27.** HPLC chromatograms showing the conversion of 4-OT A33D synthesized **3c** by PRO-ALDH(003) in the second step of the enzymatic cascade. The aldehyde functionality was derivatized to the corresponding (O)-benzyl oxime.



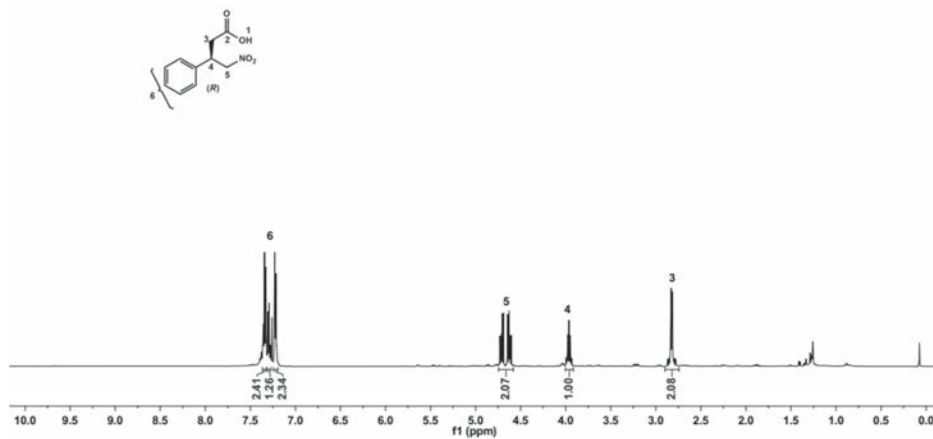
**Figure S28.** UV spectra showing the conversion of **2d** by 4-OT A33D in the first step of the enzymatic cascade.



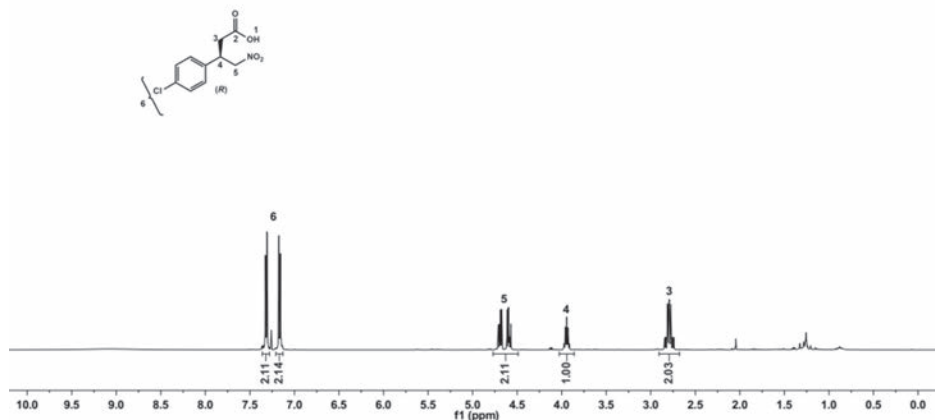
**Figure S29.** HPLC chromatograms showing the conversion of 4-OT A33D synthesized **3d** by PRO-ALDH(003) in the second step of the enzymatic cascade. The aldehyde functionality was derivatized to the corresponding (O)-benzyl oxime.

$^1\text{H}$  NMR spectra of enzymatically obtained 4a-d

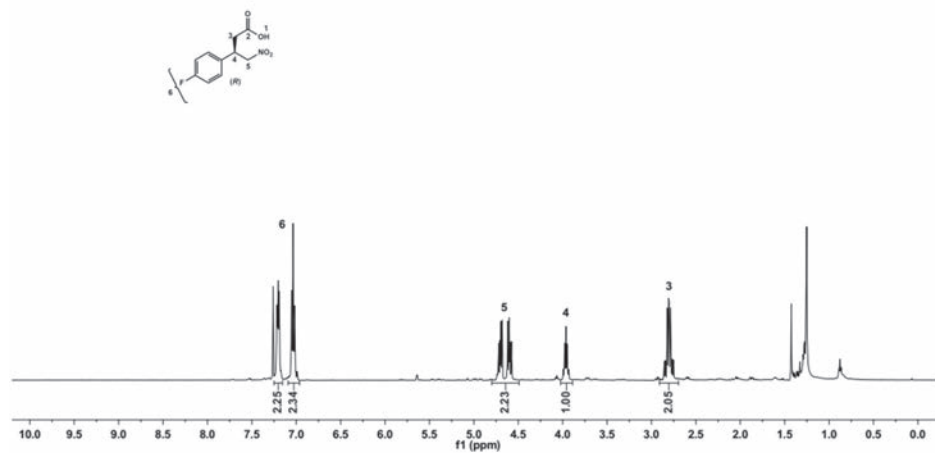
**Figure S30.**  $^1\text{H}$  NMR spectrum of (S)-5-methyl-3-(nitromethyl)hexanoic acid (**4a**) obtained by 4-OT L8Y/M45Y/F50A and PRO-ALDH(003) (400 MHz,  $\text{CDCl}_3$ ).



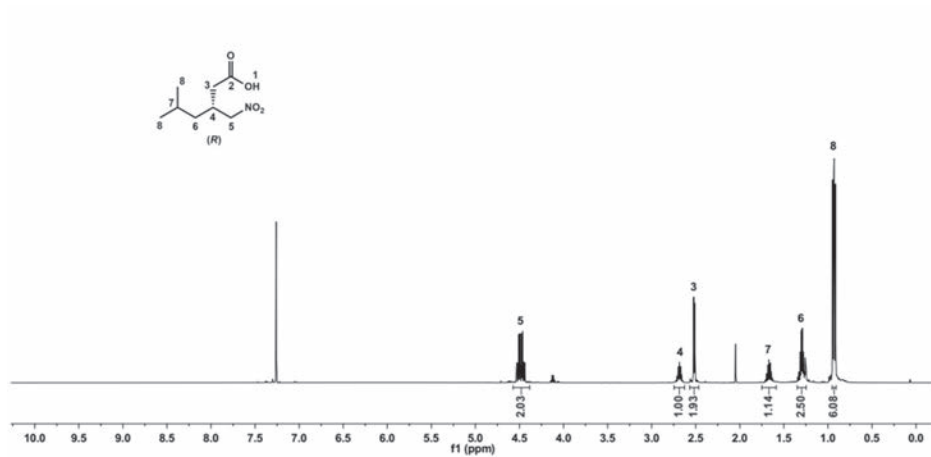
**Figure S31.**  $^1\text{H}$  NMR spectrum of (R)-4-nitro-3-phenylbutanoic acid (**4b**) obtained by 4-OT L8Y/M45Y/F50A and PRO-ALDH(003) (400 MHz,  $\text{CDCl}_3$ ).



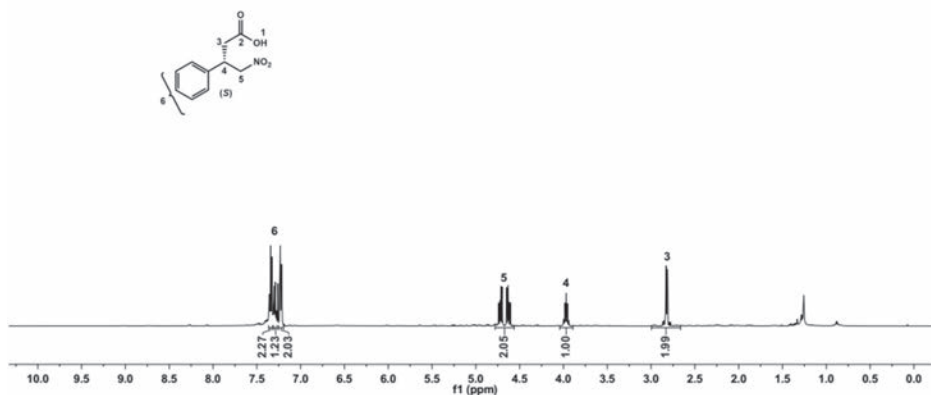
**Figure S32.**  $^1\text{H}$  NMR spectrum of  $(R)$ -3-(4-chlorophenyl)-4-nitrobutanoic acid (**4c**) obtained by 4-OT L8Y/M45Y/F50A and PRO-ALDH(003) (400 MHz,  $\text{CDCl}_3$ ).



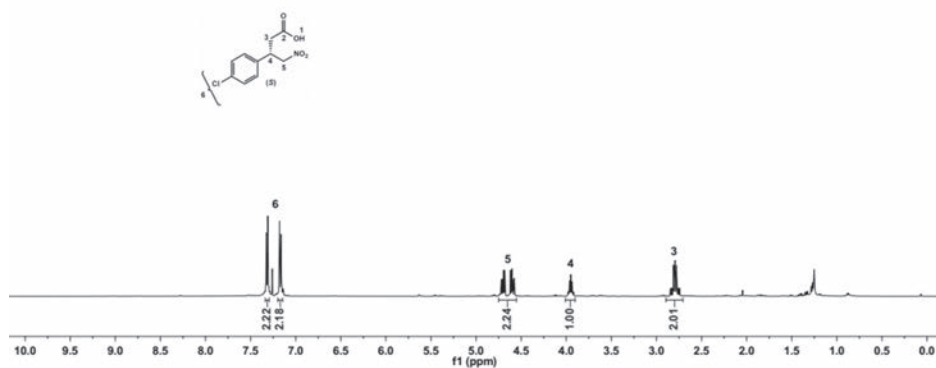
**Figure S33.**  $^1\text{H}$  NMR spectrum of  $(R)$ -3-(4-fluorophenyl)-4-nitrobutanoic acid (**4d**) obtained by 4-OT L8Y/M45Y/F50A and PRO-ALDH(003) (400 MHz,  $\text{CDCl}_3$ ).



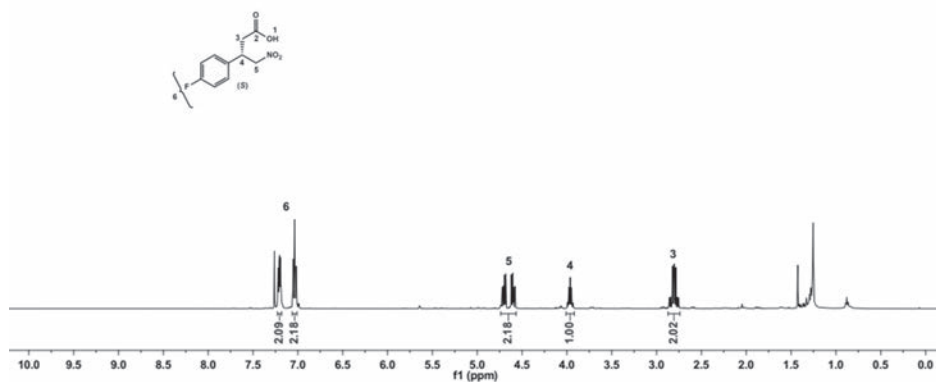
**Figure S34.**  $^1\text{H}$  NMR spectrum of (*R*)-5-methyl-3-(nitromethyl)hexanoic acid (**4a**) obtained by 4-OT A33D and PRO-ALDH(003) (400 MHz,  $\text{CDCl}_3$ ).



**Figure S35.**  $^1\text{H}$  NMR spectrum of (*S*)-4-nitro-3-phenylbutanoic acid (**4b**) obtained by 4-OT A33D and PRO-ALDH(003) (400 MHz,  $\text{CDCl}_3$ ).



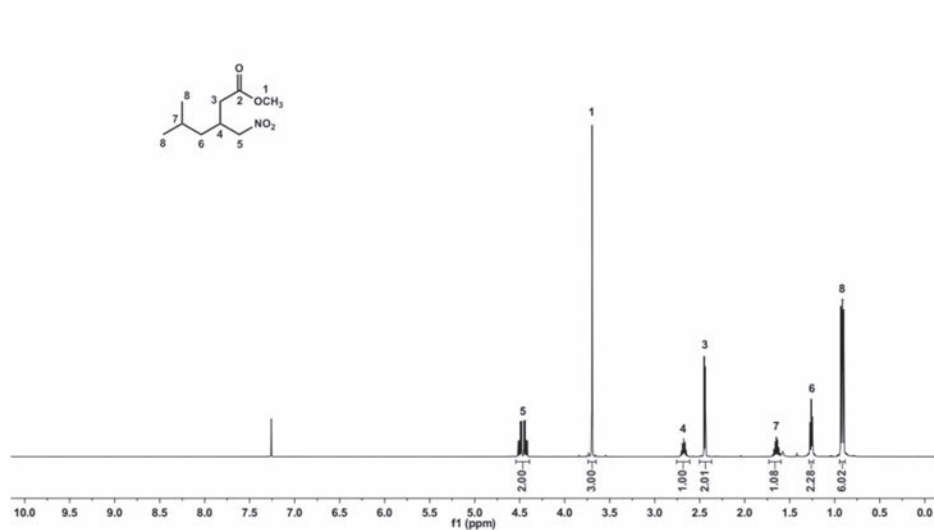
**Figure S36.**  $^1\text{H}$  NMR spectrum of (S)-3-(4-chlorophenyl)-4-nitrobutanoic acid (**4c**) obtained by 4-OT A33D and PRO-ALDH(003) (400 MHz,  $\text{CDCl}_3$ ).



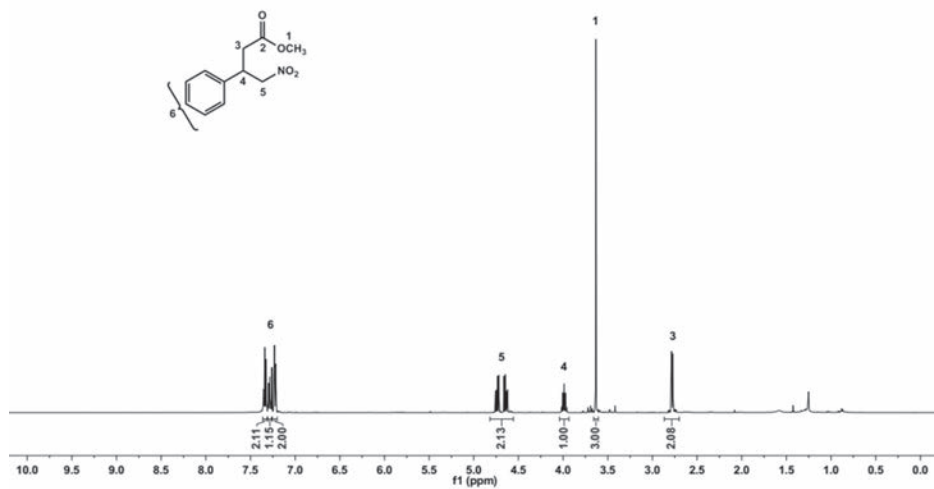
**Figure S37.**  $^1\text{H}$  NMR spectrum of (S)-3-(4-fluorophenyl)-4-nitrobutanoic acid (**4d**) obtained by 4-OT A33D and PRO-ALDH(003) (400 MHz,  $\text{CDCl}_3$ ).

2

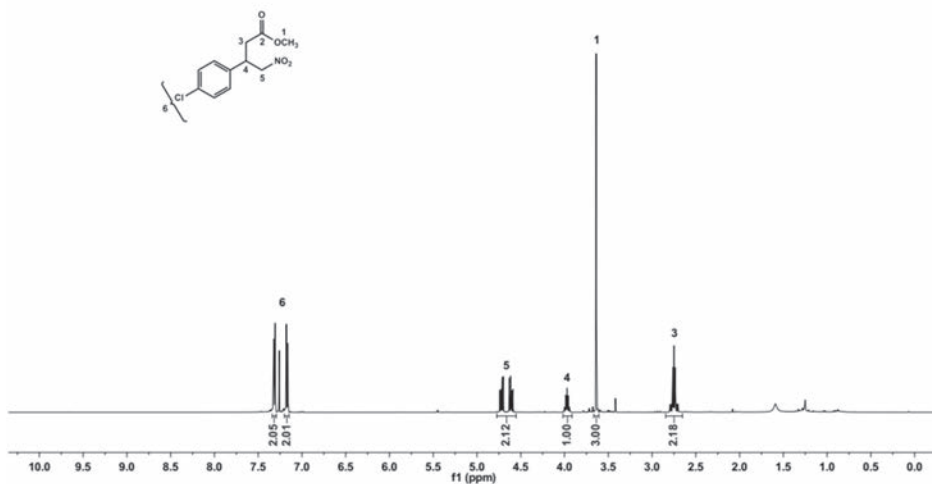


$^1\text{H}$  NMR spectra of racemic derivatized 4a-d

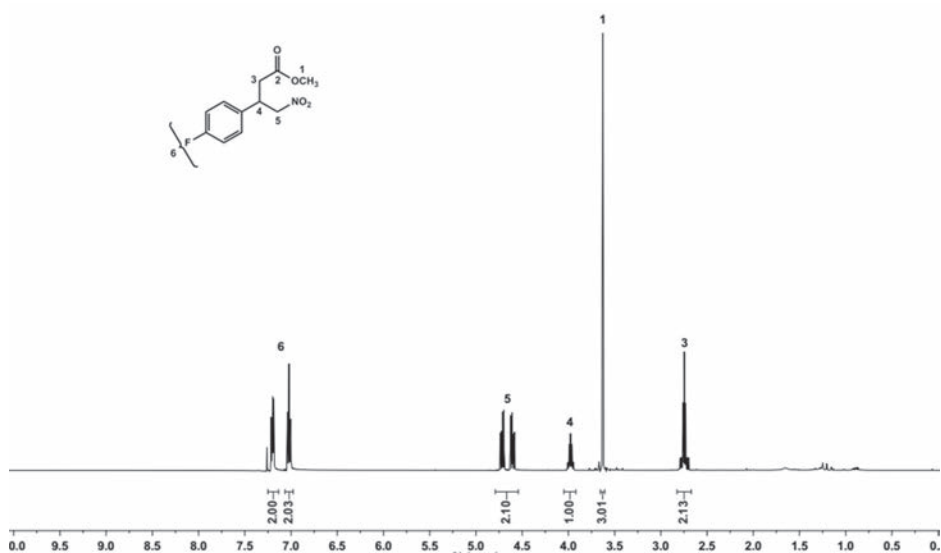
**Figure S38.**  $^1\text{H}$  NMR spectrum of racemic methyl 5-methyl-3-(nitromethyl)hexanoate (400 MHz,  $\text{CDCl}_3$ ).



**Figure S39.**  $^1\text{H}$  NMR spectrum of racemic methyl 4-nitro-3-phenylbutanoate (400 MHz,  $\text{CDCl}_3$ ).



**Figure S40.**  $^1\text{H}$  NMR spectrum of racemic methyl 3-(4-chlorophenyl)-4-nitrobutanoate (400 MHz,  $\text{CDCl}_3$ ).



**Figure S41.**  $^1\text{H}$  NMR spectrum of racemic methyl 3-(4-fluorophenyl)-4-nitrobutanoate (400 MHz,  $\text{CDCl}_3$ ).

HPLC and GC chromatograms for enantiomeric excess determination of derivatized 4a-d

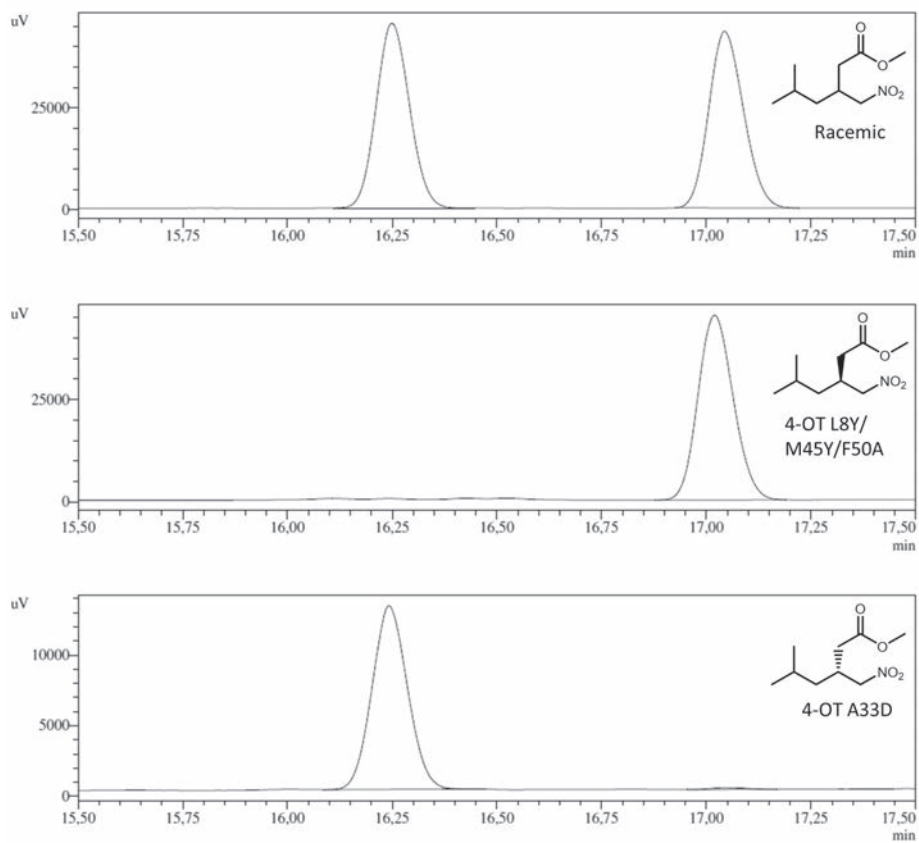


Figure S42. GC chromatograms of racemic and enzymatically obtained derivatized 4a.

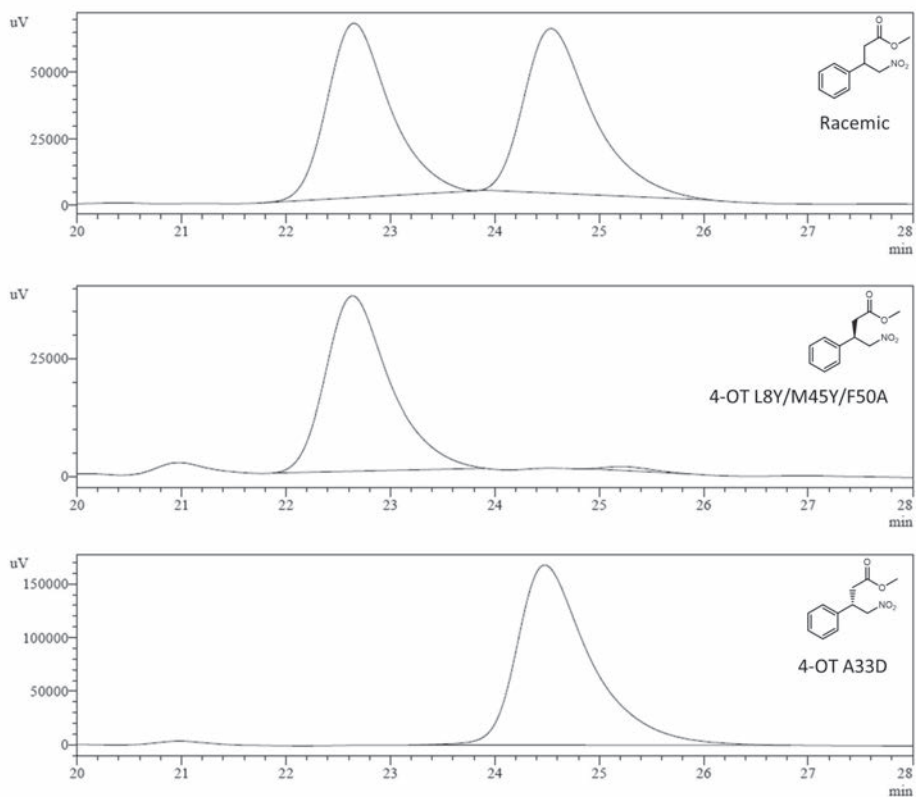
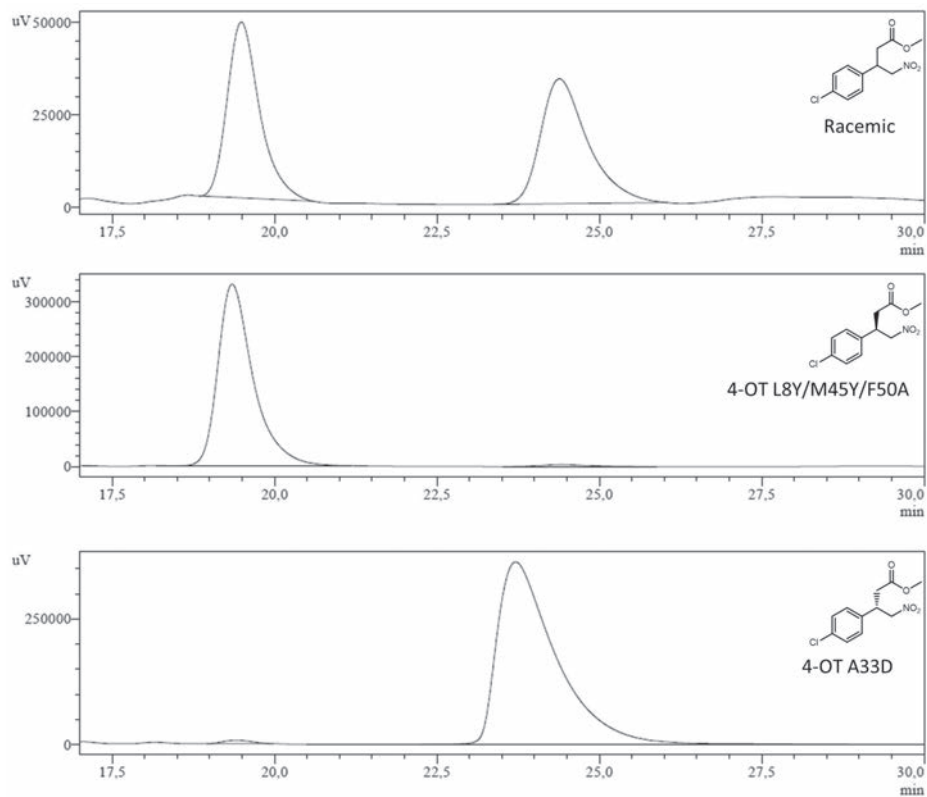


Figure S43. HPLC chromatograms of racemic or enzymatically obtained derivatized **4b**.

2

Chapter 2



**Figure S44.** HPLC chromatograms of racemic and enzymatically obtained derivatized 4c.

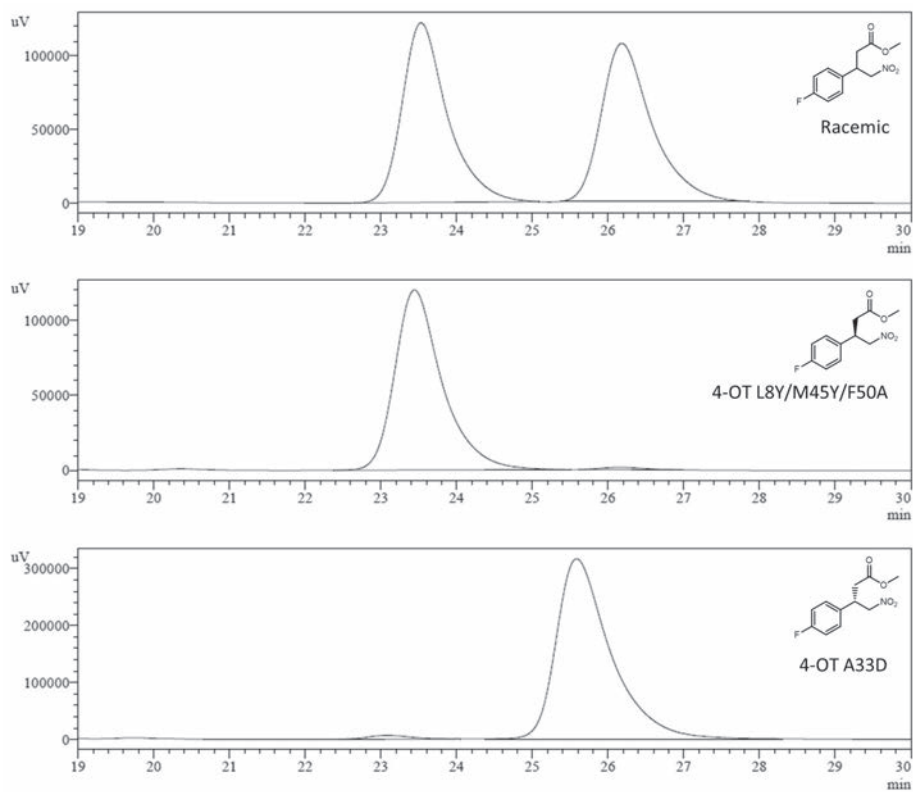
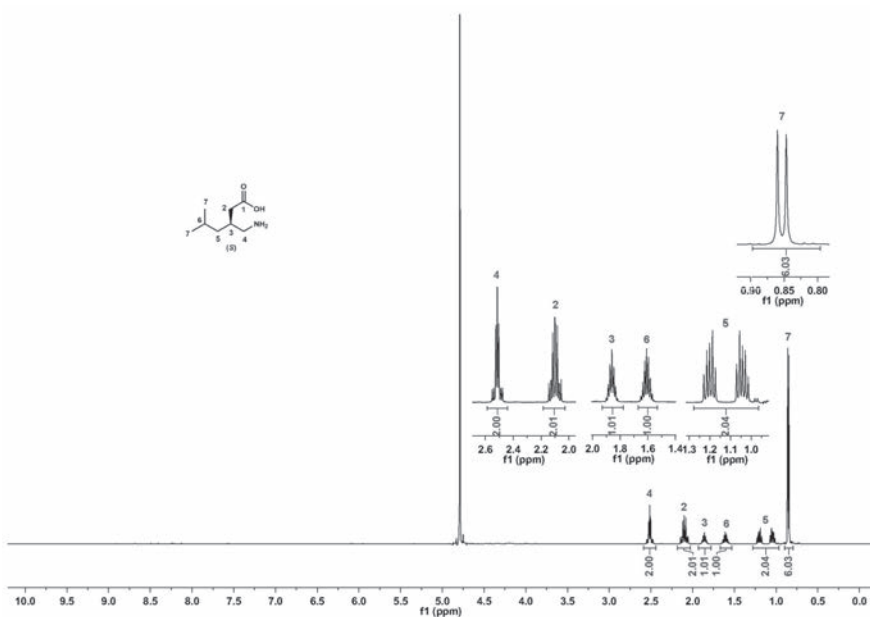
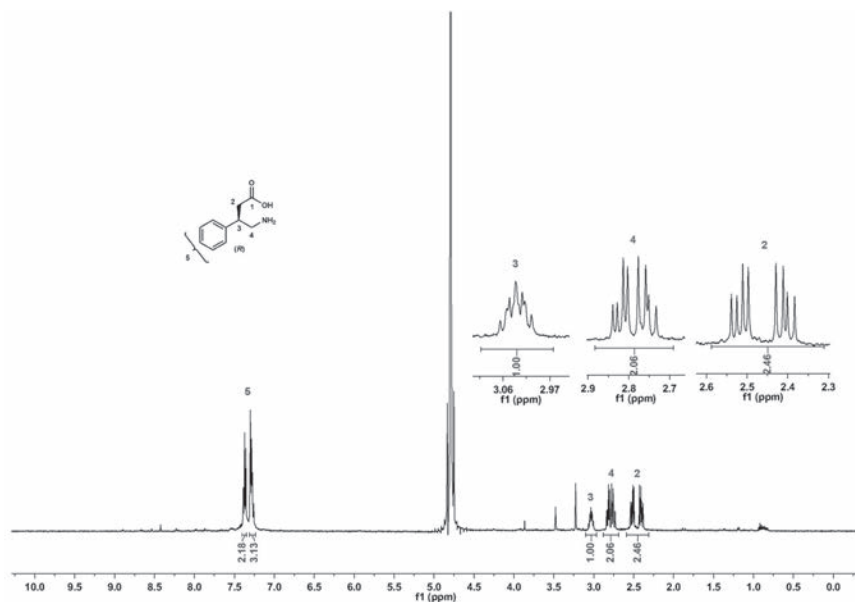


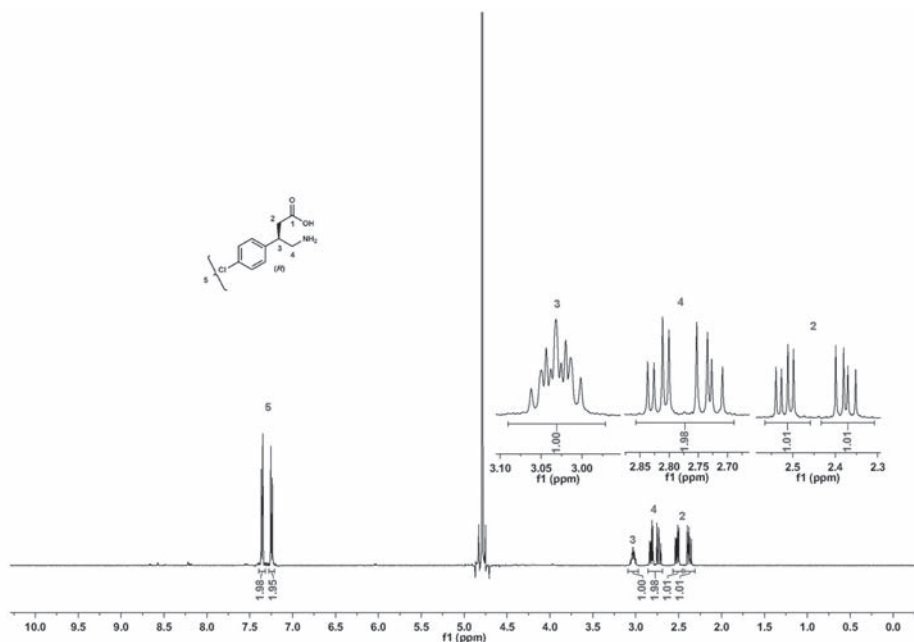
Figure S45. HPLC chromatograms of racemic and enzymatically obtained derivatized **4d**.

$^1\text{H}$  NMR spectra of chemoenzymatically obtained 5a-d

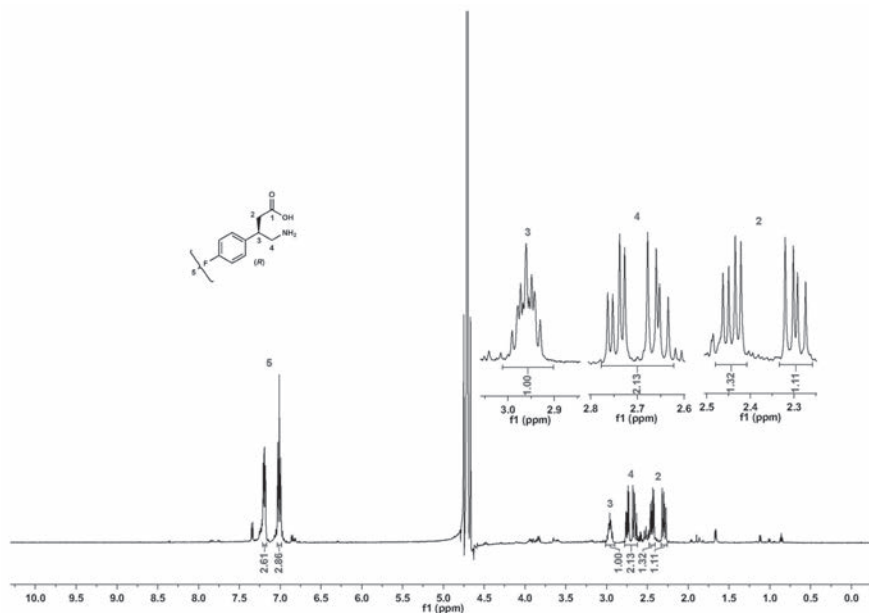
**Figure S46.**  $^1\text{H}$  NMR spectrum of chemoenzymatically prepared (*S*)-Pregabalin (**5a**) (400 MHz,  $\text{D}_2\text{O}$  and NaOD).



**Figure S47.**  $^1\text{H}$  NMR spectrum of chemoenzymatically prepared (*R*)-Phenibut (**5b**) (400 MHz,  $\text{D}_2\text{O}$  and NaOD).



**Figure S48.**  $^1\text{H}$  NMR spectrum of chemoenzymatically prepared (*R*)-Baclofen (**5c**) (400 MHz,  $\text{D}_2\text{O}$  and NaOD).



**Figure S49.**  $^1\text{H}$  NMR spectrum of chemoenzymatically prepared fluorophenibut (**5d**) (400 MHz,  $\text{D}_2\text{O}$  and NaOD).



HPLC chromatograms for enantiomeric excess determination of derivatized 5a-d

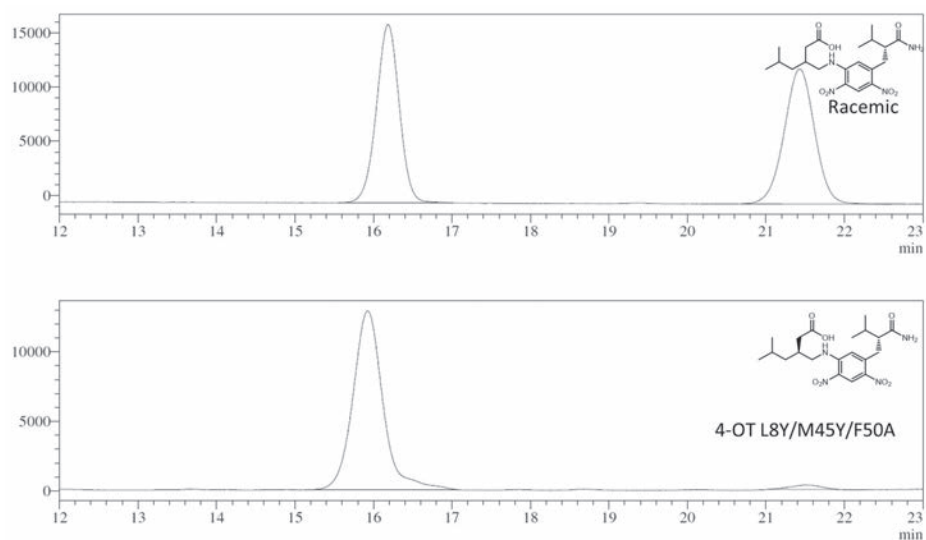


Figure S50. HPLC chromatograms of racemic and chemoenzymatically obtained derivatized 5a.

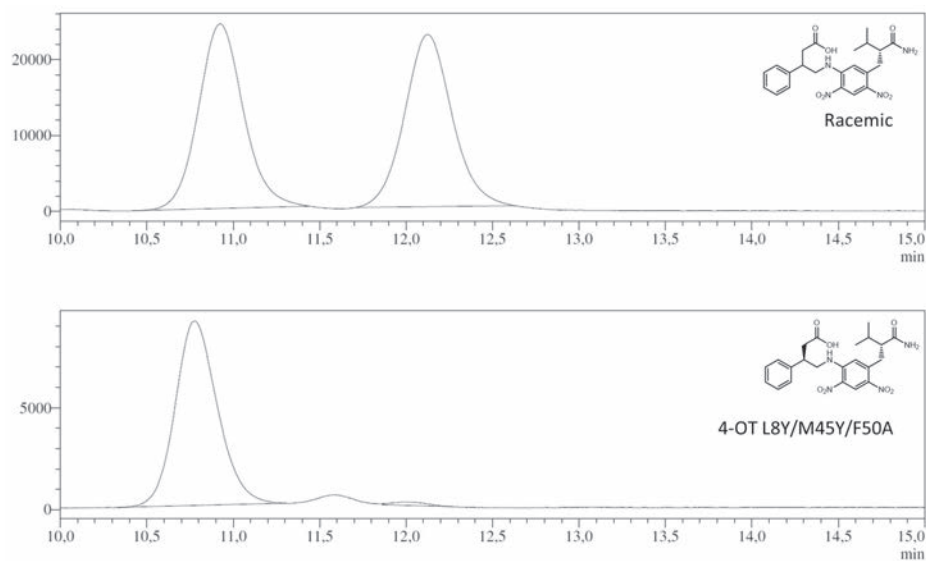


Figure S51. HPLC chromatograms of racemic and chemoenzymatically obtained derivatized 5b.

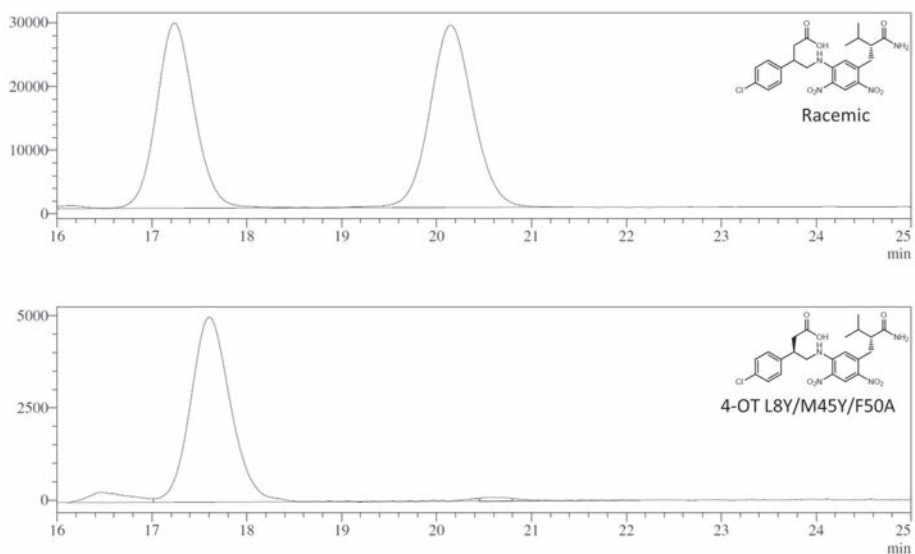


Figure S52. HPLC chromatograms of racemic and chemoenzymatically obtained derivatized 5c.

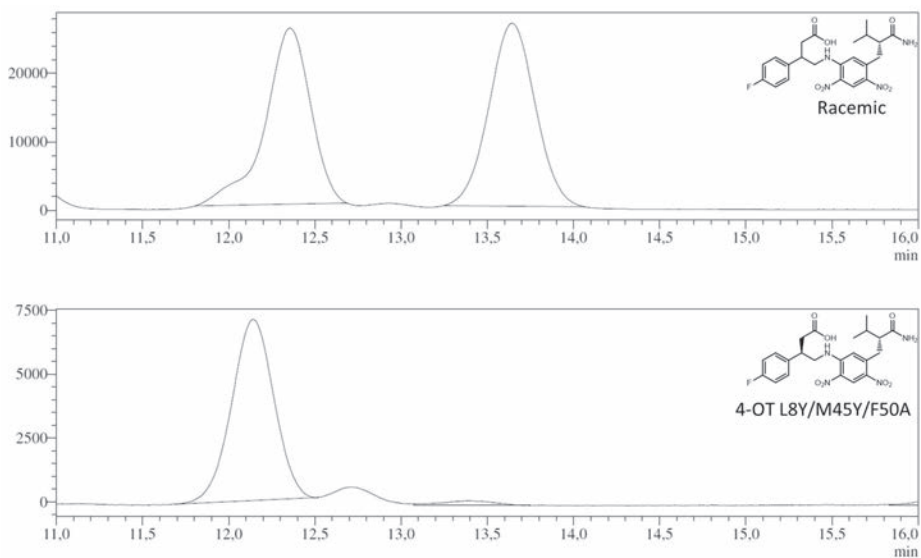


Figure S53. HPLC chromatograms of racemic and chemoenzymatically obtained derivatized 5d.

## References

1. van der Meer, J.; Poddar, H.; Baas, B.; Miao, Y.; Rahimi, M.; Kunzendorf, A.; van Merkerk, R.; Tepper, P. G.; Geertsema, E. M.; Thunnissen, A.M.W.H.; Quax, W.J.; Poelarends, G.J. Using Mutability Landscapes of a Promiscuous Tautomerase to Guide the Engineering of Enantioselective Michaelases. *Nat. Commun.* **2016**, *7*, 10911.
2. Zandvoort, E.; Baas, B.; Quax, W. J.; Poelarends, G. J. Systematic Screening for Catalytic Promiscuity in 4-Oxalocrotonate Tautomerase: Enamine Formation and Aldolase Activity. *ChemBioChem* **2011**, *12*, 602-609.
3. Waddell, W. J. A Simple Ultraviolet Spectrophotometric Method for the Determination of Protein. *J. Lab. Clin. Med.* **1956**, *48*, 311-314.
4. Geertsema, E. M.; Miao, Y.; Tepper, P. G.; de Haan, P.; Zandvoort, E.; Poelarends, G. J. Biocatalytic Michael-Type Additions of Acetaldehyde to Nitroolefins with the Proline-Based Enzyme 4-Oxalocrotonate Tautomerase Yielding Enantioenriched  $\gamma$ -Nitroaldehydes. *Chem. Eur. J.* **2013**, *19*, 14407-14410.
5. Kabsch, W. Integration, Scaling, Space-Group Assignment and Post-Refinement. *Acta Crystallogr., Sect. D* **2010**, *66*, 133-144.
6. McCoy, A. J.; Grosse-Kunstleve, R. W.; Adams, P. D.; Winn, M. D.; Storoni, L. C.; Read, R. J. Phaser Crystallographic Software. *J. Appl. Crystallogr.* **2007**, *40*, 658-674.
7. Poddar, H.; Rahimi, M.; Geertsema, E. M.; Thunnissen, A.M.W.H.; Poelarends, G. J. Evidence for the Formation of an Enamine Species During Aldol and Michael-type Addition Reactions Promiscuously Catalyzed by 4-Oxalocrotonate Tautomerase. *ChemBioChem* **2015**, *16*, 738-741.
8. Murshudov, G. N.; Vagin, A. A.; Dodson, E. J. Refinement of Macromolecular Structures by the Maximum-Likelihood Method. *Acta Crystallogr., Sect. D* **1997**, *53*, 240-255.
9. Emsley, P.; Lohkamp, B.; Scott, W. G.; Cowtan, K. Features and Development of Coot. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2010**, *66*, 486-501.
10. Li, J.; Lear, M. J.; Kawamoto, Y.; Umemiya, S.; Wong, A. R.; Kwon, E.; Sato, I.; Hayashi, Y. Oxidative Amidation of Nitroalkanes with Amine Nucleophiles Using Molecular Oxygen and Iodine. *Angew. Chem., Int. Ed.* **2015**, *127*, 13178-13182.
11. D'Oca, C. R. M.; Naciuk, F. F.; Silva, J. C.; Guedes, E. P.; Moro, C. C.; D'Oca, M. G. M.; Santos, L. S.; Natchigall, F. M.; Russowsky, D. New Multicomponent Reaction for the Direct Synthesis of  $\beta$ -Aryl- $\gamma$ -nitroesters Promoted by Hydrotalcite-Derived Mixed Oxides as Heterogeneous Catalyst. *J. Braz. Chem. Soc.* **2017**, *28*, 285-298.
12. Zandvoort, E.; Geertsema, E. M.; Baas, B.; Quax, W. J.; Poelarends, G. J. Bridging Between Organocatalysis and Biocatalysis: Asymmetric Addition of Acetaldehyde to  $\beta$ -Nitrostyrenes Catalyzed by a Promiscuous Proline-Based Tautomerase. *Angew. Chem., Int. Ed.* **2012**, *124*, 1266-1269.

13. Szekrenyi, A.; Garrabou, X.; Parella, T.; Joglar, J.; Bujons, J.; Clapés, P. Asymmetric Assembly of Aldose Carbohydrates from Formaldehyde and Glycolaldehyde by Tandem Biocatalytic Aldol Reactions. *Nat. Chem.* **2015**, *7*, 724-729.
14. Jadhav, A.; Pathare, D.; Shingare, M. Validated Enantioselective LC Method, with Precolumn Derivatization with Marfey's Reagent, for Analysis of the Antiepileptic Drug Pregabalin in Bulk Drug Samples. *Chromatographia* **2007**, *65*, 253-256.



# Chapter 3

## *Tuning Enzyme Activity for Nonaqueous Solvents: Engineering of an Enantioselective ‘Michaelase’ for Catalysis in High Concentrations of Ethanol*

---

Chao Guo,<sup>†[a]</sup> Lieuwe Biewenga,<sup>†[a]</sup> Max Lubberink,<sup>[a,b]</sup>  
Ronald van Merkerk,<sup>[a]</sup> and Gerrit J. Poelarends<sup>\*[a]</sup>

<sup>a</sup>Department of Chemical and Pharmaceutical Biology, Groningen Research Institute of Pharmacy,  
University of Groningen, Antonius Deusinglaan 1, 9713 AV Groningen (The Netherlands)

<sup>b</sup>Present address: School of Chemistry, Manchester Institute of Biotechnology, The University of  
Manchester, 131 Princess Street, M1 7DN, Manchester (UK).

<sup>†</sup>These authors contributed equally: Chao Guo and Lieuwe Biewenga.

\*Corresponding author. Tel.: +31503633354; E-mail: [g.j.poelarends@rug.nl](mailto:g.j.poelarends@rug.nl);  
Web: <http://www.rug.nl/staff/g.j.poelarends/>

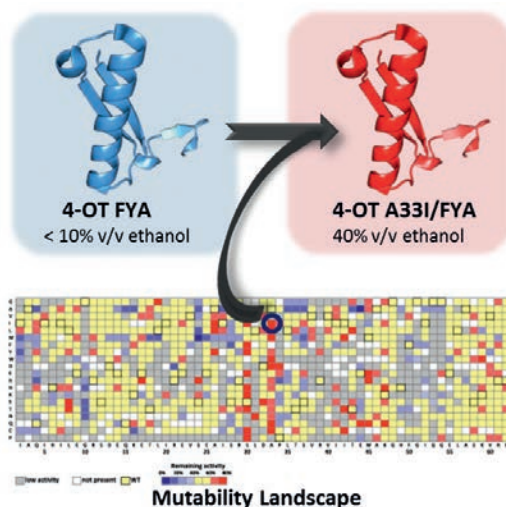
Published in *ChemBioChem* 10.1002/cbic.201900721

## Abstract

Enzymes have evolved to function under aqueous conditions and may not exhibit features essential for biocatalytic application, such as the ability to function in high concentrations of an organic solvent. Consequently, protein engineering is often required to tune an enzyme for catalysis in non-aqueous solvents. In this study, we have used a collection of nearly all single mutants of 4-oxalocrotonate tautomerase, which promiscuously catalyzes synthetically useful Michael-type additions of acetaldehyde to various nitroolefins, to investigate the effect of each mutation on the ability of this enzyme to retain its ‘Michaelase’ activity in elevated concentrations of ethanol. Examination of this mutability landscape allowed the identification of two hotspot positions, Ser-30 and Ala-33, at which mutations are beneficial for catalysis in high ethanol concentrations. The ‘hotspot’ position Ala-33 was then randomized in a highly enantioselective, but ethanol-sensitive 4-OT variant (L8F/M45Y/F50A) to generate an improved enzyme variant (L8F/A33I/M45Y/F50A) that showed great ethanol stability and efficiently catalyzes the enantioselective addition of acetaldehyde to nitrostyrene in 40% ethanol (permitting high substrate loading) to give the desired  $\gamma$ -nitroaldehyde product in excellent isolated yield (89%) and enantiopurity (ee = 98%). The presented work demonstrates the power of mutability-landscape-guided enzyme engineering for efficient biocatalysis in non-aqueous solvents.

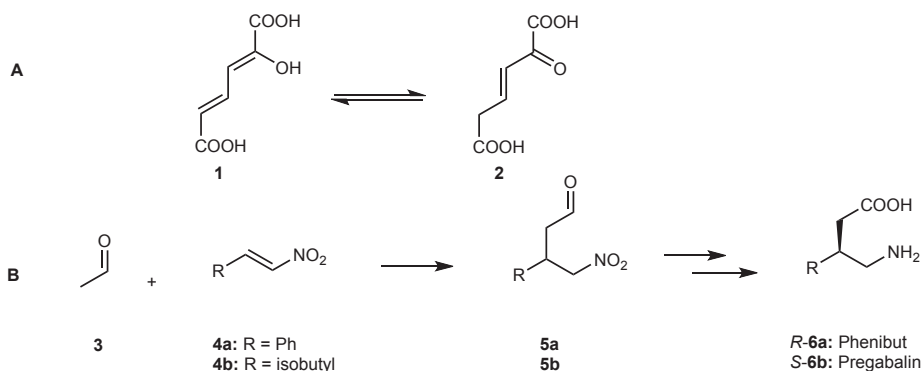
## Keywords

Enzyme engineering; biocatalysis; solvents; mutability landscape; Michael addition



## Introduction

The enzyme 4-oxalocrotonate tautomerase (4-OT) from *Pseudomonas putida* mt-2 catalyzes the tautomerization of 2-hydroxy-2,4-hexadienedioate (**1**) to 2-oxo-3-hexenedioate (**2**) as part of a metabolic pathway for the degradation of aromatic hydrocarbons (Scheme 1A).<sup>[1,2]</sup> In addition, 4-OT can promiscuously catalyze several C-C bond-forming reactions, including Michael-type additions and aldol condensations, yielding precursors for important classes of pharmaceuticals.<sup>[3-7]</sup> For instance, the 4-OT catalyzed Michael-type addition of acetaldehyde (**3**) to nitroalkenes **4a** and **4b** yields  $\gamma$ -nitroaldehydes **5a** and **5b**, important precursors for the  $\gamma$ -aminobutyric acid analogues phenibut (*R*-**6a**) and pregabalin (*S*-**6b**), respectively (Scheme 1B).<sup>[5]</sup> Hence, several enzyme engineering studies have been performed to improve the activity and enantioselectivity of 4-OT for this reaction.<sup>[8,9]</sup>



**Scheme 1.** A) Tautomerization reaction naturally catalyzed by 4-OT. B) Michael-type addition of acetaldehyde (**3**) to nitroalkenes **4a** and **4b**, promiscuously catalyzed by 4-OT. Products **5a** and **5b** are precursors for phenibut (*R*-**6a**) and pregabalin (*S*-**6b**), respectively.

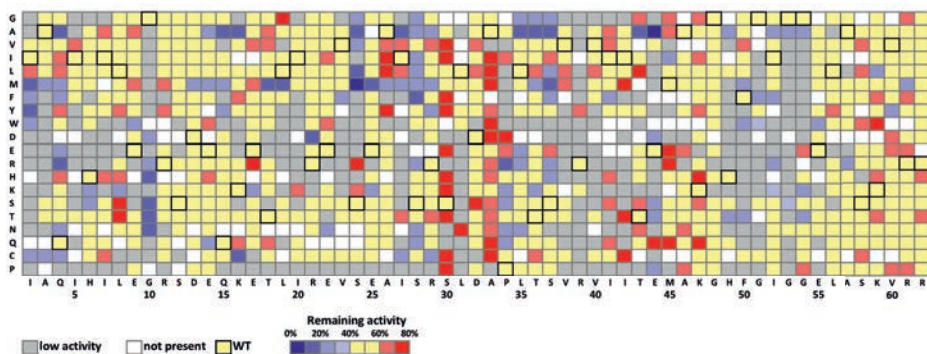
Solubilization of substrates **4a** and **4b** requires the use of cosolvents. Since enzymes have evolved to function under aqueous conditions, high concentrations of cosolvents can significantly affect their catalytic performance and eventually result in enzyme precipitation.<sup>[10]</sup> In this study, we used a collection of nearly all single-mutant variants of 4-OT to investigate the effect of each mutation on the ability of the enzyme to retain its ‘Michaelase’ activity in elevated concentrations of ethanol. Ethanol was selected as cosolvent because it is readily accessible from biorenewable sources and can also function as a precursor for **3**.<sup>[11]</sup> Randomization of the identified ‘hotspot’ position Ala-33 in the context of a previously engineered highly enantioselective, but ethanol-sensitive,



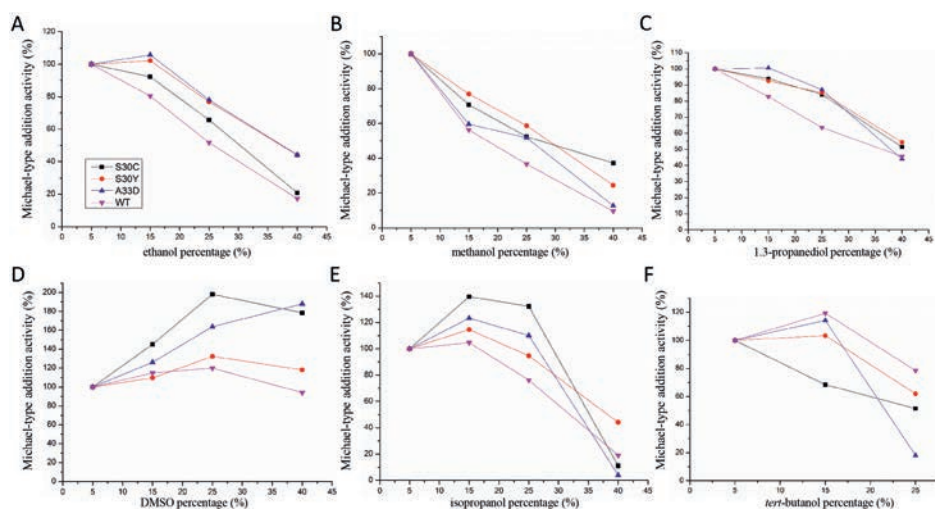
4-OT variant (L8F/M45Y/F50A) afforded an improved enzyme variant (4-OT L8F/A33I/M45Y/F50A) with high ethanol stability, allowing efficient and enantioselective Michael-type addition reactions in 40% (v/v) ethanol. As such, our work provides an interesting example of how targeted mutagenesis of a single amino acid can radically modify the cosolvent stability of an enzyme, allowing efficient catalysis in high concentrations of ethanol.

## Results

In order to identify ‘hotspot’ positions of 4-OT at which mutations are beneficial for catalysis in high concentrations of ethanol, a defined collection of 1040 single-mutant variants of 4-OT<sup>[8]</sup> was screened using cell-free extracts (CFEs) prepared from cultures each expressing a different 4-OT mutant. The Michael-type addition of **3** to **4a** was used as a model reaction in screening owing to the marked absorbance of **4a** at 320 nm. Control experiments demonstrated no significant difference between the effect of ethanol on the reaction catalyzed by purified 4-OT or 4-OT present in CFE (Figure S1). The ‘Michaelase’ activity of each single-mutant variant of 4-OT was measured using either 5% or 25% ethanol as cosolvent and the remaining activity at 25% ethanol, compared to that at 5% ethanol, was graphically represented in a mutability landscape for solvent tolerance (Figure 1). Increasing the concentration of ethanol from 5% to 25% reduced the ‘Michaelase’ activity of wild-type 4-OT by approximately 50% (Figures 1 and 2A). Interestingly, analysis of the ethanol-tolerance mutability landscape revealed two ‘hotspot’ positions, Ser-30 and Ala-33, at which single mutations resulted in enzyme variants that showed more than 70% residual ‘Michaelase’ activity at 25% ethanol (Figure 1). Notably, the crystal structure of wild-type 4-OT does not provide an immediate explanation for the improved ethanol tolerance caused by mutations at these two positions (Figure S2), illustrating the importance of mutability-landscape navigation to identify functional ‘hotspot’ positions. Three single mutants, 4-OT S30C, S30Y and A33D, which showed high ethanol tolerance, were purified and the effect of ethanol and other cosolvents on the ‘Michaelase’ activity was tested (Figure 2). Interestingly, 4-OT S30C, S30Y and A33D also showed tolerance towards other cosolvents such as DMSO and isopropanol, suggesting that these mutations convey general cosolvent resistance. Notably, while the 4-OT variants perform well up to 40% DMSO, visible protein precipitation with concomitant loss of activity was observed at DMSO concentrations  $\geq 50\%$  (v/v).



**Figure 1.** Ethanol-tolerance mutability landscape of 4-OT. The horizontal axis of the data matrix represents the residue positions of 4-OT. The vertical axis represents all 20 canonical amino acids. The wild-type amino acid at each position is indicated with a bold square. White squares indicate that this mutant is not present in the collection. The color of the square indicates the residual 'Michaelase' activity of a specific single-mutant variant of 4-OT for the addition of **3** to **4a** in 25% ethanol, compared to that in 5% ethanol. Grey boxes indicate that the 'Michaelase' activity was too low to determine the remaining activity.



**Figure 2.** Michael-type addition of **3** to **4a** catalyzed by purified wild-type 4-OT or 4-OT mutants in the presence of different cosolvent concentrations. A) ethanol, B) methanol, C) 1,3-propanediol, D) DMSO, E) isopropanol, F) *tert*-butanol. The activity of each mutant is normalized to the activity in the presence of 5% cosolvent.

We next investigated if we could use the information from the solvent-tolerance mutability landscape to engineer a previously constructed highly enantioselective 4-OT variant, L8F/M45Y/F50A (4-OT FYA),<sup>[9]</sup> to function in high concentrations of ethanol. As single mutants

at ‘hotspot’ position Ala-33 generally exhibited higher ‘Michaelase’ activity than those at ‘hotspot’ position Ser-30, we focused our mutagenesis strategy on position Ala-33.<sup>[8]</sup> In the context of 4-OT FYA, residue Ala-33 was mutated to all possible amino acids and the nineteen enzyme variants were expressed and purified to homogeneity. Initially, we tested all variants for visible precipitation upon incubation (1 h) of the enzyme with increasing concentrations of ethanol (up to 50%). The parental enzyme 4-OT FYA rapidly precipitated when incubated with ethanol concentrations equal to or greater than 10% (Figure 3D). Interestingly, enzyme variants with isoleucine (A33I/FYA), leucine (A33L/FYA) or valine (A33V/FYA) at position 33 could tolerate up to 50% ethanol without any visible protein precipitation after 1 hour of incubation. Notably, substitution of Ala-33 to aspartate, glutamate or cysteine in the context of 4-OT FYA also strongly improved the stability of the enzyme in high concentrations of ethanol, tolerating up to 40% ethanol without visible protein precipitation.

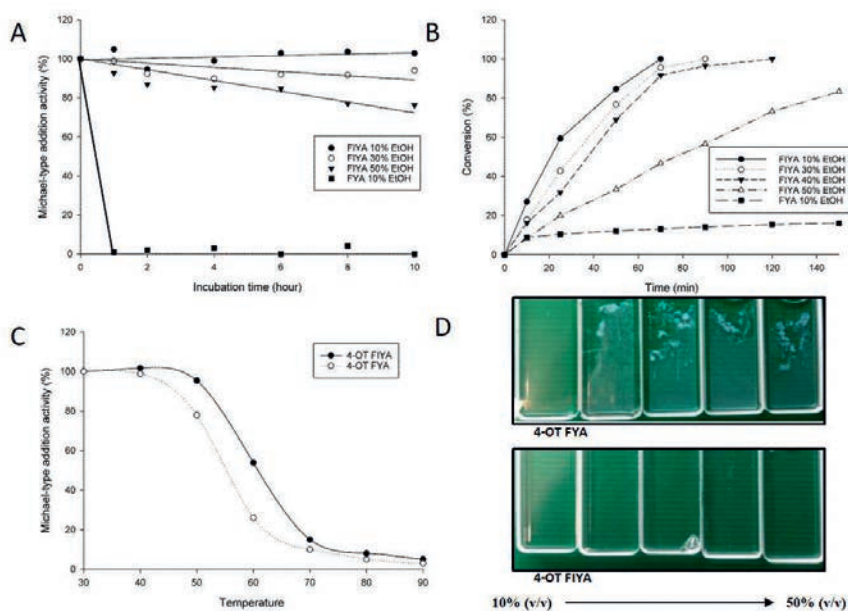
Next, we tested the activity and enantioselectivity of these six quadruple mutants in the presence of 5%, 30% or 50% ethanol. The Michael-type addition of **3** to **4b** was used as model reaction because the optical purity of product **5b** can easily be analyzed by gas chromatography. All six quadruple mutants proved to be highly enantioselective giving nearly enantiopure product **5b**, and, importantly, increasing ethanol concentrations do not negatively affect enzyme enantioselectivity (Table 1). In the presence of 30% ethanol, the reactions with the six quadruple mutants were completed within 35-70 min, whereas the reaction with the parental enzyme (4-OT FYA) showed no conversion due to rapid protein precipitated (Table 1).

**Table 1.** Biocatalytic addition of **3** to **4b** catalyzed by 4-OT mutants in different ethanol concentrations.<sup>[a]</sup>

Enzymes	5% v/v ethanol		30% v/v ethanol		50% v/v ethanol	
	<i>ee</i> <sup>[b]</sup>	Reaction time (min) <sup>[c]</sup>	<i>ee</i> <sup>[b]</sup>	Reaction time (min) <sup>[c]</sup>	<i>ee</i> <sup>[b]</sup>	Reaction time (min) <sup>[c]</sup>
FYA	98 (S)	40	– <sup>[d]</sup>	– <sup>[d]</sup>	– <sup>[d]</sup>	– <sup>[d]</sup>
A33D/FYA	98 (S)	50	98 (S)	70	92 (S)	> 360 <sup>[e]</sup>
A33E/FYA	96 (S)	30	96 (S)	60	98 (S)	> 360 <sup>[e]</sup>
A33C/FYA	98 (S)	55	98 (S)	70	– <sup>[d]</sup>	– <sup>[d]</sup>
A33I/FYA	98 (S)	30	98 (S)	35	98 (S)	90
A33L/FYA	98 (S)	35	98 (S)	50	96 (S)	180
A33V/FYA	98 (S)	35	98 (S)	45	98 (S)	120

<sup>[a]</sup> Assay conditions: 3 mM **4b**, 100 mM **3**, 73  $\mu$ M 4-OT, 20 mM NaH<sub>2</sub>PO<sub>4</sub> (pH 7.3), 0.3 mL reaction volume <sup>[b]</sup> Determined by GC using a chiral stationary phase; the absolute configuration was determined by literature comparison<sup>[8,9]</sup> <sup>[c]</sup> Reaction progress was monitored by following the depletion in absorbance at 249 nm. <sup>[d]</sup> No data due to protein precipitation. <sup>[e]</sup> Reaction was not finished after 360 min.

The best mutant, 4-OT A33I/FYA, catalyzed the Michael-type addition reaction practically as efficient in 40% ethanol as in 10% ethanol (Figure 3B). Moreover, pre-incubation of 4-OT A33I/FYA in 50% ethanol for 10 h resulted in only 25% loss of activity (Figure 3A). On the contrary, the parental enzyme (4-OT FYA) rapidly lost its activity upon incubation in 10% ethanol (Figure 3A, B). Interestingly, an increase in the  $T_{50}^{60}$  of approximately 6 °C was observed for 4-OT A33I/FYA compared to 4-OT FYA, indicating that 4-OT A33I/FYA is also somewhat more thermostable than 4-OT FYA (Figure 3C). Finally, to further demonstrate the synthetic usefulness of 4-OT A33I/FYA, a semi-preparative scale reaction was performed, using 40% ethanol as cosolvent, which allowed for the solubilization of 15 mM **4a**. Using a 6.7-fold excess of **3** over **4a**, the reaction was finished within 200 min. Product **R-5a** was obtained in excellent isolated yield (89%) and enantiopurity ( $ee = 98\%$ ) (Figures S3 and S4). Taken together, these results demonstrate that 4-OT can be engineered to efficiently catalyze enantioselective Michael-type reactions in ethanol concentrations up to 40%.



**Figure 3.** Characterization of 4-OT A33I/FYA (FIYA). A) Enzyme activity for the Michael-type addition of **3** to **4a** after pre-incubation (1, 2, 4, 6, 8 or 10 h) in the presence of 10, 30 or 50% ethanol. The data is normalized to the activity of the enzyme without pre-incubation. B) Progress curves of 4-OT A33I/FYA- and 4-OT FYA-catalyzed Michael-type additions in the presence of different ethanol concentrations. C) Temperature-induced inactivation profiles of 4-OT A33I/FYA and 4-OT FYA. The enzyme activity after incubation for 60 min at 30 °C was set as 100%. D) Photograph of cuvettes in which 4-OT FYA (190 μg/mL, top five cuvettes) and 4-OT A33I/FYA (190 μg/mL, bottom five cuvettes) were incubated for 1 h in 20 mM NaH<sub>2</sub>PO<sub>4</sub> buffer (pH 7.3) containing 10%, 20%, 30%, 40% or 50% (v/v) ethanol (cuvette 1 to 5, respectively, from left to right). 4-OT FYA shows rapid precipitation in buffer containing >10% ethanol.

## Discussion

Enzymes are highly attractive catalysts for organic synthesis because of their unparalleled enantio-, regio- and chemoselectivity. Given that enzymes have evolved to operate in the mild aqueous environment of the cell, they are usually not fit for preparative biocatalysis in the presence of high concentrations of organic cosolvents required for substrate solubilization.<sup>[10]</sup> A solution to this problem is the engineering of enzymes to improve their cosolvent tolerance. Rational enzyme engineering towards increased cosolvent tolerance is still very challenging due to our relatively poor understanding of the interactions between enzymes and solvent molecules.<sup>[12-14]</sup> Currently employed rational engineering strategies include stabilization of flexible regions, introduction of new cysteine bridges and modification of access tunnels.<sup>[14-17]</sup>

An important strategy to guide enzyme-engineering efforts is to make use of mutability landscapes.<sup>[18-21]</sup> By screening a large collection of nearly all single mutants of an enzyme, important information is obtained on single mutations or residue positions that influence a desired characteristic of the enzyme. Here we have applied mutability-landscape guided enzyme engineering to improve the ethanol tolerance of 4-OT. Screening of a collection of nearly all single-mutant variants of 4-OT revealed that mutations at particularly positions Ser-30 and Ala-33 resulted in improved ethanol tolerance. Interestingly, a previously reported 4-OT variant with 3.5-fold increased 'Michaelase' activity, 4-OT A33D, also showed improved ethanol tolerance.<sup>[8]</sup> We used the information from the ethanol-tolerance mutability landscape to further engineer a previously constructed highly enantioselective 4-OT variant (FYA) that exhibits poor ethanol stability. All 19 possible variants at position Ala-33 in the context of 4-OT FYA were constructed, expressed and purified. Remarkably, from this small, focused set of quadruple mutants, six mutants showed strongly improved ethanol tolerance. It is interesting to note that all six variants, including 4-OT A33D/FYA, are highly enantioselective towards the synthesis of *S*-**5b**, similar to the parental mutant 4-OT FYA.<sup>[9]</sup> Conversely, the single mutant 4-OT A33D markedly improved the enantioselectivity towards the opposite enantiomer *R*-**5b**.<sup>[8]</sup>

Incubation of the best mutant, 4-OT A33I/FYA, for 10 h in the presence of 50% ethanol resulted in only 25% loss of activity, whereas the parental enzyme 4-OT FYA lost all its activity upon incubation for 1 h in 10% ethanol. 4-OT A33I/FYA also showed an increase in thermostability compared to 4-OT FYA, an effect that has also been observed for other enzymes that have been engineered towards increased solvent tolerance.<sup>[22,23]</sup> We further show that 4-OT A33I/FYA can be used to efficiently catalyze the Michael-type addition

of **3** to **4a** in the presence of 40% ethanol, which permitted the use of a higher substrate loading (up to 15 mM **4a**). Product *R*-**5a** could be obtained in good isolated yield (89%) and with excellent enantiopurity (ee = 98%).

In conclusion, our results demonstrate the power of mutability landscapes to guide engineering efforts to improve the cosolvent tolerance of enzymes. By specifically targeting the identified ‘hotspot’ position Ala-33, we could engineer an ethanol-sensitive mutant, 4-OT FYA, into a highly ethanol-resistant mutant 4-OT A33I/FYA. Further tuning of 4-OT A33I/FYA might lead to new synthetic opportunities in almost neat organic solvents.

## Experimental section

**Production of cell-free extract.** Cell-free extracts (CFE) of 4-OT single mutants were prepared according to a reported procedure.<sup>[8]</sup>

**Construction of the ethanol-tolerance mutability landscape.** The CFEs prepared from cells producing 4-OT single-mutant variants were used in two reactions, containing either 5% or 25% v/v ethanol. The following reaction conditions were used: CFE (20% v/v), **3** (50 mM), **4a** (0.5 mM) in 20 mM NaH<sub>2</sub>PO<sub>4</sub> buffer (pH 7.3), 100 μL final volume. The reactions were performed in 96-well microtiter plates (MTP) (UV-star μclear, Greiner Bio-one), covered with UV-transparent plate seals (VIEWseal™, Greiner Bio-one). To ensure proper mixing of the reagents, the plate was shaken (60 s at 500 rpm) immediately after all reaction components were added. The reaction progress was monitored in a plate reader by measuring the depletion in absorbance at 320 nm, corresponding to the concentration of **4a**, for 60 min with a 60 s data interval. The slope of the linear part of the curve was determined for both the reactions. The remaining enzymatic activity was determined by dividing the slope of the reaction in 25% v/v ethanol by the slope of the reaction in 5% v/v ethanol.

**4-OT purification.** The purification of 4-OT single mutants<sup>[24]</sup> and 4-OT quadruple mutants<sup>[9]</sup> are based on previously reported procedures. All purified proteins were >90% pure as assessed by SDS-PAGE. All purified mutants were analyzed by electron spray ionization (ESI) mass spectrometry to confirm the correct mass of the enzyme. The purified protein was flash-frozen in liquid nitrogen and stored at -80 °C until further use.

**UV-spectroscopic assay for the enzymatic activity of 4-OT single mutants in different organic solvents.** The enzymatic activities of the 4-OT mutants and wild-type 4-OT was monitored by following the decrease in absorbance at 320 nm, which corresponds to the depletion of **4a**. Purified enzyme (150  $\mu\text{g}$ , 73  $\mu\text{M}$ ) was incubated in a 1 mm cuvette with **3** (50 mM) and **4a** (2 mM) in 20 mM  $\text{NaH}_2\text{PO}_4$  (pH 7.3; 0.3 mL final volume).

**Construction of nineteen Ala-33 mutants of 4-OT L8F/M45Y/F50A.** Ala-33 was randomized by Quikchange technology using the gene encoding 4-OT L8F/M45Y/F50A cloned in the pET20b vector as the template. The following two primers were used: 5'-GCTCCCTGGATNNKCCGCTGACCAG-3' and 5'-CTGGTCAGCGMNNATCCAGGGAGC-3'. After transformation of the DNA into *E. coli* cells, random colonies were picked from an agar plate, and the mutant 4-OT genes were sequenced by Macrogen Europe (Meibergdreef 31, 1105AZ, Amsterdam, the Netherlands) until all of the 19 quadruple mutants were obtained.

**Activity assays of the six best quadruple mutants.** The enzymatic activities of the 4-OT quadruple mutants and 4-OT L8F/M45Y/F50A were monitored by following the decrease in absorbance at 249 nm, which corresponds to the depletion of **4b**. Purified enzyme (150  $\mu\text{g}$ , 73  $\mu\text{M}$ ) was incubated in a 1 mm cuvette with **3** (100 mM) and **4b** (3 mM) in 20 mM  $\text{NaH}_2\text{PO}_4$  buffer (pH 7.3; 0.3 mL final volume). After the reactions were completed, product **5b** was extracted with ethyl acetate (400  $\mu\text{L}$ ) and analyzed by gas chromatography using an Astec CHIRALDEX G-TA column, isocratic 125  $^\circ\text{C}$ . Retention time *S-5b*: 25.6 min, retention time *R-5b*: 26.9. The assignment of the absolute configuration was based on earlier reported data.<sup>[9]</sup>

**Determination of  $T_{50}^{60}$ .** 4-OT L8F/M45Y/F50A and 4-OT L8F/A33I/M45Y/F50A (50  $\mu\text{L}$  of 2 mg/mL in 20 mM  $\text{NaH}_2\text{PO}_4$ , pH 7.3) were incubated in 0.2-mL PCR tubes at temperatures ranging from 30  $^\circ\text{C}$  to 90  $^\circ\text{C}$  for 60 min in a thermal cycler. After incubation, the enzymes were cooled on ice for 10 min followed by equilibration at 25  $^\circ\text{C}$  for 10 min. Samples were centrifuged to remove any precipitated protein. The residual 'Michaelase' activity for the addition of **3** to **4a** was tested in a plate reader. Following conditions were used: 25  $\mu\text{L}$  of enzyme supernatant, 50 mM **3**, 0.5 mM **4a**, 5% v/v ethanol in 20 mM  $\text{NaH}_2\text{PO}_4$  buffer (pH 7.3), 100  $\mu\text{L}$  final volume. The 'Michaelase' activities were normalized to that obtained after 60 min incubation at 30  $^\circ\text{C}$ .

**Stability of 4-OT L8F/M45Y/F50A and 4-OT L8F/A33I/M45Y/F50A upon incubation with increasing ethanol concentrations.** 4-OT L8F/M45Y/F50A and 4-OT L8F/A33I/M45Y/F50A (1 mL of 1.5 mg/mL in 20 mM  $\text{NaH}_2\text{PO}_4$ , pH 7.3) were incubated in 20 mM

NaH<sub>2</sub>PO<sub>4</sub> buffer (pH 7.3) containing 10%, 30% or 50% v/v ethanol in a water bath of 25 °C. Aliquots of enzyme (80 μL) were taken at different time intervals and centrifuged to remove any aggregated protein. 50 μL of the supernatant was used to test the residual enzymatic activity. The reaction mixture consisted of the following: **3** (50 mM), **4a** (2 mM, from a 40 mM stock solution in 100% (v/v) ethanol) in 20 mM NaH<sub>2</sub>PO<sub>4</sub> buffer (pH 7.3), 0.3 mL final volume. Depletion of **4a** was monitored by following the decrease in absorbance at 320 nm in time. The activities were normalized to the activity measured without incubation of the enzyme.

**Semi-preparative scale synthesis.** To a 50 mL round bottom flask was added: 6 mL ethanol, 112 μL **3**, 12 mL buffer (20 mM NaH<sub>2</sub>PO<sub>4</sub>, pH 6.5) containing 4-OT L8F/A33I/M45Y/F50A. The reaction was initiated by the addition of 2 mL ethanol containing 150 mM **4a**. The final concentrations were: **3** (100 mM), **4a** (15 mM), 4-OT L8F/A33I/M45Y/F50A (75 μM, based on monomer concentration), and 40% (v/v) ethanol in 20 mM NaH<sub>2</sub>PO<sub>4</sub> buffer (pH 6.5). The reaction progress was monitored using UV-spectrophotometric analysis. At timely intervals, a sample of 30 μL was collected from the reaction mixture and diluted to 300 μL with 20 mM NaH<sub>2</sub>PO<sub>4</sub> buffer and a full spectrum from 200 nm to 500 nm was recorded. After 200 min, the reaction was finished. The reaction mixture was extracted 3x with ethyl acetate. The combined organic layers were washed with brine and dried over anhydrous Na<sub>2</sub>SO<sub>4</sub>. The organic layer was concentrated in vacuo, yielding **5a** without any further purification (51.5 mg, 89% yield). The aldehyde functionality of **5a** was derivatized to a cyclic acetal according to a reported procedure<sup>[5]</sup>. The enantiopurity of derivatized **5a** was determined by reverse phase HPLC using a chiralpak AD-RH column (150 mm x 4.6 mm, Daicel) MeCN/water 70:30. Retention time *R*-**5a**: 7.8 min, *S*-**5a**: 10.8 min.

### Acknowledgement

We acknowledge financial support from the Netherlands Organization of Scientific Research (VICI grant 724.016.002) and the European Union's Horizon 2020 Research and Innovation Programme under grant agreement No 635595 (CarbaZymes).

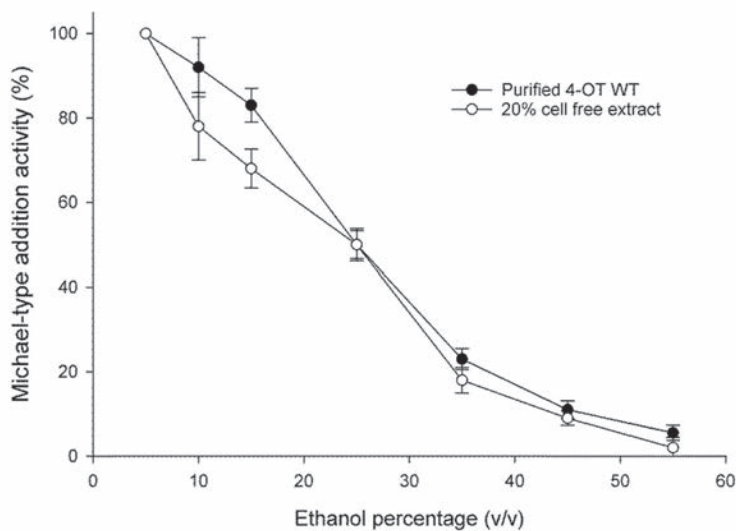


## References

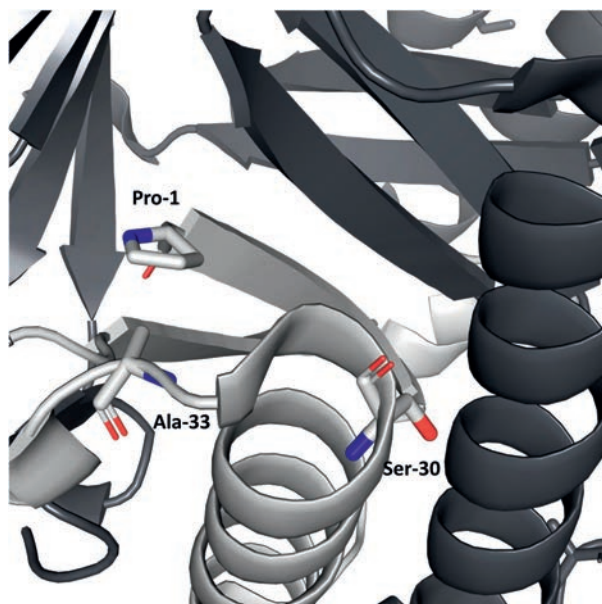
1. S. Harayama, M. Rekik, K. L. Ngai, L. N. Ornston, *J. Bacteriol.* **1989**, *171*, 6251 – 6258.
2. C. P. Whitman, B. A. Aird, W. R. Gillespie, N. J. Stolowich, *J. Am. Chem. Soc.* **1991**, *113*, 3154–3162.
3. M. Rahimi, J. Y. van der Meer, E. M. Geertsema, H. Poddar, B. J. Baas, G. J. Poelarends, *ChemBioChem* **2016**, *17*, 1225–1228.
4. M. Rahimi, J. Y. van der Meer, E. M. Geertsema, G. J. Poelarends, *ChemBioChem* **2017**, *18*, 1435–1441.
5. E. M. Geertsema, Y. Miao, P. G. Tepper, P. de Haan, E. Zandvoort, G. J. Poelarends, *Chem. Eur. J.* **2013**, *19*, 14407–14410.
6. E. Zandvoort, E. M. Geertsema, B. J. Baas, W. J. Quax, G. J. Poelarends, *Angew. Chem. Int. Ed.* **2012**, *51*, 1240–1243.
7. C. Guo, M. Saifuddin, T. Saravanan, M. Sharifi, G. J. Poelarends, *ACS Catal.* **2019**, *9*, 4369–4373.
8. J. Y. van der Meer, H. Poddar, B. J. Baas, Y. Miao, M. Rahimi, A. Kunzendorf, R. van Merkerk, P. G. Tepper, E. M. Geertsema, A. M. W. H. Thunnissen, W. J. Quax, G. J. Poelarends, *Nat. Commun.* **2016**, *7*, 10911.
9. L. Biewenga, T. Saravanan, A. Kunzendorf, J. Y. van der Meer, T. Pijning, P. G. Tepper, R. van Merkerk, S. J. Charnock, A.M. W. H. Thunnissen, G. J. Poelarends, *ACS Catal.* **2019**, *9*, 1503–1513.
10. V. Stepankova, S. Bidmanova, T. Koudelakova, Z. Prokop, R. Chaloupkova, J. Damborsky, *ACS Catal.* **2013**, *3*, 2823–2836.
11. D. Prat, O. Pardigon, H. W. Flemming, S. Letestu, V. Ducandas, P. Isnard, E. Guntrum, T. Senac, S. Ruisseau, P. Cruciani, P. Hosek, *Org. Process Res. Dev.* **2013**, *17*, 1517–1525.
12. R. J. Kazlauskas, U. T. Bornscheuer, *Nat. Chem. Biol.* **2009**, *5*, 526–529.
13. F. H. Arnold, *FASEB J.* **1993**, *7*, 744–749.
14. V. G. H. Eijnsink, A. Bjørk, S. Gåseidnes, R. Sirevåg, B. Synstad, B. van den Burg, G. Vriend, *J. Biotechnol.* **2004**, *113*, 105–120.
15. M. T. Reetz, J. D. Carballeira, A. Vogel, *Angew. Chem. Int. Ed.* **2006**, *45*, 7745–7751.
16. T. Koudelakova, R. Chaloupkova, J. Brezovsky, Z. Prokop, E. Sebestova, M. Hesseler, M. Khabiri, M. Plevaka, D. Kulik, I. Kuta Smatanova, P. Rezacova, R. Etrich, U. T. Bornscheuer, J. Damborsky, *Angew. Chem. Int. Ed.* **2013**, *52*, 1959–1963.
17. M. Karpusas, W. A. Baase, M. Matsumura, B. W. Matthews, *Proc. Natl. Acad. Sci.* **1989**, *86*, 8237 – 8241.
18. A. Fulton, V. J. Frauenkron-Machedjou, P. Skoczinski, S. Wilhelm, L. Zhu, U. Schwaneberg, K. E. Jaeger, *ChemBioChem* **2015**, *16*, 930–936.

19. P. A. Romero, T. M. Tran, A. R. Abate, *Proc. Natl. Acad. Sci.* **2015**, *112*, 7159 – 7164.
20. N. Palackal, Y. Brennan, W. N. Callen, P. Dupree, G. Frey, F. Goubet, G. P. Hazlewood, S. Healey, Y. E. Kang, K. A. Kretz, E. Lee, X. Tan, G. L. Tomlinson, J. Verruto, V. W. K. Wong, E. J. Mathur, J. M. Short, D. E. Robertson, B. A. Steer, *Protein Sci.* **2004**, *13*, 494–503.
21. J. Y. van der Meer, L. Biewenga, G. J. Poelarends, *ChemBioChem* **2016**, *17*, 1792–1799.
22. B. Rasekh, K. Khajeh, B. Ranjbar, N. Mollania, B. Almasinia, H. Tirandaz, *Eng. Life Sci.* **2014**, *14*, 442–448.
23. T. Börner, S. Rämisch, S. Bartsch, A. Vogel, P. Adlercreutz, C. Grey, *ChemBioChem* **2017**, *18*, 1482–1486.
24. E. Zandvoort, B. J. Baas, W. J. Quax, G. J. Poelarends, *ChemBioChem* **2011**, *12*, 602–609.

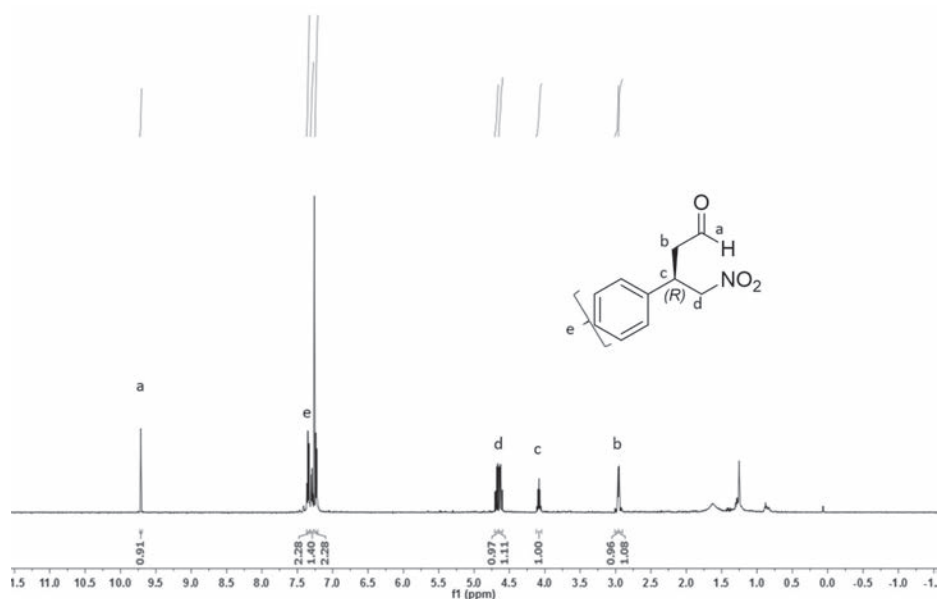
## Supporting Information



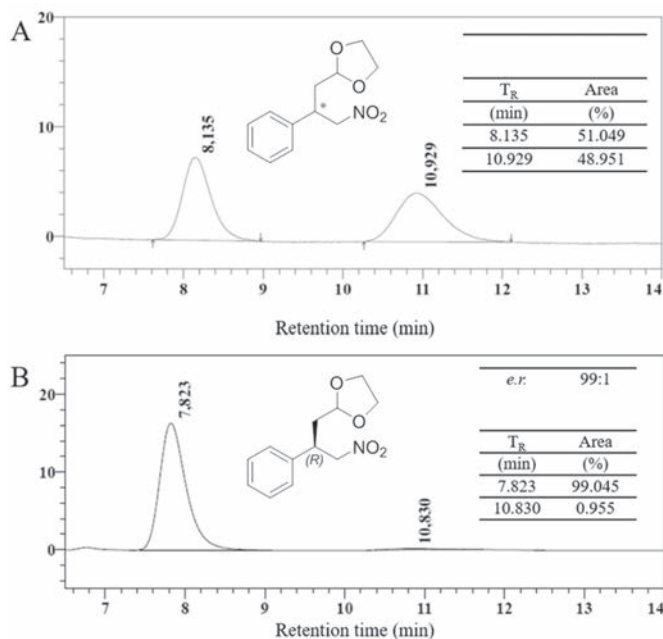
**Figure S1.** Activity of purified wild-type 4-OT and wild-type 4-OT in cell free extract for the Michael-type addition of **3** to **4a** in the presence of different ethanol concentrations.



**Figure S2.** Crystal structure of wild-type 4-OT (PDB: 4X1C). In light grey, a single monomer of 4-OT is indicated. Residues Pro-1, Ser-30 and Ala-33 are indicated as stick representation.



**Figure S3.**  $^1\text{H}$  NMR spectrum of enzymatically obtained **5a** (4-nitro-3-phenylbutanal).  $^1\text{H}$  NMR (500 MHz, Chloroform-*d*)  $\delta$  9.71 (s, 1H), 7.37 – 7.33 (m, 2H), 7.31 – 7.27 (m, 1H), 7.25 – 7.21 (m, 2H), 4.68 (dd,  $J = 12.5, 7.2$  Hz, 1H), 4.62 (dd,  $J = 12.5, 7.6$  Hz, 1H), 4.08 (p,  $J = 7.3$  Hz, 1H), 3.01 – 2.96 (m, 1H), 2.96 – 2.90 (m, 1H).



**Figure S4.** HPLC chromatograms of derivatized racemic (A) and enzymatically (B) obtained **5a**.



# Chapter 4

## *A Selective Colorimetric “Turn-on” Probe for Efficient Engineering of Iminium Biocatalysis*

---

Lieuwe Biewenga<sup>a</sup>, Michele Crotti<sup>a</sup>, Mohammad Saifuddin<sup>a</sup>, and Gerrit J. Poelarends<sup>\*a</sup>

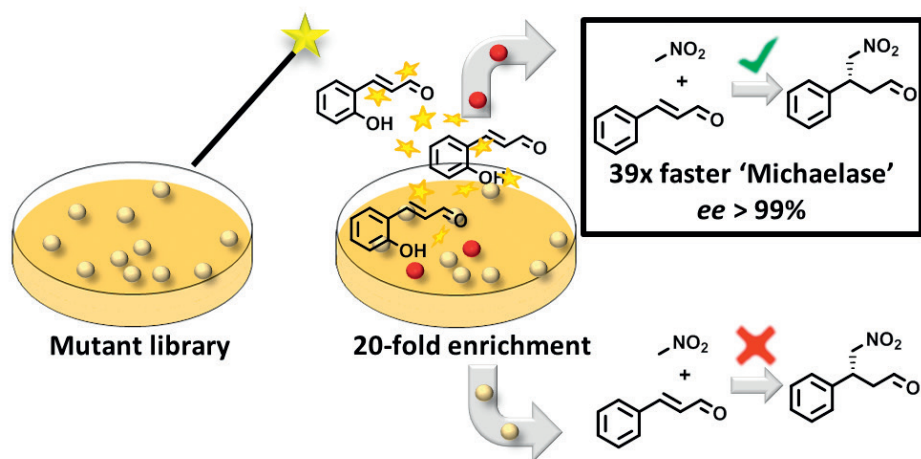
<sup>a</sup>*Department of Chemical and Pharmaceutical Biology, Groningen Research Institute of Pharmacy, University of Groningen, Antonius Deusinglaan 1, 9713 AV Groningen, The Netherlands.*

*\*Corresponding author. Tel.: +31503633354; E-mail: [g.j.poelarends@rug.nl](mailto:g.j.poelarends@rug.nl);  
Web: <http://www.rug.nl/staff/g.j.poelarends/>*

Published in ACS Omega 10.1021/acsomega.9b03849

## Abstract

The efficient engineering of iminium biocatalysis has drawn considerable attention, with many applications in pharmaceutical synthesis. Here we report a tailor-made iminium-activated colorimetric “turn-on” probe, specifically designed as a pre-screening tool to facilitate engineering of iminium biocatalysis. Upon complexation of the probe with the catalytic Pro-1 residue of the model enzyme 4-oxalocrotonate tautomerase (4-OT), a brightly colored merocyanine-dye-type structure is formed. 4-OT mutants that formed this brightly colored species upon incubation with the probe proved to have substantial activity for the iminium-based Michael-type addition of nitromethane to cinnamaldehyde, whereas mutants that showed no staining by the probe exhibited no or very low-level ‘Michaelase’ activity. This system was exploited in a solid-phase pre-screening assay termed as Activated Iminium Colony Staining (AICS) to enrich libraries for active mutants. AICS pre-screening reduced the screening effort up to 20-fold. After two rounds of directed evolution, two artificial ‘Michaelases’ were identified with up to 39-fold improvement in activity for the addition of nitromethane to cinnamaldehyde, yielding the target  $\gamma$ -nitroaldehyde product with excellent isolated yield (up to 95%) and enantiopurity (up to >99% ee). The colorimetric activation of the “turn-on” probe could be extended to the class I aldolase 2-deoxy-D-ribose 5-phosphate aldolase, implicating a broader application of AICS in engineering iminium biocatalysis.



## Introduction

Iminium-based catalysts are among the most versatile catalysts, finding applications in a myriad of C-C bond-forming transformations<sup>1</sup>. Many of these transformations lead to the formation of one or more asymmetric carbon atoms. Hence, researchers have focused on the discovery and engineering of iminium-based enzymes that could be applied as biocatalysts<sup>2-4</sup>. Unfortunately, nature’s repertoire of iminium-based carboligases is limited. This has motivated researchers to supplement natural iminium-based biocatalysts with artificial biocatalysts, either obtained by computational design<sup>5,6</sup> or inspired by natural scaffolds<sup>7,8</sup>. Regardless of the source of the biocatalyst, enzyme engineering is often required to improve the catalytic rate, enantioselectivity or stability. In the Hilvert-laboratory, the computationally designed retroaldolase RA95.0 was engineered toward Michael-type additions<sup>9</sup>, aldol reactions<sup>10,11</sup>, knoevenagel condensations<sup>12</sup> and Henry condensations<sup>13</sup>. In our group, we have focused on the engineering of a natural tautomerase, 4-oxalocrotonate tautomerase (4-OT). By exploiting the nucleophilic character of the key catalytic residue Pro-1, we have engineered 4-OT towards several enamine or iminium mediated C-C bond-forming reactions, including Michael-type additions and aldol reactions<sup>14-18</sup>.

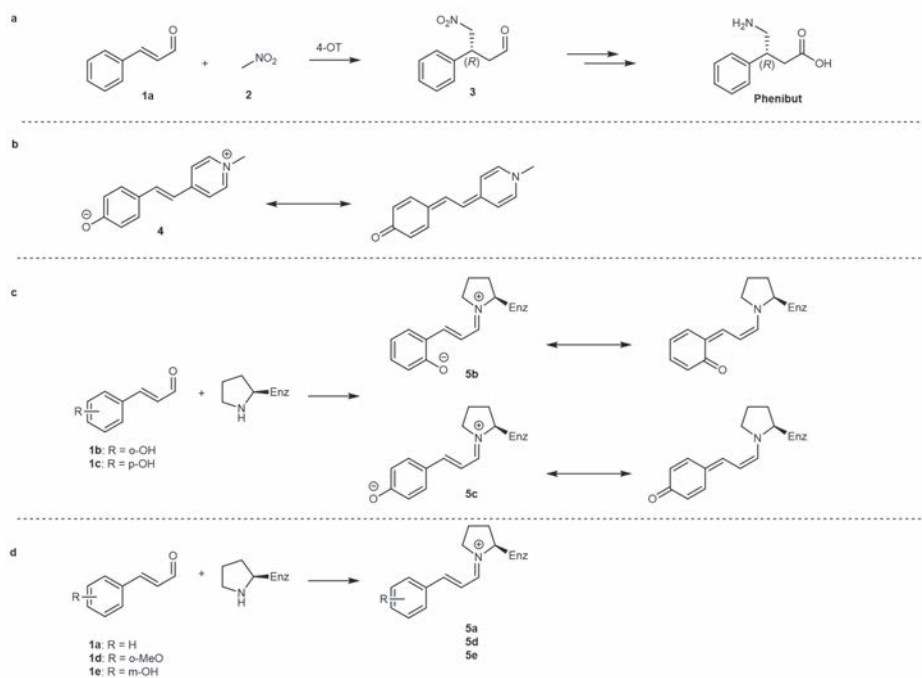
Despite large steps that have been made in techniques to construct small smart libraries, typically only a small fraction of the mutants within a library display the desired improved characteristic<sup>19-21</sup>. Especially improving the catalytic rate is challenging, as it is difficult to predict how mutations influence the catalytic machinery of the enzyme. As a consequence, the screening of mutants is still considered to be the bottleneck in enzyme engineering<sup>20</sup>.

Ultrahigh-Throughput screening techniques, such as Fluorescence-Activated Cell Sorting (FACS)<sup>22</sup> and Fluorescence-Activated Droplet Sorting (FADS)<sup>23,24</sup>, have enabled the screening of very large mutant libraries to enrich these libraries for active mutants. The resulting enriched libraries are small enough that they can be screened using lower throughput screening techniques such as 96-well assays, or HPLC/GC-based assays to identify individual mutants with the desired characteristic. The effectiveness of this approach was elegantly demonstrated by the FADS-based directed evolution of retroaldolase RA95.5-8, improving the catalytic efficiency >20-fold<sup>11,25</sup>. However, these techniques require specialized equipment and non-trivial reaction setup.

Here we present a simple pre-screening technique based on the iminium-activated colorimetric “turn-on” probe **1b** to pre-screen large 4-OT mutant libraries for improved



activity towards the iminium-mediated Michael-type addition of **2** to **1a** (Figure 1). The principle behind the pre-screening technique is based on the iminium-species that is selectively formed between **1b** and Pro-1 of 4-OT (Figure 1). This structure resembles a merocyanine dye and results in a new relatively bathochromic absorption peak that can easily be observed by the naked eye. This staining is not, or to a much lesser extent, observed in many inactive mutants which allows for those mutants to be excluded from subsequent screenings. We applied this technique in a solid-phase pre-selection assay that was termed as Activated Iminium Colony Staining (AICS). AICS pre-screening reduced screening efforts up to 20-fold and allowed for the screening of large mutant libraries. After 2 rounds of directed evolution we could improve the ‘Michaelase’ activity of 4-OT for the addition of **2** to **1a** up to 39-fold. We show that the principle behind AICS can be extended to other iminium-based biocatalysts like the class I aldolase DERA, which potentially paves the road for the engineering of novel biocatalysts for iminium-mediated C-C bond-forming reactions.



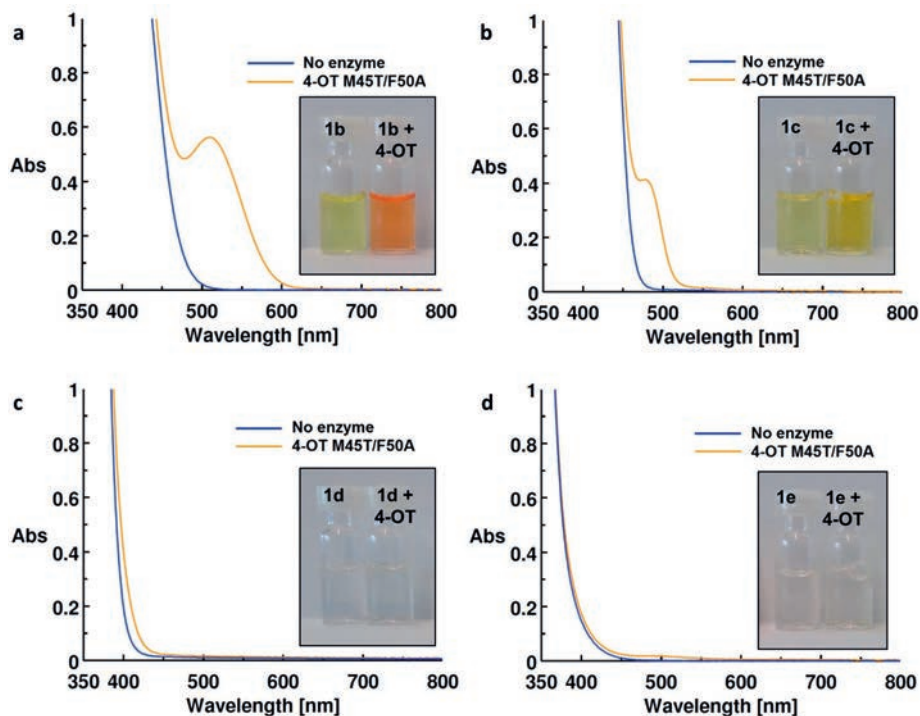
**Figure 1. 4-OT catalyzed Michael-type addition, Brooker's merocyanine and the reaction between Pro-1 of 4-OT and cinnamaldehydes. (a)** 4-OT catalyzed Michael-type addition of **2** to **1a**. Product **R-3** is a precursor for Phenibut. **(b)** Resonance forms of Brooker's merocyanine. **(c)** "Turn-on" probes **1b** and **1c** reacting with Pro-1 of 4-OT, forming a zwitterionic resonating species. **(d)** Compounds **1a**, **1d** and **1e** reacting with Pro-1 of 4-OT, forming a species without zwitterionic resonance.

## Results

### Identification of a selective “turn-on” probe for 4-OT

Recently, it was shown that Pro-1 of 4-OT can react with **1a** to form complex **5a**, a catalytic intermediate in the Michael-type addition of **2** to **1a** yielding **3**<sup>18</sup>. Analogously, we hypothesized that **1b** and **1c** could react with Pro-1 of 4-OT to form **5b** and **5c**, respectively, structures that bear chemical resemblance to Brooker’s merocyanine (**4**) (Figure 1). Compound **4** is a member of the diverse class of merocyanine dyes, which are characterized by an electron donating N-atom connected via a streptopolymethine chain to an electron withdrawing carbonyl group<sup>26,27</sup>. As a consequence, merocyanine dyes possess a neutral and a zwitterionic resonance form, and have a relatively high  $\lambda_{\text{max}}$ . As compound **1b** nor **1c** nor free 4-OT bear this characteristic chemical scaffold, incubation of **1b** and **1c** with 4-OT is expected to induce the formation of complexes **5b** and **5c** respectively, each having a new relatively bathochromic absorption peak.

To test this hypothesis, **1b** and **1c** were chemically synthesized (Figure S1). As a control, we synthesized compounds **1d** and **1e** that, like substrate **1a**, should not be able to form merocyanine dye-type structures upon condensation with Pro-1 of 4-OT because they lack an o-OH or p-OH group (Figure 1d). For testing we used 4-OT M45T/F50A, a well-studied 4-OT mutant that was constructed for the synthesis of **1a** via the aldol condensation of acetaldehyde with benzaldehyde<sup>15</sup>. Gratifyingly, upon incubation of **1b** or **1c** with 4-OT M45T/F50A, the color of the solution changed within a few seconds (Figure 2). Spectroscopic analysis revealed the appearance of a novel absorption peak at 516 nm for **1b** and 478 nm for **1c**, whereas incubation of 4-OT M45T/F50A with **1d** and **1e** did not reveal any new absorption peaks (Figure 2). Especially the bright red species that was formed upon incubation of **1b** with 4-OT M45T/F50A could clearly be seen by the naked eye, which made **1b** a powerful “turn on” probe for 4-OT. A single-site binding relationship between **1b** and 4-OT M45T/F50A was observed, with an estimated  $K_D$  of  $0.119 \pm 0.0055$  mM, which indicates that binding of **1b** to 4-OT is site specific (Figure S2).



**Figure 2. Absorption spectra of 1b-e in the presence and absence of 4-OT M45T/F50A.** (a) Absorption spectrum of **1b** with and without 4-OT M45T/F50A. (b) Absorption spectrum of **1c** with and without 4-OT M45T/F50A. (c) Absorption spectrum of **1d** with and without 4-OT M45T/F50A. (d) Absorption spectrum of **1e** with and without 4-OT M45T/F50A. Conditions: 2.5% DMSO, 20 mM sodium phosphate (pH 7.3), 1 mM **1b-e**, 50  $\mu$ M 4-OT M45T/F50A, 1 cm quartz cuvetts.

To gain further evidence for the formation of complex **5b**, 4-OT M45T/F50A was incubated with **1b** for 15 minutes, followed by the addition of  $\text{NaCNBH}_3$  to reduce any iminium species and covalently trap the interaction of the probe and the protein. Subsequent analysis by ESI-MS showed a mass consistent with that of 4-OT M45T/F50A modified with a single molecule of **1b** (Figure S3). Control reactions with 4-OT M45T/F50A incubated with only **1b** or only  $\text{NaCNBH}_3$  did not show any modification. The labeled 4-OT was digested with the endoproteinase Glu-C and the resulting peptides were analyzed by ESI-MS, revealing labeling of **1b** on the peptide fragment Pro-1 to Glu-9 (Figure S4). As Pro-1 is the only likely candidate on this peptide to form an iminium species with **1b**, Pro-1 was reckoned as the site of labeling.

Interestingly, incubation of **1b** with the enzyme 2-deoxy-D-ribose 5-phosphate aldolase (DERA), a well-studied class I aldolase<sup>4,28</sup>, also resulted in an increase in absorbance at 516 nm (Figure S5), implicating a broader applicability of probe **1b**.

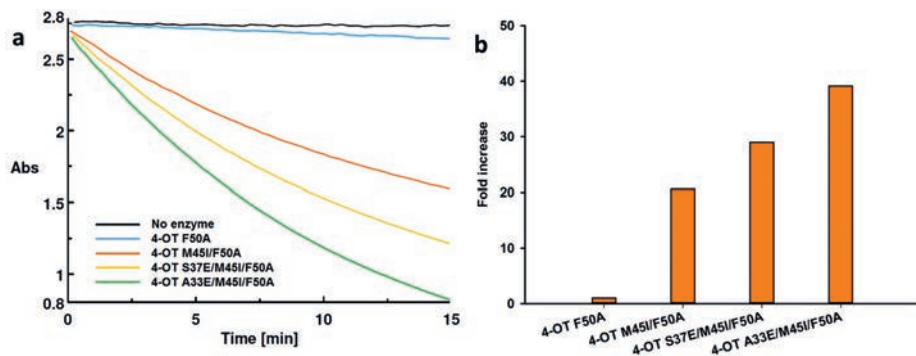
#### **Mutants with improved ‘Michaelase’ activity show strong staining with 1b**

Having developed a selective “turn-on” probe for 4-OT, we were interested in applying this system to aid our directed evolution efforts. Therefore the effect of mutations of 4-OT on the staining with **1b** was investigated. Cells producing a defined set of 53 single mutants, covering nearly all single mutants at positions 6, 45 and 50 were grown in a deep 96-well plate. These positions were selected because mutations at these positions have shown to influence the activity of 4-OT for several promiscuous reactions<sup>14–16,18</sup>. Incubation of cells producing the 53 single mutants with **1b** demonstrated that single mutations can significantly affect the staining with **1b** (Figure S6). Most notably, single mutants 4-OT F50A, F50I and F50V showed strong staining upon incubation with **1b**. Interestingly, these three mutants were previously found in a screen for mutants with enhanced activity for the Michael-type addition of **2** to **1a**, forming **3**<sup>18</sup>. These results prompted us to investigate if **1b** could be used as a pre-screening tool to enrich libraries for mutants with activity towards the Michael-type addition of **2** to **1a** yielding **3**, a precursor for phenibut, an important gamma-aminobutyric acid (GABA) analogue.

#### **Activated Iminium Colony Staining-based directed evolution of 4-OT**

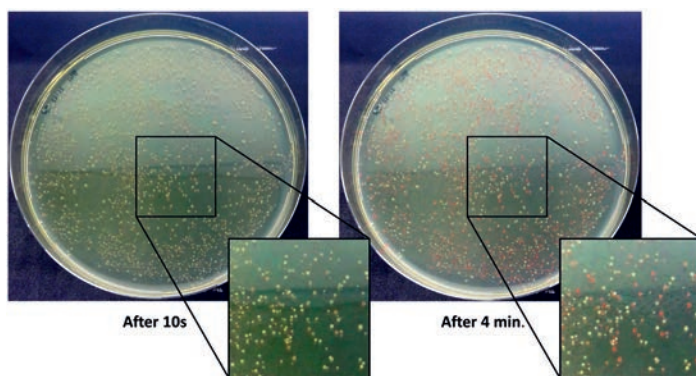
To improve the ‘Michaelase’ activity of 4-OT for the addition of **2** to **1a**, we started by constructing a triple-site NNK-library. We randomized the active site residues Gln-4, Leu-8 and Met-45 in the context of 4-OT F50A, the reported best mutant for this reaction<sup>18</sup>, yielding library Q4X/L8X/M45X/F50A. The addition of 0.2% lactose to agar plates induced the expression of 4-OT in freshly transformed cells, without any effect on the transformation efficiency. A solution of 0.75% agarose containing **1b** was poured on the agar plate and was allowed to solidify. Within 1 minute, ~5% of the colonies stained bright red and could be distinguished from other colonies by the naked eye.

A total of 1472 stained colonies were picked and screened for improved activity for the addition of **2** to **1a**. Several mutants with improved activity were identified. Almost all mutants harbored the M45I mutation, without any obvious additive effect of mutations at position 4 and 8. Hence, double mutant 4-OT M45I/F50A was constructed which showed a 20-fold improvement in activity compared to 4-OT F50A, using optimized reaction conditions (Figure 3).



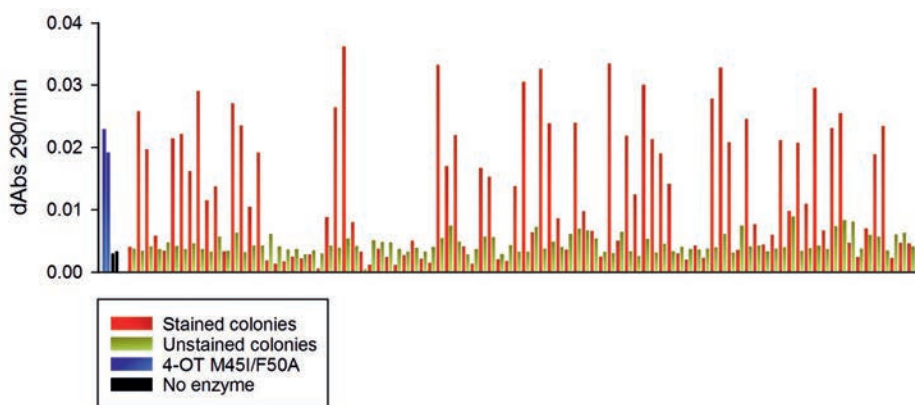
**Figure 3.** Activity for the Michael-type addition of 2 to 1a catalyzed by different 4-OT mutants. (a) Progress curve of the Michael-type addition of 2 to 1a catalyzed by different 4-OT mutants. (b) Initial rate of the Michael-type addition of 2 to 1a catalyzed by different 4-OT mutants. The data is presented as fold increase compared to the initial rate of starting mutant 4-OT F50A. Assay conditions: 5% ethanol, 20 mM HEPES (pH 6.5), 1 mM 1a, 100 mM 2, 50 mM sodium formate.

In a second round of directed evolution two libraries, 2-20\_IA and 21-40\_IA, were constructed containing a single random mutation from position 2-20 and 21-40 respectively in the context of 4-OT M45I/F50A. No mutations were introduced in the C-terminal part of 4-OT (position 41-62) as this region is highly important for stabilization of the hexameric structure<sup>29</sup>. The library was transformed into *Escherichia coli* cells, and colonies were stained using the AICS assay (Figure 4). Comparing the activity of 92 stained colonies with 92 unstained colonies of library 21-40\_IA demonstrated the high sensitivity of AICS pre-screening, as 32% of the stained mutants showed activity comparable or higher than the parental mutant 4-OT M45I/F50A, whereas all of the unstained mutants showed no activity or less than 50% of the activity of the parental mutant 4-OT M45I/F50A (Figure 5).



**Figure 4.** Agar plates 10 s (left) and 4 min (right) after performing the AICS assay. Transformants of Library 21-40\_IA were stained with the AICS assay as described in the experimental section.

From each library, 368 stained colonies were picked and screened for improved activity for the addition of **2** to **1a**. Although library 2-20\_IA yielded many active mutants, unfortunately none were significantly more active than parental mutant 4-OT M45I/F50A. Gratifyingly, library 21-40\_IA yielded 2 improved triple mutants: 4-OT S37E/M45I/F50A and 4-OT A33E/M45I/F50A with 29-fold and 39-fold improved activity compared to 4-OT F50A, respectively (Figure 3). The synthetic usefulness of the newly engineered mutants 4-OT S37E/M45I/F50A and 4-OT A33E/M45I/F50A was demonstrated in semi-preparative scale Michael-type addition reactions using substrates **1a** and **2**. Target product *R*-**3** could be obtained in good to excellent isolated yield (up to 95%) and with excellent enantiopurity (up to >99% enantiomeric excess, *ee*) (Table S1, Figures S7-S11). Given that only a small number of colonies were assayed, these results demonstrate the power of using the AICS pre-screening tool for rapid engineering of iminium biocatalysis.



**Figure 5. Activity for the Michael-type addition of **2** to **1a** catalyzed by CFEs of 92 stained and 92 unstained colonies.** Colonies from cells that were freshly transformed with library 21-40\_IA were stained using the AICS procedure. In total 92 stained colonies and 92 unstained colonies were picked, grown overnight and the CFE was assayed for ‘Michaelase’ activity using substrates **2** and **1a**. The ‘Michaelase’ activity was measured in a platereader by following the depletion of **1a** at 290 nm. Assay conditions: 5% ethanol, 30% CFE, 0.5 mM **1a**, 100 mM **2**, 50 mM sodium formate. Analysis of CFEs of unstained colonies by SDS PAGE showed that 8 out of 16 show significant 4-OT expression, indicating that the AICS procedure can efficiently be used to eliminate colonies not producing enzyme or producing enzyme with low-level activity.

## Discussion

The use and development of novel, higher throughput screening systems has drawn considerable attention<sup>30</sup>. As screening of large mutant libraries is often limiting enzyme engineering efforts, systems that reduce screening efforts, or allow for higher throughput of screening, have significantly attributed to the success of enzyme engineering studies<sup>25,31,32</sup>. The design of novel smart (pre-)screening techniques is therefore expected to have a strong impact. We wanted to contribute to this field by designing a smart iminium-based “turn on” probe that could be used to remove a considerable part of 4-OT mutants with no or low activity from large mutant libraries.

Although there are a few reports on “turn-on” probes that make use of an activated lysine in the acceptor protein<sup>33–35</sup>, these studies focused on non-catalytic proteins. Instead, we aimed to design the first iminium-activated “turn-on” probe for an enzyme, exploiting the characteristic nucleophilic N-terminal proline of 4-OT. In the past we showed that **1a** binds specifically to Pro-1 of 4-OT, forming **5a** (Figure 1d)<sup>18</sup>. Hence, we designed “turn-on” probes **1b** and **1c**, which are close mimics of substrate **1a**. Incubation of **1b** and **1c** with 4-OT M45T/F50A resulted in the formation of a novel relatively bathochromic absorption peak characteristic of merocyanine dyes. Especially incubation of “turn-on” probe **1b** with 4-OT M45T/F50A induced a remarkable spectroscopic shift that could easily be observed by the naked eye. Labeling studies further supported our hypothesis that Pro-1 of 4-OT reacts with **1b** to form merocyanine dye **5b**. Note that substrate **1a** and analogues **1d** and **1e**, which were prepared as controls, can not form merocyanine dye-type structures upon condensation with Pro-1 of 4-OT because they lack an o-OH or p-OH group (Figure 1d).

Merocyanine dye **5b** is very similar to **5a**, a key catalytic intermediate for the Michael-type addition of **2** to **1a**. Therefore we investigated if there is a correlation between 4-OT single mutants that upon incubation with probe **1b** show strong staining and 4-OT single mutants with improved ‘Michaelase’ activity. Indeed, we found that, from a subset of 4-OT single mutants, specifically mutants with enhanced ‘Michaelase’ activity showed strong staining upon incubation with **1b**. This inspired us to use “turn-on” probe **1b** as a pre-screening tool in the engineering of improved artificial ‘Michaelases’. We developed the AICS technique, a colorimetric solid-phase pre-screening procedure based on selective staining of colonies producing an active 4-OT mutant with **1b**, and used it to enrich large libraries for active 4-OT mutants. The developed AICS pre-screening technique proved to be highly effective, reducing the screening effort for a single round of directed evolution up to 20-fold. In addition, the sensitivity of AICS pre-screening

proved to be very high, as from a subset of 92 unstained colonies all showed less than 50% activity compared to the parental mutant.

After two rounds of directed evolution two triple mutants, 4-OT S37E/M45I/F50A and 4-OT A33E/M45I/F50A, were identified with a 29-fold and 39-fold improvement in activity compared to starting mutant 4-OT F50A. Semi-preparative scale reactions of the Michael-type addition of **2** to **1a** catalyzed by 4-OT S37E/M45I/F50A and 4-OT A33E/M45I/F50A resulted in isolation of target product **R-3** with excellent yield (up to 95%) and enantiopurity (*ee* up to >99%).

The developed AICS pre-screening technique provides a novel application of “turn-on” probes, adding to the usage of “turn-on” probes as tools for non-invasive tissue imaging or colorimetric pH sensors<sup>36-38</sup>. Contrary to many reported colorimetric assays (e.g. <sup>39-42</sup>), AICS does not rely on a (coupled) chemical reaction producing a chromogenic product, but instead solely relies on a conjugate formed between the probe and the enzyme. As such, AICS is very quick and nicely complements available reported colorimetric assays.

It would be interesting to investigate the application of AICS to a broader range of iminium biocatalysts, as preliminary results indicate that “turn-on” probe **1b** was also activated by the lysine-dependent aldolase DERA. The further development of “turn-on” probes as tools for enzyme engineering potentially opens up new possibilities for the efficient engineering of novel classes of iminium-based biocatalysts.

## Experimental Section

### Spectrophotometric analysis of **1b-e** incubated with purified 4-OT

A stock solution of 40 mM **1b-e** was prepared in 100% DMSO. 25  $\mu$ L was added to a solution of 20 mM sodium phosphate buffer (pH 7.3) with and without 50  $\mu$ M 4-OT M45T/F50A, 1 mL final volume. The mixture was transferred to a 1 cm quartz cuvet and a spectrum from 350 – 800 nm was recorded. To analyze the binding behavior of **1b** to 4-OT M45T/F50A a binding curve was constructed. **1b** (4 mM – 0 mM, 2-fold dilution series in 20 mM sodium phosphate pH 7.3, 5% DMSO) was incubated with 4-OT M45T/F50A (50  $\mu$ M) for 15 minutes (300  $\mu$ L final volume). The mixtures were transferred to 1 mm quartz cuvetts and the absorbance at 516 nm was recorded and plotted against the concentration of **1b**. The experiment was performed in triplo. SigmaPlot was used to fit the data against the one-site saturation model and to calculate the apparent  $K_D$  value (Figure S2).



### Covalent labeling of **1b** to 4-OT

To gain further evidence for the formation of **5b**, 4 mM of **1b** was incubated with 150  $\mu$ M 4-OT M45T/F50A in 10 mM sodium phosphate (pH 7.3), 5% DMSO at room temperature, 500  $\mu$ L total volume. After 15 minutes, 167  $\mu$ L 100 mM NaCNBH<sub>3</sub> in H<sub>2</sub>O was added to reduce any iminium species. The reaction was incubated at room temperature for 30 minutes. The buffer was exchanged to 10 mM ammonium formate in H<sub>2</sub>O using a pre-packed PD-10 sephadex G-25 gel filtration column. A sample of the collected enzyme was analyzed by ESI-MS (Figure S3). From the remaining fraction, 36  $\mu$ L was digested with GluC (New England BioLabs, 4  $\mu$ L from a 100 ng/ $\mu$ L) in a total volume of 80  $\mu$ L for 24 h at 37 °C using the supplied buffer. A sample from the digested mixture was analyzed by ESI-MS. As controls, a reaction without **1b** and a reaction without the addition of NaCNBH<sub>3</sub> (no reduction) were included. In both cases no modified 4-OT could be detected by ESI-MS (Figure S3 and S4).

### Staining of 4-OT single mutants with **1b**

From a reported set of 4-OT single mutants<sup>16</sup>, BL21 DE3 glycerol stocks of all single mutants with a mutation at position 6, 45 or 50 were used to inoculate 1 mL of LB medium supplemented with ampicillin (100  $\mu$ g/mL) and IPTG (100  $\mu$ M) and grown overnight at 37 °C in 96-deep well plates (Greiner Bio-one, 96-well Masterblock), covered with sterile gas-permeable seals (Greiner Bio-one, BREATHseal). The cells were harvested by centrifugation and the supernatant was removed. The cell pellets were re-suspended in 200  $\mu$ L 20 mM sodium phosphate buffer (pH 7.3) and 50  $\mu$ L was transferred to a round bottom 96-well plate (96 well U-bottom, Greiner Bio-One). To this 2.5  $\mu$ L of 40 mM **1b** in 100% DMSO was added. The cells were allowed to settle on the bottom of the 96-well plate.

### Cloning and mutagenesis

Triple-site library 4-OT Q4X/L8X/M45X/F50A was constructed by PCR with degenerate primers. The first part of the 4-OT gene was amplified using the following primers: 5'-GGAGATATACATATGCCTATTGCC**NKATCCACATC**NKGAAGGCCGCAGCGACG****-3' and 5'-CAGGGAGCGCGAGATGGCC-3'. In bold are indicated the mutated codons Gln-4 and Leu-8. Met-45 was randomized using primers 5'-GTGATTATCACGGAG**NKGCCAAGGGCCAC**-3' and 5'-GTGGCCCTTGG**C**MNNCTCCGTGATAATCAC****-3' by means of quikchange technology. DNA of 4-OT F50A was used as a template. In bold is indicated the mutated codon Met-45. The second part of 4-OT was amplified from the quikchange product using the following primers: 5'-GGCCATCTCGCGCTCCCTG-3' and 5'-GCTAGGGGGATCCTCAGCGTCTGACCTTGCTGGCCAGTTCCGCCG-3', using the quikchange product as template. The amplified parts

of the 4-OT gene were combined by overlap extension PCR and further amplified using primers 5'-TTGAAGGAGATATACATATGCCT-3' and 5'-GCTAGGGGGATCCTCAGCG-3'. An empty pET20b backbone was generated by amplification of a pET20b plasmid using the following primers: 5'-TGGGCAATAGGCATATGTATATCTCCTTCTTAAGTTAAAC-3' and 5'-AAGGTCAGACGCTGAGGATCCGAATTCGAGCTCCG-3'. The assembled 4-OT gene and the empty pET20b backbone were purified and digested with restriction enzymes NdeI, BamHI and DpnI. The empty pET20b backbone was dephosphorylated using alkaline phosphatase. The digested DNA products were purified and ligated by T4 DNA ligase. The ligation mixture was purified and used to electroporate electrocompetent *E. coli* Dh5 $\alpha$  cells. After 1 h of outgrowth with SOC medium a small sample of the transformed dh5 $\alpha$  cells was spread on an agar plate to determine transformation efficiency. The remainder of the transformation mixture was added to 5 mL LB medium, supplemented with ampicillin (100  $\mu$ g/mL) and grown overnight. The overnight culture was harvested and the plasmids were isolated and used to transform BL21 DE3 cells.

For the construction of library 2-20\_IA a previously reported set of nearly all single mutants was used<sup>16</sup>. From a glycerol stock, each DH10B culture that carried a plasmid harboring a 4-OT gene with a mutation at position 2-20 was separately grown overnight in 300  $\mu$ L LB medium supplemented with ampicillin (100  $\mu$ g/mL). From all individual cultures 100  $\mu$ L was mixed together. From a sample of this mixed culture plasmid DNA was isolated to obtain a library of 4-OT genes with a random single mutation at position 2-20. The second part of the 4-OT gene was amplified using the following primers: '5-GATTATCACGGAGATT**GTCCAAGGGCCACGCC**GCATCGGCGGCG-3' and 5'-GCTAGGGGGATCCTCAGCG-3'. In bold are indicated the mutated residues Ile-45 and Ala-50. Using the library of 4-OT genes with a random single mutation at position 2-20 as a template, the first part of the 4-OT gene was amplified using the following primers: '5-TTGAAGGAGATATACATATGCCT-3' and 5'-CGCCGCCGATGCCG**GCGTGGCCCTTGGCAATCTCCGTGATAATC**-3'. Using overlap extension PCR a full length 4-OT gene was obtained containing a random mutation at position 2-20, combined with the M45I and F50A mutation. The full length 4-OT gene was digested, ligated and cloned into the pET20b plasmid according to the procedure described for the construction of the Q4X/L8X/M45X/F50A library, to obtain library 2-20\_IA. The same procedure was followed for the construction of library 21-40\_IA, using the single mutants with a mutation at position 21-40.

### Expression and purification of enzymes

All 4-OT mutants were expressed and purified according to a reported procedure<sup>7</sup>. DERA with C-terminal His-tag was expressed and purified from *E. coli* using the pET20b(+) expression system. A colony of freshly transformed BL21 DE3 cells was used to inoculate 5 mL of LB medium supplemented with 100 µg/mL ampicillin and grown for 8 h at 37 °C. This starter culture was used to inoculate 100 mL of LB medium supplemented with ampicillin (100 µg/mL) and lactose (0.2% w/v). The cells were incubated at 30 °C overnight at 200 rpm shaking. Cells were harvested by centrifugation, re-suspended in 30 mL 20 mM potassium phosphate buffer (pH 7.0) and lysed by sonication. The lysate was cleared by centrifugation and transferred to a clean tube. Ni-sepharose (0.5 mL) was added to the lysate and incubated for 4 °C under slow rotation for 30 min. The lysate with Ni-sepharose was loaded on a column, washed with 8 mL of 20 mM potassium phosphate buffer (pH 7.0), followed by 8 mL of 20 mM potassium phosphate buffer (pH 7.0) containing 30 mM imidazole. Final elution was performed with 2.5 mL of 20 mM potassium phosphate buffer (pH 7.0) containing 250 mM imidazole (>95% purity as assessed by SDS PAGE). The buffer was exchanged to 20 mM potassium phosphate buffer (pH 7.0) using a pre-packed PD-10 sephadex G-25 gelfiltration column. The concentration of purified enzyme was determined by the Waddell method<sup>43</sup>. DERA K167L was purified according to the same protocol. Purified DERA and DERA K167L were immediately after purification used in subsequent experiments.

### Activated Iminium Colony Staining (AICS)

BL21 DE3 cells were transformed with library DNA and plated on LB agar plates supplemented with ampicillin (100 µg/mL) and lactose (0.2% w/v). Colonies were grown for ~16 h at 37 °C, after which the agar plate was stored at room temperature. A solution of 1.5% agarose in H<sub>2</sub>O was prepared by heating in a microwave, and the solution was kept at 60 °C. A second solution of 1 mM **1b** dissolved in 20 mM sodium phosphate (pH 7.3) was filter sterilized and stored at room temperature. 7 mL of the 1.5% agarose solution was mixed with 7 mL of the solution containing 1 mM **1b** in 20 mM sodium phosphate and immediately after mixing poured on the agar plate. Typically, after 10 seconds staining of the first colonies was observed. The plate was incubated at room temperature for 5 minutes to complete the staining and allow the agarose to solidify. The staining is stable for at least several hours. For screening of library 2-20\_IA and 21-40\_IA the concentration of **1b** in the second solution was reduced to 0.5 mM to increase the stringency of selection.

### **Library screening**

Colonies that showed clear staining after the AICS assay were picked with sterile tooth picks and were used to inoculate 1 mL of LB medium supplemented with ampicillin (100 µg/mL) and lactose (0.2 % w/v) in 96-deep well plates (Greiner Bio-one, 96-well Masterblock). Two wells in the 96-deep well plate were used to inoculate the parental mutant (4-OT F50A or 4-OT M45I/F50A). The plates were sealed with sterile gas-permeable seals (Greiner Bio-one, BREATHseal) and incubated at 37 °C, with overnight shaking at 250 rpm. 80 µL of the culture was mixed with 20 µL of a sterile solution of 80% glycerol in H<sub>2</sub>O and stored at -80 °C for later reference. The remainder of the culture was centrifuged. The supernatant was discarded and the pellets were re-suspended with 200 µL BugBuster (Novagen) supplemented with 15 U/mL benzonase (Novagen). After 20 minutes of incubation at room temperature the lysates were cleared by centrifugation after which the CFE was obtained. For the screening of library 2-20\_IA and 21-40\_IA, prior to centrifugation the lysed cells were diluted with 200 µL of sterile H<sub>2</sub>O. For the screening of library Q4X/L8X/M45X/F50A, the final reaction mixture for monitoring the addition of **2** to **1a** consisted of 40% CFE, 25 mM **2**, 250 µM **1a**, 5% ethanol in 20 mM HEPES pH 6.5, 100 µL final volume. For the screening of library 2-20\_IA and 21-40\_IA, the final reaction mixture for monitoring the addition of **2** to **1a** consisted of 30% CFE, 100 mM **2**, 1 mM **1a**, 50 mM sodium formate, 5% ethanol in 20 mM HEPES pH 6.5, 100 µL final volume. Sodium formate was added to the reaction mixture, as this was found to positively affect the catalytic rate of 4-OT M45I/F50A. The reaction was performed in a 96-well plate (Greiner Bio-one, UV-star F-bottom microplate) and reaction progress was followed in a platereader by following the depletion in absorbance at 290 nm. For the screening of library 2-20\_IA and 21-40\_IA the initial absorbance was too high for the platereader, so the time point at which the absorbance went below the absorbance threshold of the platereader was used as reference point for activity. For the comparison of activity of 92 stained mutants to 92 unstained mutants, the same procedure as described for library 21-40\_IA was followed. However, the reaction mixtures consisted of 0.5 mM **1a** instead of 1 mM **1a**.

### **Progress curves of the enzymatic reactions**

For monitoring the progress of the enzymatic addition of **2** to **1a** the following reaction setup was used. To a solution of ethanol, **1a** and **2** in 20 mM HEPES (pH 6.5) purified 4-OT F50A, 4-OT M45I/F50A, 4-OT S37E/M45I/F50A or 4-OT A33E/M45I/F50A was added. The final reaction mixture consisted of the following: 5% ethanol, 20 mM HEPES (pH 6.5), 1 mM **1a**, 100 mM **2**, 50 mM sodium formate, 20 µM 4-OT. The reaction mixtures were transferred to 1 mm quartz cuvetts and the absorption at 290 nm was recorded every 20 seconds.

### Semi-preparative scale reactions

Semi-preparative scale experiments were performed using 4-OT S37E/M45I/F50A and 4-OT A33E/M45I/F50A as biocatalysts. The reaction mixtures consisted of the following: 5 mL ethanol (5% v/v), 20 mg **1a**, 25 mM **2**, 50 mM sodium formate, 20 mM HEPES (pH 6.5), 7.5  $\mu$ M enzyme, 100 mL final volume. At timely intervals a sample was withdrawn from the reaction mixture and the reaction progress was monitored by following the depletion in absorbance at 290 nm. After 10.5 h for 4-OT S37E/M45I/F50A and 7.5 h for 4-OT A33E/M45I/F50A the reaction was finished and the reaction mixture was extracted 3x with 20 mL ethyl acetate. The combined organic layers were washed with brine and dried over anhydrous Na<sub>2</sub>SO<sub>4</sub>. For the reaction catalyzed by S37E/M45I/F50A, the organic layer was concentrated in vacuo, yielding **3** (27.8 mg, 95% isolated yield). For the 4-OT A33E/M45I/F50A catalyzed reaction the organic layer was concentrated in vacuo and the product was purified by silica gel column chromatography (hexane/ethyl acetate 4:1) to obtain **3** (16.0 mg, 55% yield).

### Condensation of **1b** to DERA and to 4-OT

The condensation of **1b** with DERA and 4-OT was monitored in time by following the absorbance at 516 nm. For the condensation with DERA and DERA K167L, a 2x stock solution was prepared consisting of the following: 20 mM potassium phosphate (pH 7.0), 10% DMSO and 8 mM **1b**. 500  $\mu$ L of the stock solution was mixed in a 1 cm quartz cuvet with 500  $\mu$ L of 200  $\mu$ M DERA or DERA K167L (5.54 mg/mL) in 20 mM potassium phosphate (pH 7.0). Immediately after mixing the absorbance at 516 nm was measured every 0.5 s for 190 seconds. For 4-OT, a 2x stock solution was prepared consisting of the following: 20 mM HEPES (pH6.5), 10% ethanol and 0.6 mM **1b**. 500  $\mu$ L of the stock solution was mixed in a 1 cm quartz cuvet with 500  $\mu$ L of 100  $\mu$ M 4-OT in 20 mM HEPES pH 6.5. Immediately after mixing the absorbance at 516 nm was measured every 0.1 s for 190 seconds.

### Derivatization of **3** for enantiomeric excess determination

The aldehyde functionality of **3** was derivatized to a cyclic acetal according to a literature procedure<sup>44,45</sup>. The enantiopurity of derivatized **3** was analyzed by reverse phase HPLC using a column with a chiral stationary phase (chiralpak AD-RH, 150 mm x 4.6 mm, Daicel). Detection at 220 nm, retention time *R*-**3**: 8.8 min, *S*-**3**: 12.2 min. The absolute configuration was determined by literature comparison<sup>16</sup>.

## **Associated Content**

### **Supporting Information**

Synthesis of starting materials and reference compounds, additional Tables and Figures including ESI-MS, NMR, HPLC and UV-VIS spectra.

### **Acknowledgement**

We acknowledge financial support from the Netherlands Organization of Scientific Research (ECHO grant 713.015.003 and VICI grant 724.016.002), the European Research Council (PoC grant 713483), and the European Union’s Horizon 2020 Research and Innovation Programme under grant agreement No 635595 (CarbaZymes). We thank M.P. de Vries for his assistance in performing the electrospray ionization mass spectrometry experiments.

### **Author contributions**

L.B. performed 4-OT labeling experiments, developed the AICS screening assay, and performed experiments with DERA. L.B. and M.C. performed the screening of mutant libraries. M.S. synthesized cinnamaldehyde derivatives. G.J.P. supervised scientific work. All authors contributed to writing of the manuscript.

### **Competing financial interests**

The authors declare no competing financial interests

## References

1. Claraz, A.; Siitonen, J. H.; Pihko, P. M. Iminium Catalysis ( $n^{\pi} \rightarrow \pi^*$ ). In *Lewis Base Catalysis in Organic Synthesis*; Vedejs, E.; Denmark, S.E., Eds.; Wiley-VCH: Weinheim, Germany, **2016**; pp 805–856.
2. Samland, A. K.; Rale, M.; Sprenger, G. A.; Fessner, W.-D. The Transaldolase Family: New Synthetic Opportunities from an Ancient Enzyme Scaffold. *ChemBioChem* **2011**, *12* (10), 1454–1474.
3. Clapés, P.; Garrabou, X. Current Trends in Asymmetric Synthesis with Aldolases. *Adv. Synth. Catal.* **2011**, *353* (13), 2263–2283, <https://doi.org/10.1002/adsc.201100236>.
4. Fei, H.; Zheng, C.; Liu, X.; Li, Q. An Industrially Applied Biocatalyst: 2-Deoxy-d-Ribose-5-Phosphate Aldolase. *Process Biochem.* **2017**, *63*, 55–59.
5. Jiang, L.; Althoff, E. A.; Clemente, F. R.; Doyle, L.; Röthlisberger, D.; Zanghellini, A.; Gallaher, J. L.; Betker, J. L.; Tanaka, F.; Barbas, III, C. F.; Hilvert, D.; Houk, K. N.; Stoddard, B. L.; Baker, D. De Novo Computational Design of Retro-Aldol Enzymes. *Science* **2008**, *319* (5868), 1387 – 1391.
6. Althoff, E. A.; Wang, L.; Jiang, L.; Giger, L.; Lassila, J. K.; Wang, Z.; Smith, M.; Hari, S.; Kast, P.; Herschlag, D.; Hilvert, D.; Baker, D. Robust Design and Optimization of Retroaldol Enzymes. *Protein Sci.* **2012**, *21* (5), 717–726.
7. Zandvoort, E.; Baas, B.; Quax, W. J.; Poelarends, G. J. Systematic Screening for Catalytic Promiscuity in 4-Oxalocrotonate Tautomerase: Enamine Formation and Aldolase Activity. *ChemBioChem* **2011**, *12* (4), 602–609.
8. Poddar, H.; Rahimi, M.; Geertsema, E. M.; Thunnissen, A. M. W. H.; Poelarends, G. J. Evidence for the Formation of an Enamine Species during Aldol and Michael-Type Addition Reactions Promiscuously Catalyzed by 4-Oxalocrotonate Tautomerase. *ChemBioChem* **2015**, *16* (5), 738-741.
9. Garrabou, X.; Beck, T.; Hilvert, D. A Promiscuous De Novo Retro-Aldolase Catalyzes Asymmetric Michael Additions via Schiff Base Intermediates. *Angew. Chemie, Int. Ed.* **2015**, *54* (19), 5609–5612.
10. Giger, L.; Caner, S.; Obexer, R.; Kast, P.; Baker, D.; Ban, N.; Hilvert, D. Evolution of a Designed Retro-Aldolase Leads to Complete Active Site Remodeling. *Nat. Chem. Biol.* **2013**, *9*, 494–498.
11. Obexer, R.; Godina, A.; Garrabou, X.; Mittl, P. R. E.; Baker, D.; Griffiths, A. D.; Hilvert, D. Emergence of a Catalytic Tetrad during Evolution of a Highly Active Artificial Aldolase. *Nat. Chem.* **2017**, *9* (1), 50–56.
12. Garrabou, X.; Wicky, B. I. M.; Hilvert, D. Fast Knoevenagel Condensations Catalyzed by an Artificial Schiff-Base-Forming Enzyme. *J. Am. Chem. Soc.* **2016**, *138* (22), 6972–6974.
13. Garrabou, X.; MacDonald, D. S.; Hilvert, D. Chemoselective Henry Condensations Catalyzed by Artificial Carboligases. *Chem. - Eur. J.* **2017**, *23* (25), 6001–6003.

14. Rahimi, M.; van der Meer, J.Y.; Geertsema, E. M.; Poelarends, G. J. Engineering a Promiscuous Tautomerase into a More Efficient Aldolase for Self-Condensations of Linear Aliphatic Aldehydes. *ChemBioChem* **2017**, *18* (14), 1435–1441.
15. Rahimi, M.; van der Meer, J.Y.; Geertsema, E. M.; Poddar, H.; Baas, B. J.; Poelarends, G. J. Mutations Closer to the Active Site Improve the Promiscuous Aldolase Activity of 4-Oxalocrotonate Tautomerase More Effectively than Distant Mutations. *ChemBioChem* **2016**, *17* (13).
16. van der Meer, J.Y.; Poddar, H.; Baas, B.J.; Miao, Y.; Rahimi, M.; Kunzendorf, A.; van Merkerk, R.; Tepper, P. G.; Geertsema, E. M.; Thunnissen, A.M. W. H.; Quax, W. J.; Poelarends, G. J. Using Mutability Landscapes of a Promiscuous Tautomerase to Guide the Engineering of Enantioselective Michaelases. *Nat. Commun.* **2016**, *7*, 10911.
17. Biewenga, L.; Saravanan, T.; Kunzendorf, A.; van der Meer, J.Y.; Pijning, T.; Tepper, P. G.; van Merkerk, R.; Charnock, S. J.; Thunnissen, A.M. W. H.; Poelarends, G. J. Enantioselective Synthesis of Pharmaceutically Active  $\gamma$ -Aminobutyric Acids Using a Tailor-Made Artificial Michaelase in One-Pot Cascade Reactions. *ACS Catal.* **2019**, *9* (2), 1503–1513.
18. Guo, C.; Saifuddin, M.; Saravanan, T.; Sharifi, M.; Poelarends, G. J. Biocatalytic Asymmetric Michael Additions of Nitromethane to  $\alpha,\beta$ -Unsaturated Aldehydes via Enzyme-Bound Iminium Ion Intermediates. *ACS Catal.* **2019**, *9* (5), 4369–4373.
19. Goldsmith, M.; Tawfik, D. S. Enzyme Engineering by Targeted Libraries. In *Methods in Protein Design*; Keating, A. E., Ed.; Academic Press, 2013; Vol. 523, pp 257–283.
20. Acevedo-Rocha, C.G.; Ferla, M.; Reetz, M.T. Directed Evolution of Proteins Based on Mutational Scanning. In *Methods in Protein Engineering*; Bornscheuer U., Höhne M., eds.; Humana Press, New York, **2018**; Vol. 1685, pp 87–128.
21. van der Meer, J. Y.; Biewenga, L.; Poelarends, G. J. The Generation and Exploitation of Protein Mutability Landscapes for Enzyme Engineering. *ChemBioChem* **2016**, *17* (19), 1792–1799.
22. Yang, G.; Withers, S. G. Ultrahigh-Throughput FACS-Based Screening for Directed Enzyme Evolution. *ChemBioChem* **2009**, *10* (17), 2704–2715.
23. Baret, J.C.; Miller, O. J.; Taly, V.; Ryckelynck, M.; El-Harrak, A.; Frenz, L.; Rick, C.; Samuels, M. L.; Hutchison, J. B.; Agresti, J. J.; Link, D. R.; Weitz, D. A.; Griffiths, A. D. Fluorescence-Activated Droplet Sorting (FADS): Efficient Microfluidic Cell Sorting Based on Enzymatic Activity. *Lab Chip* **2009**, *9* (13), 1850–1858.
24. Mair, P.; Gielen, F.; Hollfelder, F. Exploring Sequence Space in Search of Functional Enzymes Using Microfluidic Droplets. *Curr. Opin. Chem. Biol.* **2017**, *37*, 137–144.
25. Obexer, R.; Pott, M.; Zeymer, C.; Griffiths, A. D.; Hilvert, D. Efficient Laboratory Evolution of Computationally Designed Enzymes with Low Starting Activities Using Fluorescence-Activated Droplet Sorting. *Protein Eng. Des. Sel.* **2016**, *29* (9), 355–366.
26. Kulinich, A. V.; Ishchenko, A. A. Merocyanine Dyes: Synthesis, Structure, Properties and Applications. *Russ. Chem. Rev.* **2009**, *78* (2), 141–164.
27. Mishra, A.; Behera, R. K.; Behera, P. K.; Mishra, B. K.; Behera, G. B. Cyanines during the 1990s: A Review. *Chem. Rev.* **2000**, *100* (6), 1973–2012.



28. Haridas, M.; Abdelraheem, E. M. M.; Hanefeld, U. 2-Deoxy-d-Ribose-5-Phosphate Aldolase (DERA): Applications and Modifications. *Appl. Microbiol. Biotechnol.* **2018**, *102* (23), 9959–9971.
29. Whitman, C. P. The 4-Oxalocrotonate Tautomerase Family of Enzymes: How Nature Makes New Enzymes Using a  $\beta$ - $\alpha$ - $\beta$  Structural Motif. *Arch. Biochem. Biophys.* **2002**, *402* (1), 1–13.
30. Mugge, C.; Kourist, R. Practical Considerations Regarding the Choice of the Best High-Throughput Assay. In *Protein engineering methods and protocols*; Bornscheuer, U. T., Hohne, M., Eds.; Humana Press: New York, 2018; pp 189–208.
31. Aharoni, A.; Thieme, K.; Chiu, C. P. C.; Buchini, S.; Lairson, L. L.; Chen, H.; Strynadka, N. C. J.; Wakarchuk, W. W.; Withers, S. G. High-Throughput Screening Methodology for the Directed Evolution of Glycosyltransferases. *Nat. Methods* **2006**, *3* (8), 609–614.
32. Lauchli, R.; Rabe, K. S.; Kalbarczyk, K. Z.; Tata, A.; Heel, T.; Kitto, R. Z.; Arnold, F. H. High-Throughput Screening for Terpene-Synthase-Cyclization Activity and Directed Evolution of a Terpene Synthase. *Angew. Chemie Int. Ed.* **2013**, *52* (21), 5571–5574.
33. Wang, W.; Nossoni, Z.; Berbasova, T.; Watson, C. T.; Yapici, I.; Lee, K. S. S.; Vasileiou, C.; Geiger, J. H.; Borhan, B. Tuning the Electronic Absorption of Protein-Embedded All-*Trans*-Retinal. *Science* **2012**, *338* (6112), 1340 – 1343.
34. Crist, R. M.; Vasileiou, C.; Rabago-Smith, M.; Geiger, J. H.; Borhan, B. Engineering a Rhodopsin Protein Mimic. *J. Am. Chem. Soc.* **2006**, *128* (14), 4522–4523.
35. Vasileiou, C.; Vaezeslami, S.; Crist, R. M.; Rabago-Smith, M.; Geiger, J. H.; Borhan, B. Protein Design: Reengineering Cellular Retinoic Acid Binding Protein II into a Rhodopsin Protein Mimic. *J. Am. Chem. Soc.* **2007**, *129* (19), 6140–6148.
36. Herwig, L.; Rice, A. J.; Bedbrook, C. N.; Zhang, R. K.; Lignell, A.; Cahn, J. K. B.; Renata, H.; Dodani, S. C.; Cho, I.; Cai, L.; Gradinaru, V.; Arnold, F. H. Directed Evolution of a Bright Near-Infrared Fluorescent Rhodopsin Using a Synthetic Chromophore. *Cell Chem. Biol.* **2017**, *24* (3).
37. Berbasova, T.; Nosrati, M.; Vasileiou, C.; Wang, W.; Lee, K. S. S.; Yapici, I.; Geiger, J. H.; Borhan, B. Rational Design of a Colorimetric PH Sensor from a Soluble Retinoic Acid Chaperone. *J. Am. Chem. Soc.* **2013**, *135* (43), 16111–16119.
38. Yapici, I.; Lee, K. S. S.; Berbasova, T.; Nosrati, M.; Jia, X.; Vasileiou, C.; Wang, W.; Santos, E. M.; Geiger, J. H.; Borhan, B. “Turn-On” Protein Fluorescence: In Situ Formation of Cyanine Dyes. *J. Am. Chem. Soc.* **2015**, *137* (3), 1073–1080.
39. Ferrari, A. R.; Gaber, Y.; Fraaije, M. W. A Fast, Sensitive and Easy Colorimetric Assay for Chitinase and Cellulase Activity Detection. *Biotechnol. Biofuels* **2014**, *7* (1), 37.
40. Khersonsky, O.; Tawfik, D. S. Chromogenic and Fluorogenic Assays for the Lactonase Activity of Serum Paraoxonases. *ChemBioChem* **2006**, *7* (1), 49–53.
41. Thai, Y. C.; Szekrenyi, A.; Qi, Y.; Black, G. W.; Charnock, S. J.; Fessner, W. D. Fluorogenic Kinetic Assay for High-Throughput Discovery of Stereoselective Ketoreductases Relevant to Pharmaceutical Synthesis. *Bioorg. Med. Chem.* **2018**, *26* (7), 1320–1326.

42. Baud, D.; Ladkau, N.; Moody, T. S.; Ward, J. M.; Hailes, H. C. A Rapid, Sensitive Colorimetric Assay for the High-Throughput Screening of Transaminases in Liquid or Solid-Phase. *Chem. Commun.* **2015**, 51 (97), 17225–17228.
43. Waddell, W. J. A Simple Ultraviolet Spectrophotometric Method for the Determination of Protein. *J. Lab. Clin. Med.* **1956**, 48 (2), 311–314.
44. Zandvoort, E.; Geertsema, E. M.; Baas, B.-J.; Quax, W. J.; Poelarends, G. J. Bridging between Organocatalysis and Biocatalysis: Asymmetric Addition of Acetaldehyde to  $\beta$ -Nitrostyrenes Catalyzed by a Promiscuous Proline-Based Tautomerase. *Angew. Chemie, Int. Ed.* **2012**, 51 (5), 1240–1243, S1240/1-S1240/13.
45. Geertsema, E. M.; Miao, Y.; Tepper, P. G.; Dehaan, P.; Zandvoort, E.; Poelarends, G. J. Biocatalytic Michael-Type Additions of Acetaldehyde to Nitroolefins with the Proline-Based Enzyme 4-Oxalocrotonate Tautomerase Yielding Enantioenriched  $\gamma$ -Nitroaldehydes. *Chem. - Eur. J.* **2013**, 19 (43), 14407–14410.

## Supporting Information

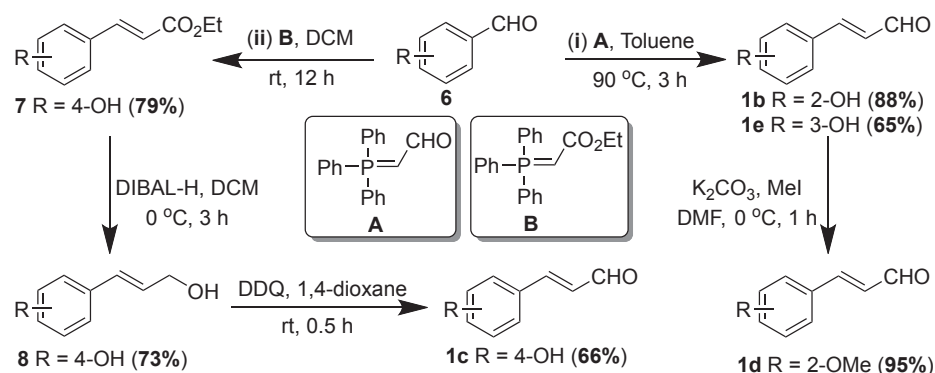
## Supporting methods

## Synthesis of chiral references

Racemic **3**, that served as a reference for analyses by HPLC on a chiral stationary phase, was synthesized according to a literature procedure<sup>1</sup>.

Synthesis of **1b-e**

Compounds **1b-e** were prepared according to a literature procedure<sup>2</sup> and their <sup>1</sup>H NMR spectra match with earlier reported NMR data<sup>3</sup>.



**Figure S1.** Synthesis of  $\alpha,\beta$ -unsaturated aldehydes **1b-e** via the Wittig reaction.

(*E*)-3-(2-hydroxyphenyl)acrylaldehyde (**1b**). <sup>1</sup>H NMR (500 MHz, CDCl<sub>3</sub>):  $\delta$  9.68 (d,  $J$  = 8.0 Hz, 1H), 7.79 (d,  $J$  = 16.0 Hz, 1H), 7.50 (dd,  $J$  = 7.8, 1.4 Hz, 1H), 7.31 (td,  $J$  = 8.1, 1.6 Hz, 1H), 7.07 – 6.93 (m, 2H), 6.89 (dd,  $J$  = 8.1, 0.8 Hz, 1H), 6.59 (s, 1H).

(*E*)-3-(4-hydroxyphenyl)acrylaldehyde (**1c**). <sup>1</sup>H NMR (500 MHz, DMSO-*d*<sub>6</sub>):  $\delta$  10.20 (s, 1H), 9.58 (d,  $J$  = 7.9 Hz, 1H), 7.67 – 7.57 (m, 3H), 6.84 (d,  $J$  = 8.6 Hz, 2H), 6.66 (dd,  $J$  = 15.8, 7.9 Hz, 1H).

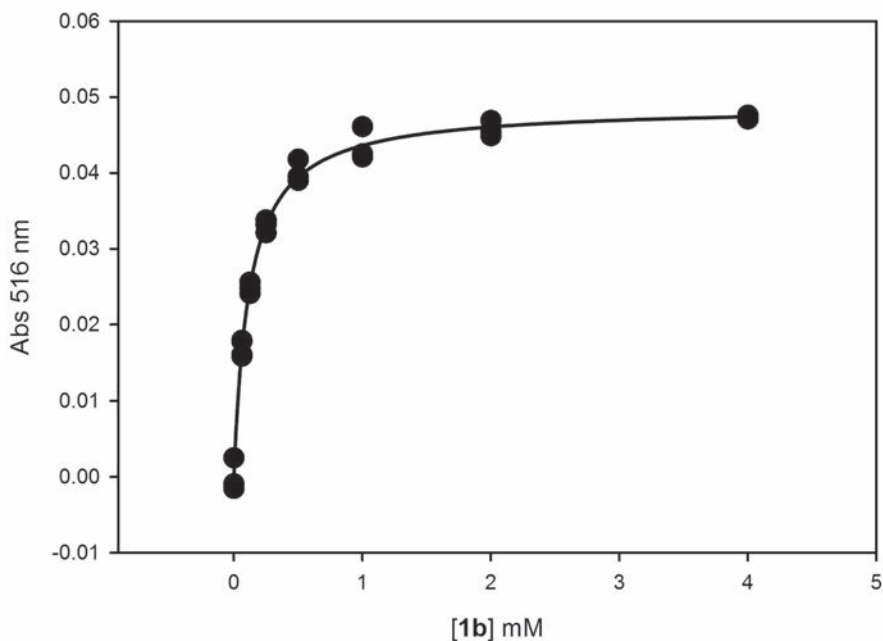
(*E*)-3-(2-methoxyphenyl)acrylaldehyde (**1d**). <sup>1</sup>H NMR (500 MHz, CDCl<sub>3</sub>):  $\delta$  9.67 (d,  $J$  = 7.9 Hz, 1H), 7.82 (d,  $J$  = 16.1 Hz, 1H), 7.52 (dd,  $J$  = 7.7, 1.6 Hz, 1H), 7.40 (ddd,  $J$  = 8.9, 7.5, 1.7 Hz, 1H), 6.98 (t,  $J$  = 7.5 Hz, 1H), 6.93 (d,  $J$  = 8.3 Hz, 1H), 6.77 (dd,  $J$  = 16.1, 7.9 Hz, 1H), 3.89 (s, 3H).

(*E*)-3-(3-hydroxyphenyl)acrylaldehyde (**1e**).  $^1\text{H}$  NMR (500 MHz, DMSO- $d_6$ )  $\delta$  9.72 (s, 1H), 9.65 (d,  $J = 7.8$  Hz, 1H), 7.66 (d,  $J = 15.9$  Hz, 1H), 7.27 (t,  $J = 7.8$  Hz, 1H), 7.18 (d,  $J = 7.7$  Hz, 1H), 7.08 (s, 1H), 6.89 (dd,  $J = 8.0, 2.0$  Hz, 1H), 6.74 (dd,  $J = 15.9, 7.8$  Hz, 1H).

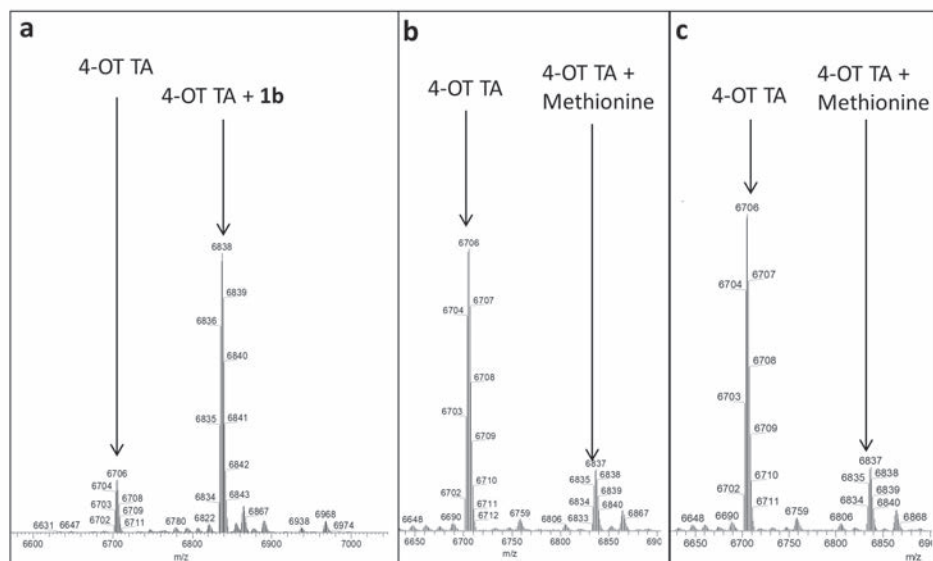
Ethyl (*E*)-3-(4-hydroxyphenyl)acrylate (**7**).  $^1\text{H}$  NMR (500 MHz,  $\text{CDCl}_3$ ):  $\delta$  7.64 (d,  $J = 16.0$  Hz, 1H), 7.40 (d,  $J = 8.6$  Hz, 2H), 7.02 (s, 1H), 6.88 (d,  $J = 8.6$  Hz, 2H), 6.29 (d,  $J = 15.9$  Hz, 1H), 4.27 (q,  $J = 7.1$  Hz, 2H), 1.34 (t,  $J = 7.1$  Hz, 3H).

(*E*)-4-(3-hydroxyprop-1-en-1-yl)phenol (**8**).  $^1\text{H}$  NMR (500 MHz, DMSO- $d_6$ ):  $\delta$  9.48 (s, 1H), 7.22 (d,  $J = 8.5$  Hz, 2H), 6.70 (d,  $J = 8.5$  Hz, 2H), 6.41 (d,  $J = 16.0$  Hz, 1H), 6.12 (dt,  $J = 15.9, 5.4$  Hz, 1H), 4.79 (t,  $J = 5.5$  Hz, 1H), 4.05 (t,  $J = 4.8$  Hz, 2H).

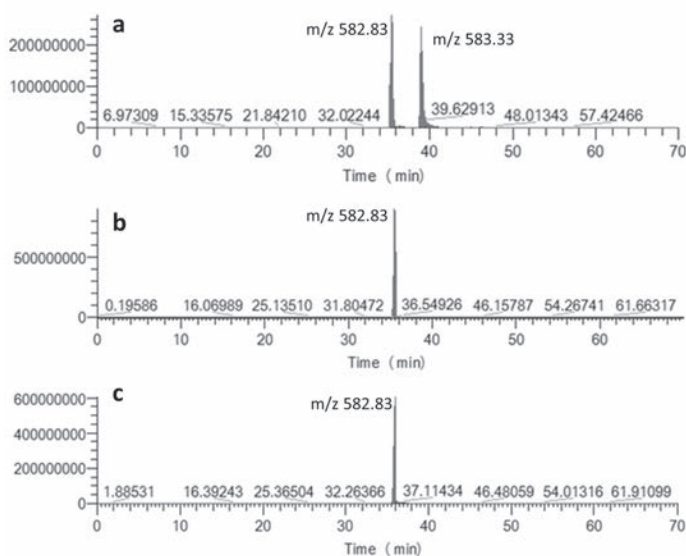
## Supporting figures



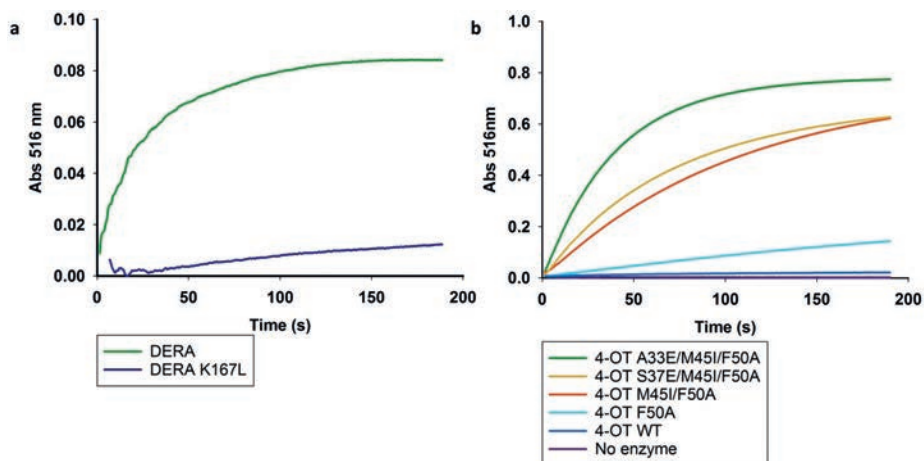
**Figure S2.** Binding curve of **1b** to 4-OT M45T/F50A. Following conditions were used: 50  $\mu\text{M}$  4-OT M45T/F50A, **1b** 4 mM – 0 mM, 2-fold dilution series, 20 mM sodium phosphate pH 7.3, 5% DMSO, 300  $\mu\text{l}$  reaction volume in 1 mm quartz cuvetts.



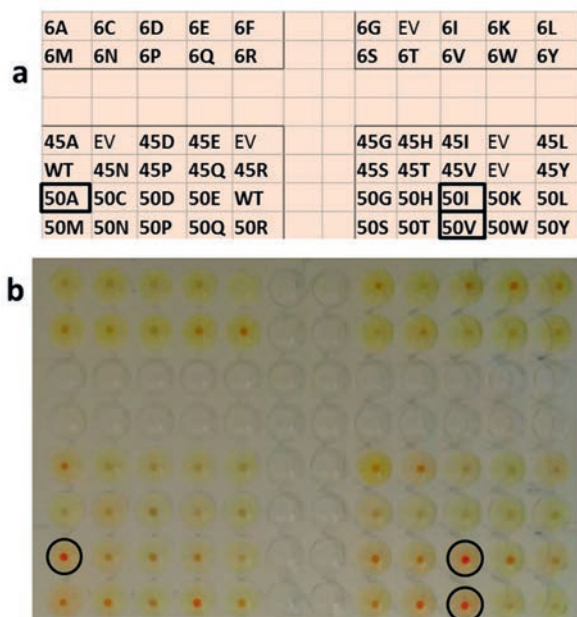
**Figure S3.** (a) ESI-MS spectrum of 4-OT M45T/F50A labeled with **1b** (expected mass free enzyme 6706 Da, expected mass labeled enzyme 6838 Da). (b) ESI-MS spectrum of 4-OT M45T/F50A no-reduction control (expected mass 6706 Da). (c) ESI-MS spectrum of unmodified 4-OT M45T/F50A treated with NaCNBH<sub>3</sub> (expected mass 6706 Da).



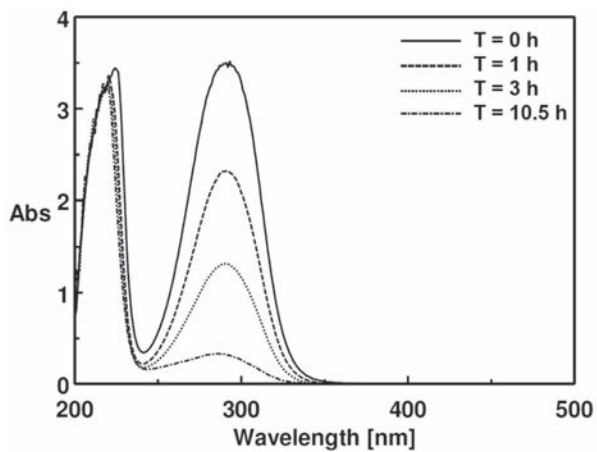
**Figure S4.** (a) ESI-MS spectrum of peptides of 4-OT M45T/F50A labeled with **1b** with  $m/z$  between 582 and 584 Da.  $m/z$  of 582.83 corresponds to the expected mass of peptide MPIAQIHILE (4-OT with an unprocessed N-terminal methionine),  $m/z$  of 583.33 corresponds to peptide PIAQIHILE labeled with a molecule of **1b**. (b) ESI-MS spectrum of peptides of 4-OT M45T/F50A no-reduction control. (c) ESI-MS spectrum of unmodified 4-OT M45T/F50A treated with NaCNBH<sub>3</sub>.



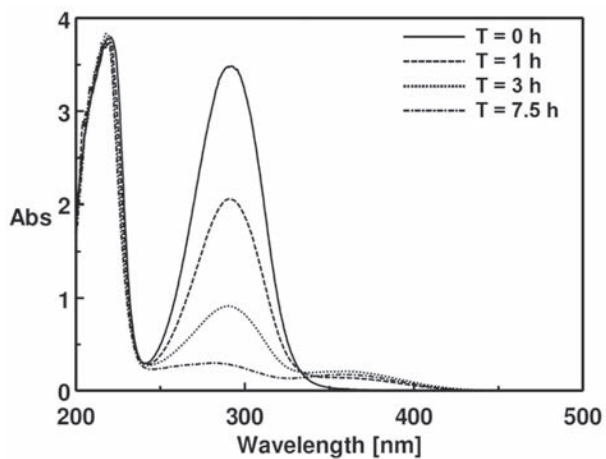
**Figure S5.** (a) The condensation of **1b** with DERA was monitored in time by following the increase in absorbance at 516 nm directly after mixing of the enzyme and **1b**. As a control, DERA K167L was used, an inactive mutant because of substitution of the active-site lysine<sup>4</sup>. Assay conditions: 20 mM potassium phosphate (pH 7.0), 5% DMSO, 4 mM **1b**, 100  $\mu$ M DERA, 1 mL in 1 cm quartz cuvetts. (b) The condensation of **1b** with different 4-OT mutants was monitored in time by following the increase in absorbance at 516 nm directly after mixing 4-OT with **1b**. Assay conditions: 20 mM HEPES (pH 6.5), 5% ethanol, 0.3 mM **1b**, 50  $\mu$ M 4-OT, 1 mL in 1 cm quartz cuvetts.



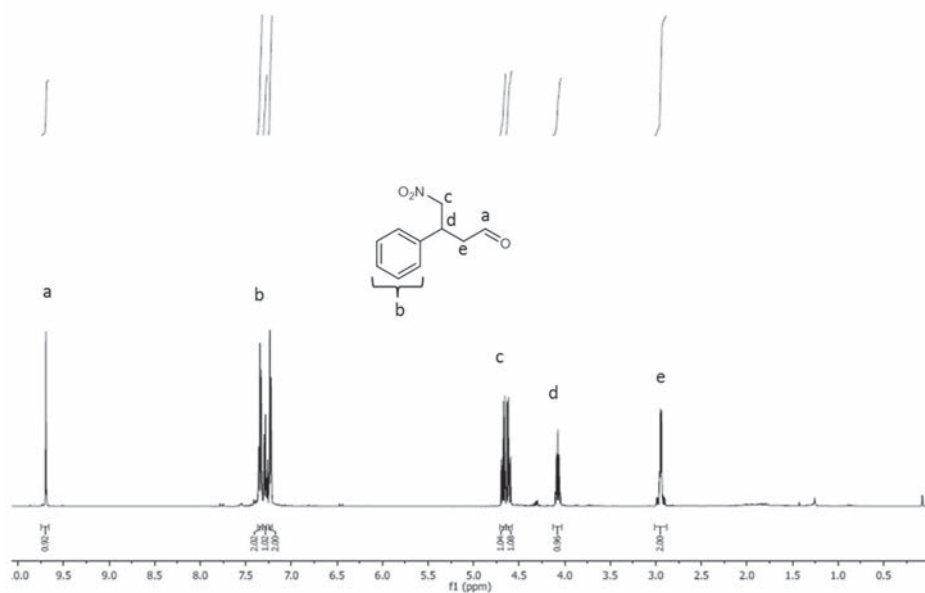
**Figure S6.** Whole cell staining of 4-OT single mutants with a mutation at position His-6, Met-45 or Phe-50. (a) Layout of 4-OT single mutants. EV: empty vector (b) Picture of cell pellets of 4-OT single mutants stained with **1b**. Indicated with a circle are 4-OT F50A, F50I and F50V.



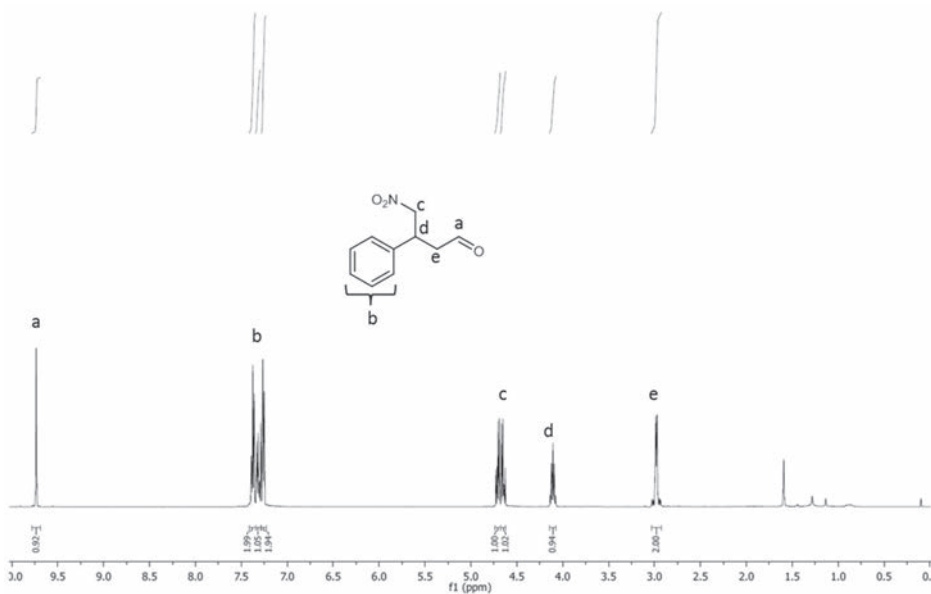
**Figure S7.** UV spectra showing the conversion of **1a** for the Michael-type addition of **2** to **1a** catalyzed by 4-OT S37E/M45I/F50A.



**Figure S8.** UV spectra showing the conversion of **1a** for the Michael-type addition of **2** to **1a** catalyzed by 4-OT A33E/M45I/F50A.

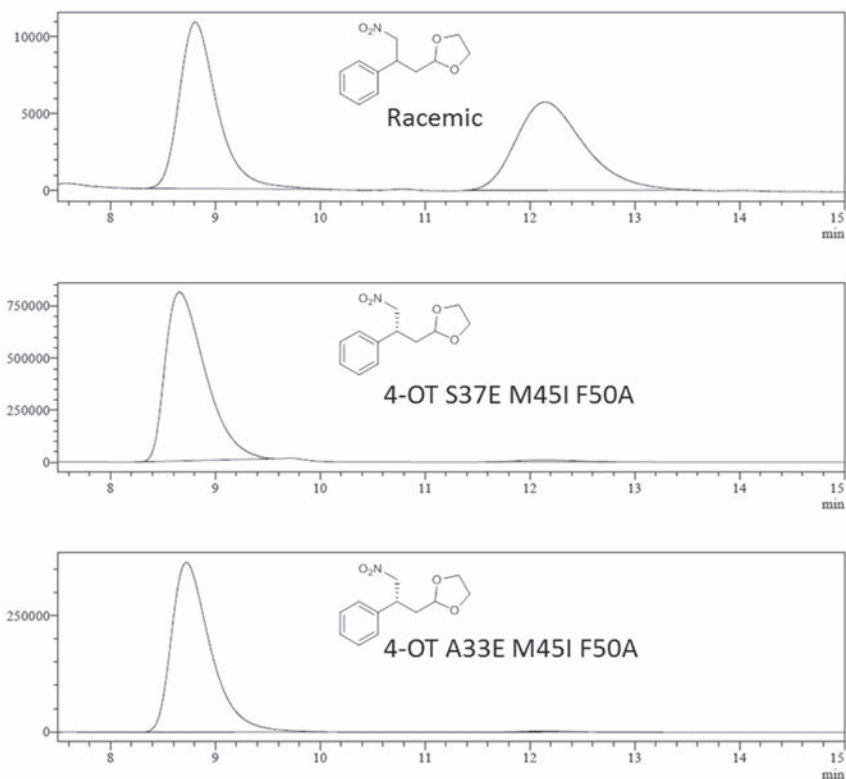


**Figure S9.**  $^1\text{H}$  NMR spectrum of 4-nitro-3-phenylbutanal (**3**), obtained from the 4-OT S37E/M45I/F50A catalyzed addition of **2** to **1a** (400 MHz,  $\text{CDCl}_3$ ).



**Figure S10.**  $^1\text{H}$  NMR spectrum of 4-nitro-3-phenylbutanal (**3**), obtained from the 4-OT A33E/M45I/F50A catalyzed addition of **2** to **1a** (400 MHz,  $\text{CDCl}_3$ ).





**Figure S11.** HPLC chromatogram of racemic and enzymatically prepared derivatized **3**.

## Supporting tables

**Table S1.** Michael-type addition of **2** to **1a** yielding **3**, catalyzed by 4-OT S37E/M45I/F50A and 4-OT A33E/M45I/F50A<sup>a</sup>.

Entry	Mutant	Reaction time (h) <sup>b</sup>	Isolated yield (%)	<i>ee</i> (%) <sup>c</sup>
1	4-OT S37E/M45I/F50A	10.5	95	98 (R)
2	4-OT A33E/M45I/F50A	7.5	55	>99 (R)

<sup>a</sup> Assay conditions: 5% ethanol, 20 mg **1a**, 25 mM **2**, 50 mM sodium formate, 20 mM HEPES (pH 6.5), 7.5  $\mu$ M enzyme, 100 mL. <sup>b</sup> Reaction progress was monitored by following the depletion in absorbance at 290 nm. <sup>c</sup> Enzymatic product **3** was first converted into the corresponding ethylene glycol acetal. The *ee* was determined by HPLC with a chiral stationary phase. The absolute configuration was based on literature comparison<sup>5</sup>.

## Supporting references

1. Geertsema, E. M.; Miao, Y.; Tepper, P. G.; Dehaan, P.; Zandvoort, E.; Poelarends, G. J. Biocatalytic Michael-Type Additions of Acetaldehyde to Nitroolefins with the Proline-Based Enzyme 4-Oxalocrotonate Tautomerase Yielding Enantioenriched  $\gamma$ -Nitroaldehydes. *Chem. - Eur. J.* **2013**, *19* (43), 14407–14410
2. Guo, C.; Saifuddin, M.; Saravanan, T.; Sharifi, M.; Poelarends, G. J. Biocatalytic Asymmetric Michael Additions of Nitromethane to  $\alpha,\beta$ -Unsaturated Aldehydes via Enzyme-Bound Iminium Ion Intermediates. *ACS Catal.* **2019**, *9* (5), 4369–4373,
3. Chuprajob, T.; Changtam, C.; Chokchaisiri, R.; Chunglok, W.; Sornkaew, N.; Suksamrarn, A. Synthesis, cytotoxicity against human oral cancer KB cells and structure–activity relationship studies of trienone analogues of curcuminoids. *Bioorg. Med. Chem. Lett.* **2014**, *24*, 2839–2844,
4. Heine, A.; DeSantis, G.; Luz, J. G.; Mitchell, M.; Wong, C. H.; Wilson, I. A. Observation of Covalent Intermediates in an Enzyme Mechanism at Atomic Resolution. *Science.* **2001** 294, 369–374,
5. van der Meer, J.Y.; Poddar, H.; Baas, B.J.; Miao, Y.; Rahimi, M.; Kunzendorf, A.; van Merkerk, R.; Tepper, P. G.; Geertsema, E. M.; Thunnissen, A.M. W. H.; Quax, W. J.; Poelarends, G. J. Using Mutability Landscapes of a Promiscuous Tautomerase to Guide the Engineering of Enantioselective Michaelases. *Nat. Commun.* **2016**, *7*, 10911,



# Chapter 5

## *In Situ Acetaldehyde Synthesis for Carboligation Reactions*

---

Lieuwe Biewenga<sup>†1</sup>, Andreas Kunzendorf<sup>†1</sup>, and Gerrit J. Poelarends<sup>\*1</sup>

<sup>1</sup>*Department of Chemical and Pharmaceutical Biology, Groningen Research Institute of Pharmacy,  
University of Groningen, Antonius Deusinglaan 1, 9713 AV Groningen, The Netherlands.*

<sup>†</sup>*These authors contributed equally to this work.*

<sup>\*</sup>*Corresponding author. Tel.: +31503633354; E-mail: g.j.poelarends@rug.nl;*

*Web: <http://www.rug.nl/staff/g.j.poelarends/>*

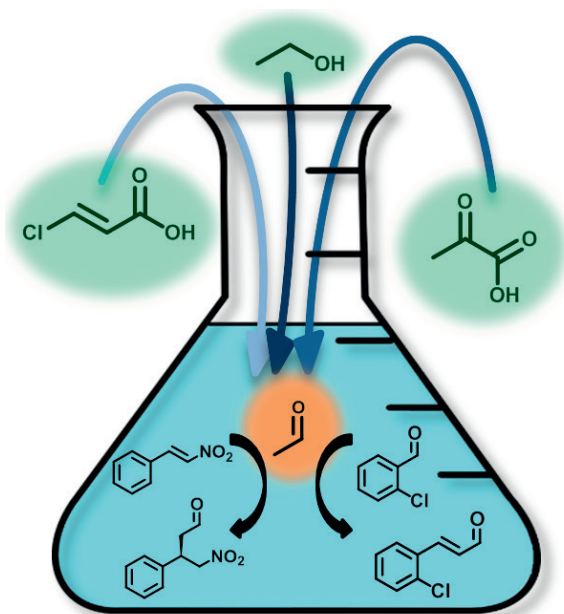
Published in *ChemBioChem* 10.1002/cbic.201900666

## Abstract

The enzyme 4-oxalocrotonate tautomerase (4-OT) can promiscuously catalyze various carboligation reactions using acetaldehyde as a nucleophile. However, the highly reactive nature of acetaldehyde requires intricate handling, which can impede its usage in practical synthesis. Therefore, we investigated three enzymatic routes to synthesize acetaldehyde *in situ* in one-pot cascade reactions with 4-OT. Two routes afforded practical acetaldehyde concentrations, using an environmental pollutant, *trans*-3-chloroacrylic acid, or a biorenewable, ethanol, as starting substrate. These routes can be combined with 4-OT catalyzed Michael-type additions and aldol condensations in one pot. This modular systems biocatalysis methodology provides a stepping stone towards the development of larger artificial metabolic networks for the practical synthesis of important chemical synthons.

## Keywords

Acetaldehyde, carboligation, biocatalysis, cascade reactions



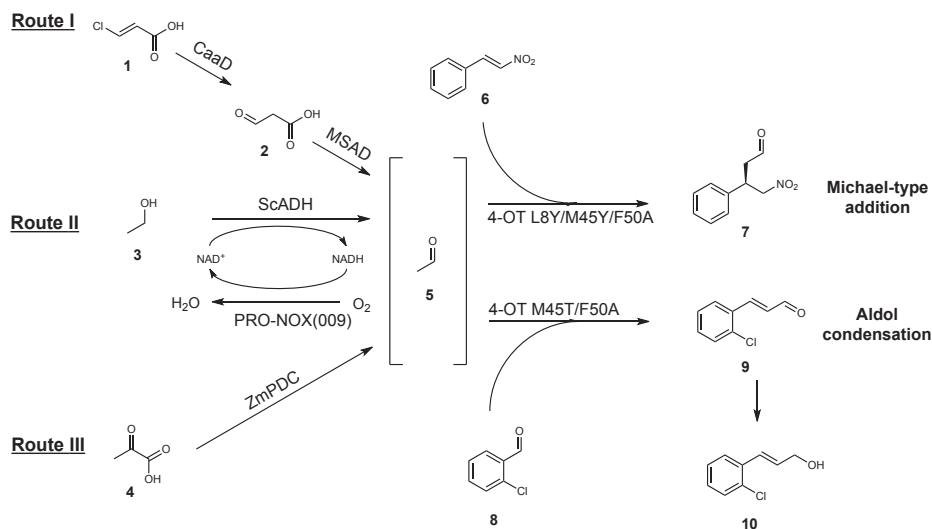
## Introduction

The enzyme 4-oxalocrotonate tautomerase (4-OT) naturally catalyzes the tautomerization of 2-hydroxy-2,4-hexadienedioate to 2-oxo-3-hexenedioate using the N-terminal proline as key catalytic base<sup>[1]</sup>. In addition, 4-OT can promiscuously catalyze several carbon-carbon bond-forming reactions, including Michael-type additions and aldol condensations, employing Pro-1 as a nucleophile<sup>[2-4]</sup>. Most notably, 4-OT can catalyze Michael-type additions and aldol condensations using the highly reactive acetaldehyde (**5**) as a nucleophile (Scheme 1).

However, the use of **5** in enzymatic reactions causes several practical challenges. Compound **5** is toxic, highly volatile and reactive, which requires intricate handling. Hence, *in situ* generation of **5** from less reactive and less toxic compounds is an attractive concept to combine with 4-OT in one-pot cascade reactions. Previous studies have primarily focused on *in situ* generation of longer aldehydes, for instance from their corresponding carboxylic acids or alcohols using carboxylic acid reductases (CAR) or alcohol dehydrogenases respectively<sup>[5-8]</sup>.

Here we investigate three enzymatic routes for the *in situ* generation of **5** in one-pot cascade reactions with 4-OT, using **1**, **3** or **4** as starting substrates (Scheme 1). Route I involves the dehalogenation of **1** into **2**, catalyzed by the enzyme chloroacrylic acid dehalogenase (CaaD), followed by the decarboxylation of **2** into **5** by the enzyme malonate semialdehyde decarboxylase (MSAD). Route II involves the oxidation of **3** into **5** catalyzed by the alcohol dehydrogenase from *Saccharomyces cerevisiae* (ScADH), in combination with a commercially available NADH oxidase, PRO-NOX(009), to shift the unfavorable reaction equilibrium and to recycle the co-factor. Route III involves the decarboxylation of **4** using the pyruvate decarboxylase from *Zymomonas mobilis* (ZmpDC).

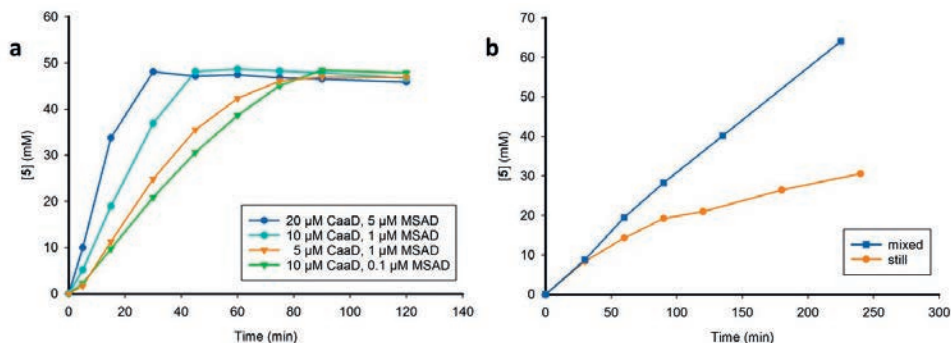
Routes I and II proved to be effective routes for *in situ* generation of **5** and could be used in combination with 4-OT to synthesize **7** and **9** in one pot. Interestingly, route II afforded a mixture of **9** and **10**. The presented work provides a stepping stone towards the construction of artificial enzymatic metabolic routes for the synthesis of valuable chemical commodities.



**Scheme 1.** Three envisioned enzymatic routes for *in situ* generation of acetaldehyde (5). The *in situ* synthesized 5 is used as substrate in a 4-OT L8Y/M45Y/F50A catalyzed Michael-type addition and a 4-OT M45T/F50A catalyzed aldol condensation reaction in one pot. Abbreviations: CaaD: chloroacrylic acid dehalogenase, MSAD: malonate semialdehyde decarboxylase, ScADH: alcohol dehydrogenase, PRO-NOX(009): NADH oxidase, ZmPDC: pyruvate decarboxylase.

## Results

There are several enzymes that naturally produce 5 as their main product, as 5 is at the crossway of several metabolic routes. We selected three complementary enzymatic routes (route I, II and III, Scheme 1) for *in situ* generation of 5, using the readily accessible starting materials 1, 3 and 4. We initially focused on the enzymatic synthesis of 5 via route I. The release of HCl and carbon dioxide in the first and second step respectively, is expected to shift the equilibrium towards the formation of 5. Indeed, initial experiments showed that a cascade reaction with CaaD and MSAD resulted in near quantitative conversion of 1 into 5 (Figure 1). Notably, no significant enzyme inhibition occurred at this concentration of 5. Optimizing enzyme concentrations afforded 50 mM 5 in 90 minutes using 5  $\mu$ M CaaD and 1  $\mu$ M MSAD (Figure 1).



**Figure 1.** Progress curves of enzymatic synthesis of **5** using route I and route II. (a) Formation of **5** via route I. Assay conditions: 50 mM **1**, 100 mM sodium phosphate (pH 7.3). (b) Formation of **5** via route II. Assay conditions: 100 mM sodium phosphate pH 7.3, 2 mg/mL PRO-NOX(009), 10 U/mL ScADH, 10% v/v **3**, 2 mM NAD<sup>+</sup>, 5 mL reaction volume, under oxygen atmosphere, performed in a 25 mL flask.

We next focused on the enzymatic oxidation of **3** (Scheme 1, route II), which can be catalyzed by both alcohol oxidases and alcohol dehydrogenases. Since alcohol oxidases are prone to over-oxidation of aldehydes into their corresponding carboxylic acids<sup>[9]</sup>, we opted for the use of an alcohol dehydrogenase instead. The commercial available Yeast alcohol dehydrogenase, (ScADH, *Saccharomyces cerevisiae*) was selected as biocatalyst because of its high selectivity towards short-chain primary alcohols<sup>[10,11]</sup>. The commercially available NADH oxidase PRO-NOX(009) was selected to recycle NAD<sup>+</sup> and to overcome the unfavorable reaction equilibrium.

Concentrations of up to 10% v/v **3** were well tolerated by both enzymes, but higher concentrations rapidly inactivated at least one of the enzymes. Vigorous stirring of the reaction mixture under an oxygen atmosphere significantly improved the rate of **5** production compared to an unstirred reaction (Figure 1). Reaction conditions were optimized to reach a concentration of approximately 50 mM **5** after 150 minutes.

Lastly, we focused on the enzymatic decarboxylation of **4** to **5** (Scheme 1, route III). The pyruvate decarboxylase from *Z. mobilis* (ZmPDC) was cloned, expressed and purified to homogeneity. However, we were unable to find conditions that afforded concentrations of **5** above 30 mM, due to either inactivation or inhibition of the enzyme by **5**. This concentration was considered too low for effective catalysis with 4-OT and hence we abandoned this route.



Having established two effective enzymatic routes for the *in situ* synthesis of **5**, we next focused our attention on the combination of these routes with 4-OT in one-pot multistep cascade reactions. We initially focused on the 4-OT catalyzed Michael-type addition of **5** to **6** yielding **7**, an important precursor for the  $\gamma$ -aminobutyric acid analogue phenibut<sup>[4,12]</sup>. As a catalyst, the previously engineered 4-OT L8Y/M45Y/F50A variant was selected<sup>[13]</sup>. Combination of 4-OT L8Y/M45Y/F50A and **6** with *in situ* synthesized **5** (via both route I and II) afforded *R*-**7** in high enantiopurity (e.r. up to 98:2) and good to excellent isolated yields (up to 96% compared to **6**) (Table 1).

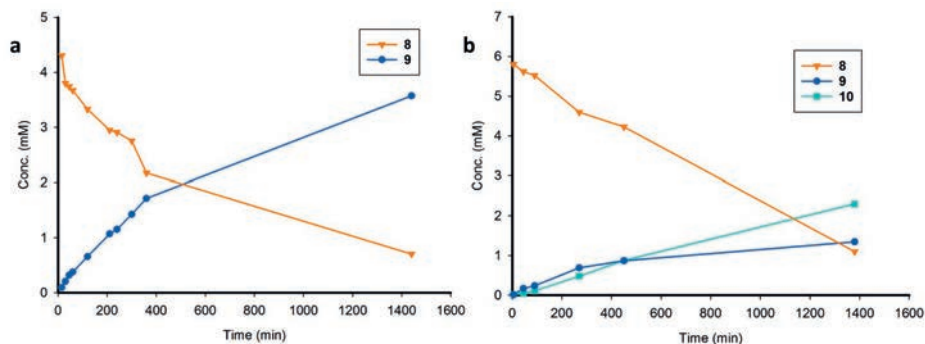
**Table 1.** Stepwise enzymatic cascade synthesis of **7**.

Reaction time				
	Step I (min) <sup>c</sup>	Step II (min) <sup>d</sup>	e.r. <sup>e</sup>	Isolated yield (%) <sup>f</sup>
Route I	150	75	98:2	62
Route II	90	85	98:2	96

<sup>a</sup>Dehalogenation of **1** into **2** catalyzed by CaaD, followed by decarboxylation of **2** into **5** catalyzed by MSAD, followed by the Michael-type addition of **5** to **6** yielding **7** catalyzed by 4-OT L8Y/M45Y/F50A. The reaction mixture consisted of 50 mM **1**, 4 mM **6**, 1  $\mu$ M MSAD, 5  $\mu$ M CaaD, 56  $\mu$ M 4-OT L8Y/M45Y/F50A, 100 mM sodium phosphate pH 7.3, 10% v/v ethanol. <sup>b</sup>Oxidation of **3** into **5** catalyzed by ScADH, followed by the Michael-type addition of **5** to **6** yielding **7** catalyzed by 4-OT L8Y/M45Y/F50A. PRO-NOX(009) was used for co-factor recycling. The reaction mixture consisted of 10% v/v **3**, 4 mM **6**, 10 U/mL ScADH, 2 mg/mL PRO-NOX(009), 2 mM NAD<sup>+</sup>, 56  $\mu$ M 4-OT L8Y/M45Y/F50A, 100 mM sodium phosphate pH 7.3. <sup>c</sup>Monitored by HPLC. <sup>d</sup>Monitored by UV spectroscopy. <sup>e</sup>Determined by GC with chiral stationary phase or derivatized into a cyclic acetal and determined by HPLC with a chiral stationary phase. The absolute configuration was determined by comparison to literature<sup>[14,15]</sup>. <sup>f</sup>Isolated yield compared to **6**.

Next, we focused on combining route I and II with the 4-OT catalyzed aldol condensation of **5** with **8**, forming **9**. The previously engineered 4-OT M45T/F50A was used as a catalyst for this reaction<sup>[16]</sup>. Combination of route I with **8** and 4-OT M45T/F50A afforded **9** in good isolated yield (56% compared to **8**) (Figure 2 and Table 2). Interestingly, the combination of route II with **8** and 4-OT M45T/F50A afforded a mixture of **9** (26% isolated yield compared to **8**) and cinnamyl alcohol **10** (28% isolated yield compared to **8**) (Figure 2 and Table 2).

These results demonstrate some exciting opportunities available for the combination of natural metabolic routes with unnatural carboligation reactions in one pot.



**Figure 2.** Progress curves of the enzymatic cascade synthesis of **9** and **10**. (a) Progress curve of the 4-OT M45T/F50A catalyzed aldol condensation of **5** with **8** using *in situ* synthesized **5** via route I. Assay conditions: 50 mM **1**, 4 mM **8**, 1  $\mu$ M MSAD, 10  $\mu$ M CaaD, 75  $\mu$ M 4-OT M45T/F50A, 150 mM sodium phosphate pH 7.3, 5% v/v ethanol, 17.8 mL reaction volume. (b) Progress curve of the 4-OT M45T/F50A catalyzed aldol condensation of **5** with **8** using *in situ* synthesized **5** via route II. Assay conditions: 10% v/v **3**, 5.7 mM **8**, 2 mM NAD<sup>+</sup>, 133  $\mu$ M 4-OT M45T/F50A, 10 U/mL ScADH, 2 mg/mL PRO-NOX(009), 100 mM sodium phosphate pH 7.3, 30 mL reaction volume.

**Table 2.** Enzymatic cascade synthesis of **9**.

	Reaction time (h)	Conversion (%) <sup>c</sup>	Isolated yield (%) <sup>d</sup>
Route I	23.5	71	56
Route II	24	85	26 <sup>e</sup>

<sup>a</sup>Dehalogenation of **1** into **2** catalyzed by CaaD, followed by decarboxylation of **2** into **5** catalyzed by MSAD, followed by the aldol condensation of **5** with **8** yielding **9** catalyzed by 4-OT M45T/F50A. The reaction mixture consisted of 50 mM **1**, 4 mM **8**, 1  $\mu$ M MSAD, 10  $\mu$ M CaaD, 75  $\mu$ M 4-OT M45T/F50A, 150 mM sodium phosphate pH 7.3, 5% v/v ethanol. <sup>b</sup>Oxidation of **3** into **5** catalyzed by ScADH, followed by the aldol condensation of **5** with **8** yielding **9** catalyzed by 4-OT M45T/F50A. The reaction mixture consisted of 10% v/v **3**, 5.7 mM **8**, 2 mM NAD<sup>+</sup>, 133  $\mu$ M 4-OT M45T/F50A, 10 U/mL ScADH, 2 mg/mL PRO-NOX(009), 100 mM sodium phosphate pH 7.3. <sup>c</sup>Determined by HPLC. <sup>d</sup>Compared to **8**. <sup>e</sup>The low yield is caused by partial reduction of **9** into **10** (Figure 2b). Cinnamyl alcohol **10** could be obtained with 28% isolated yield compared to **8**.

## Discussion

4-OT is unique in that it can use **5** as substrate for C-C bond-forming Michael-type additions to nitroolefin acceptors and for aldol condensations with benzaldehydes<sup>[3,4]</sup>. Previously, we have demonstrated that 4-OT can be combined with other biocatalysts and chemocatalysts to convert **6** and **8** into GABA analogues using two complementary one-pot cascade reactions<sup>[13,15]</sup>. However, both routes rely on **5** as a substrate. The inherent high reactivity and toxicity of **5** makes it a less desirable substrate and special care has to be taken to prevent polymerization or oxidation. A solution to this problem is the enzymatic *in situ* generation of **5** using less toxic and reactive starting substrates in multistep cascade reactions. To this end, we have investigated three enzymatic routes for the *in situ* generation of **5**, using different starting materials.

The first route we investigated involves the conversion of soil pollutant **1** into **5**. Recycling of **1**, a degradation product from the xenobiotic soil fumigant *trans*-1,3-dichloropropene<sup>[17]</sup>, into GABA precursors provides an interesting example of the first principle of circular chemistry and as such might be desirable for future sustainable synthesis<sup>[18]</sup>. To this end, we selected the enzymes *trans*-3-chloroacrylic acid dehalogenase and malonate semi-aldehyde decarboxylase, which are part of the metabolic pathway for the degradation of *trans*-1,3-dichloropropene in *Pseudomonas cichorii* 170<sup>[19,20]</sup>. The second route comprises of the oxidation of **3** into **5**. As an inexpensive biorenewable, **3** is an attractive precursor of **5** that can concurrently act as a co-solvent to solubilize **6** and **8**. To catalyze the oxidation of **3**, we selected yeast alcohol dehydrogenase, using an NADH oxidase for co-factor recycling and as a driving force to overcome the unfavorable reaction equilibrium between **3** and **5**. Thirdly, we investigated the readily available substrate **4** as precursor for **5**, using pyruvate decarboxylase from *Z. mobilis* as biocatalyst.

Gratifyingly, concentrations of 50 mM **5** or higher could be generated *in situ* using both route I and II. Route III only afforded lower concentrations of **5** and unfortunately, we were unable to find conditions to improve this. We speculate that ZmPDC was either inactivated or strongly inhibited by **5**<sup>[21]</sup>, and the effective use of route III for *in situ* generation of **5** awaits the discovery of 4-OT mutants with a lower  $K_M$  for **5**.

We demonstrated the modularity of our approach by showing that the *in situ* generated **5** can be used as a substrate for both 4-OT catalyzed aldol condensations and Michael-type additions. The Michael-type addition of **5** to **6** was catalyzed by the previously engineered 4-OT L8Y/M45Y/F50A and afforded *R*-**7** in good to excellent isolated yield (up to 96%

compared to **6**) and with an excellent e.r. value of 98:2. The aldol condensation of **5** with **8** was catalyzed by the previously engineered 4-OT M45T/F50A. Route I afforded **9** in 55% isolated yield, but route II afforded a mixture of **9** and **10**. The reduction of **9** to **10** is slowly catalyzed by ScADH and by an alcohol dehydrogenase present in the CFE of PRO-NOX(009).

In summary, we have investigated three enzymatic synthesis routes, using the readily accessible starting materials **1**, **3** or **4**, for *in situ* generation of **5** to be used as substrate by 4-OT in one-pot cascade syntheses. Route I and route II afforded **5** in effective concentrations for synthesis with 4-OT, but route III afforded **5** only in low concentrations. Route I and route II are fully compatible with 4-OT, which was demonstrated by the one-pot cascade synthesis of **7** and **9**. These modular enzymatic cascades provide a stepping stone towards the development of larger “artificial metabolisms” that could facilitate greener and more sustainable synthesis<sup>[22]</sup>.

### **Acknowledgement**

We acknowledge financial support from the Netherlands Organization of Scientific Research (VICI grant 724.016.002 and ECHO grant 713.015.003), and the European Union’s Horizon 2020 Research and Innovation Programme under grant agreement No 635595 (CarbaZymes).

### **Conflict of Interest**

The authors declare no conflict of interest.

## References

1. C. P. Whitman, *Arch. Biochem. Biophys.* **2002**, *402*, 1–13.
2. H. Poddar, M. Rahimi, E. M. Geertsema, A. M. W. H. Thunnissen, G. J. Poelarends, *ChemBioChem* **2015**, *16*(5), 738–741.
3. E. Zandvoort, B.J. Baas, W. J. Quax, G. J. Poelarends, *ChemBioChem* **2011**, *12*, 602–609.
4. E. Zandvoort, E. M. Geertsema, B. J. Baas, W. J. Quax, G. J. Poelarends, *Angew. Chemie* **2011**, *124*, 1266–1269.
5. S. P. France, S. Hussain, A. M. Hill, L. J. Hepworth, R. M. Howard, K. R. Mulholland, S. L. Flitsch, N. J. Turner, *ACS Catal.* **2016**, *6*, 3753–3759.
6. J. I. Ramsden, R. S. Heath, S. R. Derrington, S. L. Montgomery, J. Mangas-sanchez, K. R. Mulholland, N. J. Turner, *J. Am. Chem. Soc.* **2019**, *141*, 1201–1206.
7. M. Moura, D. Pertusi, S. Lenzini, N. Bhan, L. J. Broadbelt, K. E. J. Tyo, *Biotechnol. Bioeng.* **2016**, *113*, 944–952.
8. M. Fuchs, K. Tauber, J. Sattler, H. Lechner, J. Pfeffer, W. Kroutil, K. Faber, *RSC Adv.* **2012**, *2*, 6262–6265.
9. M. Pickl, M. Fuchs, S. M. Glueck, K. Faber, *Appl. Microbiol. Biotechnol.* **2015**, *99*, 6617–6642.
10. R. Pietruszko, K. Crawford, D. Lester, *Arch. Biochem. Biophys.* **1973**, *159*, 50–60.
11. F. M. Dickinson, G. P. Monger, *Biochem. J.* **1973**, *131*, 261–270.
12. O. V. Maltsev, A. S. Kucherenko, I. P. Beletskaya, V. A. Tartakovskiy, S. G. Zlotin, *European J. Org. Chem.* **2010**, *15*, 2927–2933.
13. L. Biewenga, T. Saravanan, A. Kunzendorf, J.Y. van der Meer, T. Pijning, P. G. Tepper, R. van Merkerk, S. J. Charnock, A. M. W. H. Thunnissen, G. J. Poelarends, *ACS Catal.* **2019**, *9*, 1503–1513.
14. E. M. Geertsema, Y. Miao, P. G. Tepper, P. de Haan, E. Zandvoort, G. J. Poelarends, *Chem. Eur. J.* **2013**, *19*, 14407–14410.
15. C. Guo, M. Saifuddin, T. Saravanan, M. Sharifi, G. J. Poelarends, *ACS Catal.* **2019**, *9*, 4369–4373.
16. M. Rahimi, J. Y. van der Meer, E. M. Geertsema, H. Poddar, B. J. Baas, G. J. Poelarends, *ChemBioChem* **2016**, *17*, 1225–1228.
17. T. R. Roberts, G. Stoydin, *Pestic. Sci.* **1976**, *7*, 325–335.
18. T. Keijer, V. Bakker, J. C. Slootweg, *Nat. Chem.* **2019**, *11*, 190–195.
19. G. J. Poelarends, M. Wilkens, M. J. Larkin, J. D. van Elsas, D. B. Janssen, *Appl. Environ. Microbiol.* **1998**, *64*, 2931 – 2936.
20. G. J. Poelarends, W. H. Johnson, A. G. Murzin, C. P. Whitman, *J. Biol. Chem.* **2003**, *278*, 48674–48683.

21. H. Bruhn, M. Pohl, J. Grotzinger, M. R. Kula, *Eur. J. Biochem.* **1995**, 234, 650–655.
22. W. D. Fessner, *N. Biotechnol.* **2015**, 32, 658–664.

## Supporting Information

### General Methods

#### Recombinant enzyme expression and purification

##### 4-Oxalocrotonate Tautomerase

4-oxalocrotonate tautomerase (4-OT) variants were expressed in *E. coli* BL21(DE3) harboring a pJExpress414 vector (4-OT L8Y/M45Y/F50A) or pET20b vector (4-OT M45T/F50A) with the respective 4-OT gene. Cells were grown overnight at 37 °C in 1 L LB medium substituted with 0.5% glycerol, 100 µg/ml ampicillin and 100 µM IPTG. After the expression of 4-OT the cells were collected and the enzymes were purified according to previously reported methods<sup>1,2</sup>. The enzyme concentration was determined using the Waddell method, the enzymes were flash frozen in liquid nitrogen and stored at -20 °C until further use<sup>3</sup>. For each purified 4-OT variant the removal of the N-terminal methionine was confirmed by analysis of the correct protein mass using electron spray ionization mass spectrometry.

##### Trans-3-chloroarylic acid dehalogenase

The enzyme *trans*-3-chloroarylic acid dehalogenase (CaaD) from *P. pavonaceae* 170 (previously known as *P. cichorii* 170) was expressed in *E. coli* BL21(DE3) harboring a pET5a vector with the CaaD gene. A single colony was used to inoculate 50 mL LB medium supplemented with 100 µg/ml ampicillin and grown overnight at 37 °C. 10 mL of this overnight culture was used to inoculate 1 L LB medium supplemented with 100 µg/ml ampicillin and incubated at 37 °C. After the culture reached an OD of 0.4 – 0.6 the expression of CaaD was induced by the addition of 0.5 mM IPTG. The cells were grown for 6 h at 37 °C and collected by centrifugation (15 min, 4 °C, 10000 rpm). CaaD was purified by a modification of a published procedure<sup>4</sup>. The cell pellet was resuspended in 10 mM sodium phosphate buffer (pH 8) and the cells were disrupted by ultrasonication. Insoluble cell fragments were removed by centrifugation. The supernatant was collected and solid (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub> crystals were slowly added to a final concentration of 1.6 M (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub>. The solution was kept overnight at 4 °C and then centrifuged for 30 min. The supernatant was filtered and loaded onto a 5 ml HiTrap Phenyl HP column equilibrated with 10 mM sodium phosphate buffer (pH 8) containing 1.6 M (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub>. The column was washed with 3 column volumes (CV) of 10 mM sodium phosphate buffer (pH 8) containing 1.6 M (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub>, after which a linear gradient of 10 mM sodium phosphate buffer (pH 8) 1.6 M to 0 M (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub> over 20 CV was applied. Individual fractions were analyzed by SDS-PAGE and the fractions containing high concentrations of CaaD were pooled. The buffer was exchanged against 20 mM Tris-HCl buffer (pH 8.5) on a

HiPrep 26/10 Desalting column. The protein solution was loaded on a 5 ml HiTrap Q FF column equilibrated with 20 mM Tris-HCl buffer (pH 8.5), washed with 3 CV of the equilibration buffer and eluted with a linear salt gradient (0 – 1 M NaCl) over 20 CV with 20 mM Tris-HCl buffer (pH 8.5). The fractions with the highest CaaD concentration were pooled together, dialyzed in 4 L 10 mM sodium phosphate buffer (pH 7.3) and concentrated using a vivaspin centrifugal concentrator (Sartorius, 5000 MWCO). The enzyme concentration was determined using the Waddell method and the enzyme was flash frozen in liquid nitrogen and stored at – 20 °C until further use<sup>3</sup>. The correct molecular mass of CaaD was confirmed by electron spray ionization mass spectrometry.

#### Malonate semialdehyde decarboxylase

The enzyme malonate semialdehyde decarboxylase (MSAD) was expressed and purified by a modification of a published procedure<sup>5</sup>. A single colony of *E. coli* BL21(DE3) harboring a pET3a vector with the MSAD gene was used to inoculate 50 mL LB medium supplemented with 100 µg/ml ampicillin. After overnight growth at 37 °C this culture was used to inoculate 1 L LB medium supplemented with 100 µg/ml ampicillin. The culture was grown at 37 °C until an OD 0.4 -0.6 was reached, after which protein expression was induced by addition of 0.5 mM IPTG. After 5 h incubation at 37 °C, the cells were collected by centrifugation (10 min, 10000 rpm). The obtained cell pellet was resuspended in 10 mM Tris-SO<sub>4</sub> buffer (pH 8) and disrupted by ultrasonication. Insoluble cell fragments were removed by centrifugation and the supernatant was applied to a 5 ml HiTrap Q FF column equilibrated with 10 mM Tris-SO<sub>4</sub> buffer (pH 8). The column was washed with 5 CV equilibration buffer and the proteins were eluted using a linear salt gradient of 0-0.5 M Na<sub>2</sub>SO<sub>4</sub> over 20 CV. Individual fractions were analyzed by SDS-PAGE and the fractions with the highest MSAD concentration were pooled together. Solid (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub> crystals were slowly added to this solution until a final concentration of 1 M (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub> and the resulting solution was stirred for 1 h at 4 °C. After centrifugation the supernatant was loaded on a 5 ml HiTrap Phenyl HP column equilibrated with 10 mM Tris-SO<sub>4</sub> buffer (pH 8) containing 1 M (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub>. The column was washed with 3 CV of the equilibration buffer and retained proteins were eluted with a linear gradient (1 – 0 M (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub>) over 20 CV. Individual fractions were analyzed by SDS-PAGE and the fractions with the highest MSAD concentration were pooled together, 2x dialyzed in 4 L 10 mM Na<sub>2</sub>HPO<sub>4</sub> buffer (pH 7.3) and concentrated using a vivaspin centrifugal concentrator (Sartorius, 5000 MWCO). The enzyme concentration was determined using the Waddell method and the enzyme was flash frozen in liquid nitrogen and stored at – 20 °C until further use<sup>3</sup>. The correct molecular mass of MSAD was confirmed by electron spray ionization mass spectrometry.



### Pyruvate decarboxylase

The enzyme pyruvate decarboxylase from *Z. mobilis* (ZmPDC) was expressed in *E. coli* BL21(DE3) harboring a pET21a vector containing an *E. coli* codon-optimized ZmPDC gene with a C-terminal His-tag. 1 L LB medium supplemented with 100 µg/ml ampicillin was inoculated with cells from an overnight culture. The cells were grown at 37 °C until the culture reached an OD 0.4 – 0.6, after which protein expression was induced by the addition of 100 µM IPTG. After growth overnight at 20 °C, the cells were collected by centrifugation and the cell pellet was resuspended in 50 mM Mes/NaOH, 0.1 mM ThDP, 5 mM MgSO<sub>4</sub> (pH 6.5). The cells were disrupted by ultrasonication and insoluble cell fragments were removed by centrifugation (45 min, 4 °C, 18000 rpm). The supernatant was incubated for 1 h at 4 °C with Ni sepharose resin, which had previously been equilibrated with lysis buffer. The resin loaded with proteins was washed with 6 mL lysis buffer, followed by 6 mL lysis buffer (pH 7) containing 20 mM imidazole. The retained proteins were eluted with 50 mM Mes/NaOH, 0.1 mM ThDP, 1 mM MgSO<sub>4</sub>, 250 mM imidazole (pH 7.6). The fractions with the highest ZmPDC concentration were pooled together and the buffer was exchanged to 10 mM Mes/NaOH, 200 mM NaCl, 5 mM MgSO<sub>4</sub>, 0.1 mM ThDP (pH 6.5) using a PD-10 desalting column. The enzyme concentration was determined using the Waddell method and the enzyme was flash frozen in liquid nitrogen and stored at – 80 °C until further use<sup>3</sup>.

### Alcohol dehydrogenase

ScADH was commercially available from Sigma Aldrich, provided as lyophilized enzyme (product number A7011, alcohol dehydrogenase from *Saccharomyces cerevisiae*). Prior to each experiment, a sample of the lyophilized enzyme was dissolved in 100 mM Na<sub>2</sub>HPO<sub>4</sub> buffer (pH 7.3) and the unit concentration was determined according to the provided protocol. Unit definition: one unit of alcohol dehydrogenase will convert 1.0 µmol of ethanol to acetaldehyde per minute in 50 mM sodium phosphate pH 8.8 at 25 °C.

### NADH oxidase

PRO-NOX(009) was provided as crude cell-free extract by Prozomix Ltd and used without further purification.

## ***In situ* synthesis of 5**

### Route I

Optimization of *in situ* synthesis of **5** via route I was performed on 1 mL scale. To a 1.5 mL Eppendorf tube containing 50 mM **1**, 10% ethanol, 1 µM MSAD in 100 mM Na<sub>2</sub>HPO<sub>4</sub> buffer (pH 7.3) was added Caad (5 µM final concentration). At timely intervals 50 µL was collected from the reaction mixture and the aldehyde product was derivatized with O-benzylhydroxylamine.

### Route II

Optimization of *in situ* synthesis of **5** via route II was performed in a 25 mL round-bottom flask. 3 mL 100 mM Na<sub>2</sub>HPO<sub>4</sub> buffer (pH 7.3) was added to the flask. The flask was sealed with a rubber stopper and purged with pure oxygen for 30 minutes to saturate the buffer and airspace. To this was added 0.5 mL **3**, PRO-NOX(009) (2 mg/mL final concentration) and ScADH (10 U/mL final concentration) and 100 mM Na<sub>2</sub>HPO<sub>4</sub> buffer (pH 7.3) into a final volume of 5 mL. The reaction was stirred at 250 rpm. At timely intervals, 50 µL was collected from the reaction mixture and the aldehyde product was derivatized with O-benzylhydroxylamine.

### Route III

The activity of ZmPDC was tested based on reported reaction conditions<sup>6</sup>. In a 1.5 mL Eppendorf tube 50 mM sodium pyruvate was added to 50 mM MES/NaOH buffer (pH 6.5) supplemented with 0.1 mM ThDP, 5 mM MgSO<sub>4</sub> and 5% DMSO and the reaction was initiated by the addition of 0.1 µM ZmPDC. The final volume of the reaction was 1 mL. At timely intervals 50 µL was collected from the reaction mixture and the aldehyde product was derivatized with O-benzylhydroxylamine.

### **Aldehyde derivatization and quantification by HPLC**

The aldehyde compounds in collected reaction samples were derivatized with O-benzylhydroxylamine and analyzed by reverse-phase HPLC according to a reported procedure<sup>1</sup>. The concentration in the reaction of **5**, **8**, **9** and **10** (without derivatization) was determined by comparing peak areas to known standards of **5**, **8**, **9** and **10**.

### **Semi-Preparative scale cascade reactions**

#### Synthesis of **7** via route I

To a 25 mL round-bottom flask was added: 90.5 mg **1** (sodium salt), 212 µg MSAD (1 µM final concentration) and 1.19 mg CaaD (5 µM final concentration) in 15 mL 100 mM sodium phosphate buffer (pH 7.3). At timely intervals, 50 µL of the reaction mixture was collected and processed for HPLC analysis to monitor reaction progress. During the reaction the pH dropped from pH 7.3 to 6.8. After 90 min, 10.14 mg **6** (in 1.7 mL ethanol, 4 mM final concentration) and 6.5 mg 4-OT L8Y/M45Y/F50A (56 µM final concentration) were added (final volume 17 mL). The reaction progress was monitored by UV-spectroscopy following the depletion in absorbance at 320 nm corresponding to the concentration of **6**. After the reaction was finished, the reaction mixture was extracted with 3 x 30 mL ethyl acetate. The organic layers were washed with brine, dried over anhydrous Na<sub>2</sub>SO<sub>4</sub> and concentrated *in vacuo*. Silica gel column chromatography (petroleum ether/ethyl acetate 9:1) gave the desired product **7** (8.3 mg, 62% yield). The

<sup>1</sup>H NMR spectroscopic data were in agreement with previously published data<sup>7</sup>. The aldehyde functionality was derivatized into the corresponding ethylene glycol acetal and the enantiomer ratio (e.r.) was determined by reverse phase HPLC (AD-RH column 150 mm x 4.6 mm, Diacel, 30:70 H<sub>2</sub>O/ACN, 0.5 mL/min, retention time *R*-7: 8.1 min, *S*-7: 10.8 min)<sup>8</sup>.

#### Synthesis of 9 via route I

To a 25 mL round-bottom flask was added: 50 mM **1** (sodium salt), 1 μM MSAD, 10 μM CaaD, 4 mM **8** and 75 μM 4-OT M45T/F50A in a final volume of 17.8 mL 150 mM Na<sub>2</sub>HPO<sub>4</sub> buffer (pH 7.3), 5% v/v ethanol. At timely intervals, 50 μL of the reaction mixture was collected and processed for HPLC analysis to monitor reaction progress. The pH of the buffer did not significantly change during the reaction. After the reaction was finished, the reaction mixture was extracted with 3 x 10 mL ethyl acetate. The organic layers were washed with brine, dried over anhydrous Na<sub>2</sub>SO<sub>4</sub> and concentrated *in vacuo*. Silica gel column chromatography (petroleum ether/ethyl acetate 9:1) gave the desired product **9** (6.7 mg, 56% yield). The <sup>1</sup>H NMR spectroscopic data were in agreement with previously published data<sup>9</sup>.

#### Synthesis of 7 via route II

10 mL of 100 mM sodium phosphate pH 7.3 was added to a 100 mL round-bottom flask sealed with a rubber stopper. The buffer and airspace were saturated with pure oxygen by purging oxygen into the solution for 1 h. To this was added: 1.5 mL **3**, 0.5 mL of 60 mM NAD<sup>+</sup>, 0.5 mL 15 mg/mL PRO-NOX(009). The reaction was initiated by the addition of 1 mL 150 U/mL ScADH. The final conditions were: 10 U/mL ScADH, 2 mg/mL PRO-NOX(009), 2 mM NAD<sup>+</sup>, 10% v/v **3** in a final volume of 15 mL. The reaction was stirred and incubated at room temperature. At timely intervals, 50 μL of the reaction mixture was collected and processed for HPLC analysis to monitor reaction progress. After 150 min, 10.1 mg of **6** dissolved in 1 mL of **3** (4 mM final concentration) and 1 mL of 100 mM sodium phosphate containing 6.5 mg/mL 4-OT L8Y/M45Y/F50A was added to the reaction mixture. At timely intervals, 300 μL of the reaction mixture was added to a 1 mm quartz cuvette and a UV/VIS spectrum from 200 to 500 nm was recorded. 75 min after the addition of 4-OT, the reaction was finished and the reaction mixture was extracted 3x with 10 mL ethyl acetate. The organic layers were combined, washed with brine and further dried over anhydrous Na<sub>2</sub>SO<sub>4</sub>. The organic layer was removed *in vacuo* yielding **7** without any further purification (12.6 mg, 96% isolated yield). The <sup>1</sup>H NMR spectroscopic data were in agreement with previously published data<sup>7</sup>. A sample of **7** was dissolved in ethyl acetate and injected on a gas chromatograph with a chiral stationary phase using a reported method to determine the enantiopurity<sup>9</sup>.

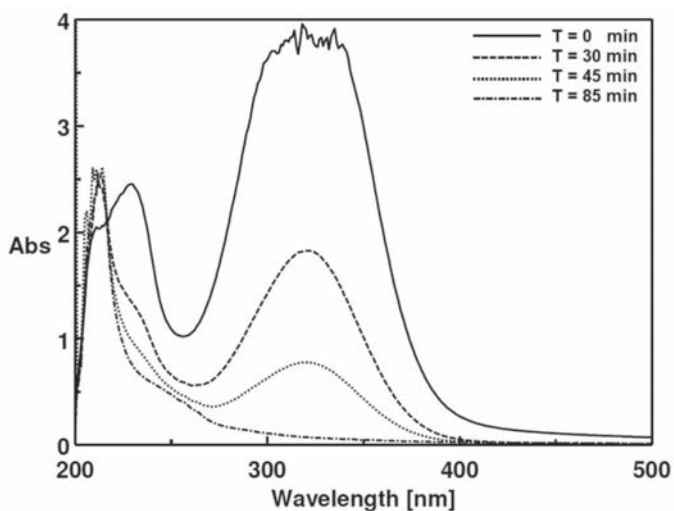
### Synthesis of **9** via route II

20 mL of 100 mM sodium phosphate pH 7.3 was added to a 100 mL round-bottom flask sealed with a rubber stopper. The buffer and airspace were saturated with pure oxygen by purging oxygen into the solution for 1 h. To this was added: 3 mL of **3** containing 24 mg of **8**, 1 mL 60 mM NAD<sup>+</sup>, 3 mL of 100 mM sodium phosphate pH 7.3 containing 20 mg/mL PRO-NOX(009) and 2 mL 100 mM sodium phosphate pH 7.3 containing 13.5 mg/mL 4-OT M45T/F50A. The reaction was started by the addition of 1 mL 100 mM Sodium phosphate pH 7.3 containing 300 U/mL ScADH. The final cascade conditions were 10 U/mL ScADH, 2 mg/mL PRO-NOX(009), 0.89 mg/mL 4-OT M45T/F50A, 2 mM NAD<sup>+</sup>, 5.7 mM **8**, 10% v/v **3** in a final volume of 30 mL. The reaction was stirred and incubated at room temperature. At timely intervals, 50  $\mu$ L of the reaction mixture was collected and processed for HPLC analysis to monitor reaction progress. After 24 h, the reaction was extracted 3x with 20 mL of ethyl acetate. The organic layers were combined, washed with brine, and dried over anhydrous Na<sub>2</sub>SO<sub>4</sub>. The dried organic layer was concentrated in vacuo and the residue was purified by silica gel column chromatography (hexane/ethyl acetate 9:1 for **9** and 4:1 for **10**) to obtain 7.5 mg of **9** and 8.0 mg of **10** (26% and 28% isolated yield respectively). The <sup>1</sup>H NMR spectroscopic data were in agreement with previously published data<sup>9,10</sup>.

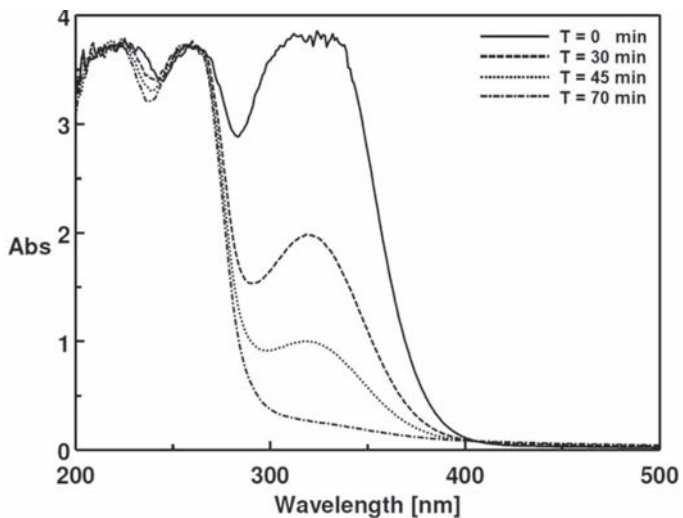
### **Synthesis of racemic reference compounds**

Racemic **7** was synthesized according to a reported procedure<sup>11</sup>.

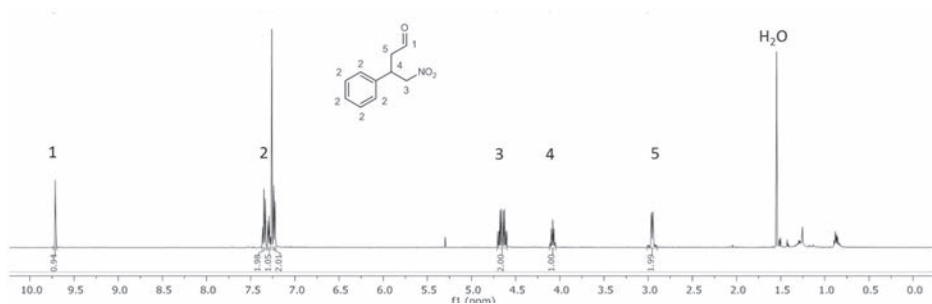
## Supporting figures



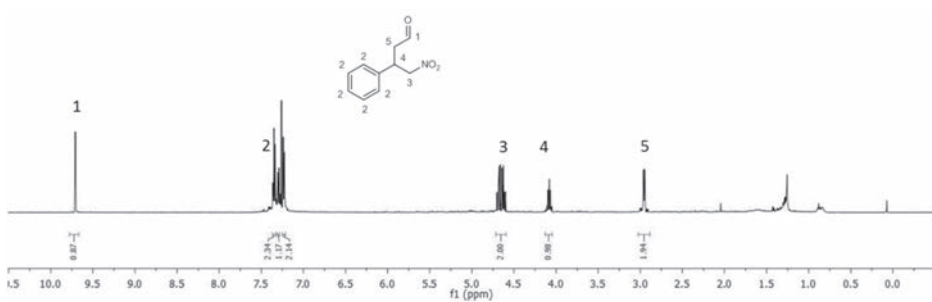
**Supplementary Figure 1.** UV spectra showing the conversion of **6** catalyzed by 4-OT L8Y/M45Y/F50A using *in situ* generated **5** via route I.



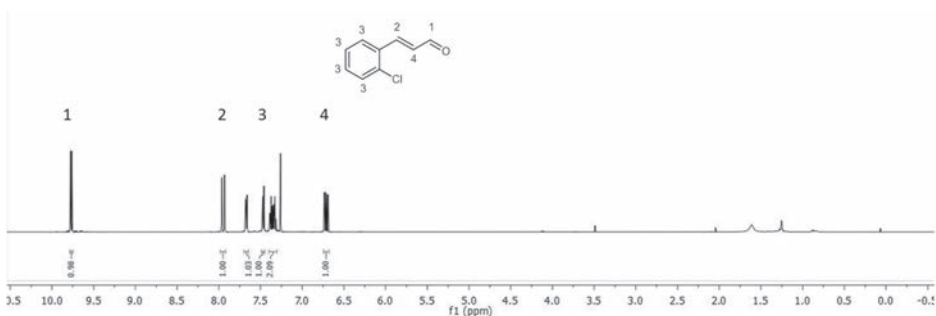
**Supplementary Figure 2.** UV spectra showing the conversion of **6** catalyzed by 4-OT L8Y/M45Y/F50A using *in situ* generated **5** via route II.



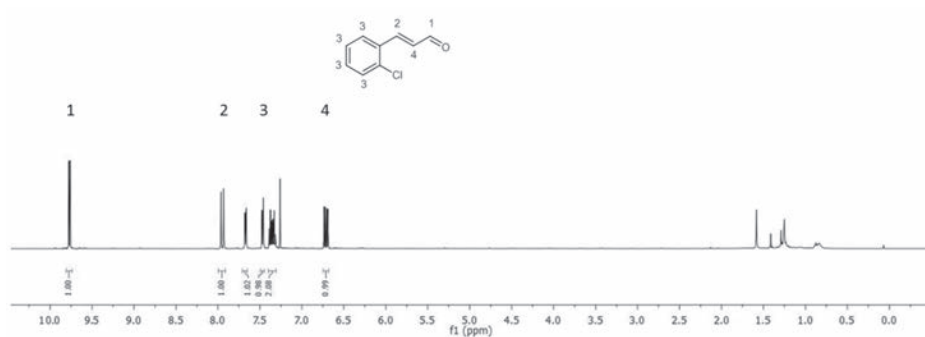
**Supplementary Figure 3.** <sup>1</sup>H NMR spectrum of enzymatically obtained 4-nitro-3-phenylbutanal (7) using *in situ* generated 5 via route I (400 MHz, CDCl<sub>3</sub>).



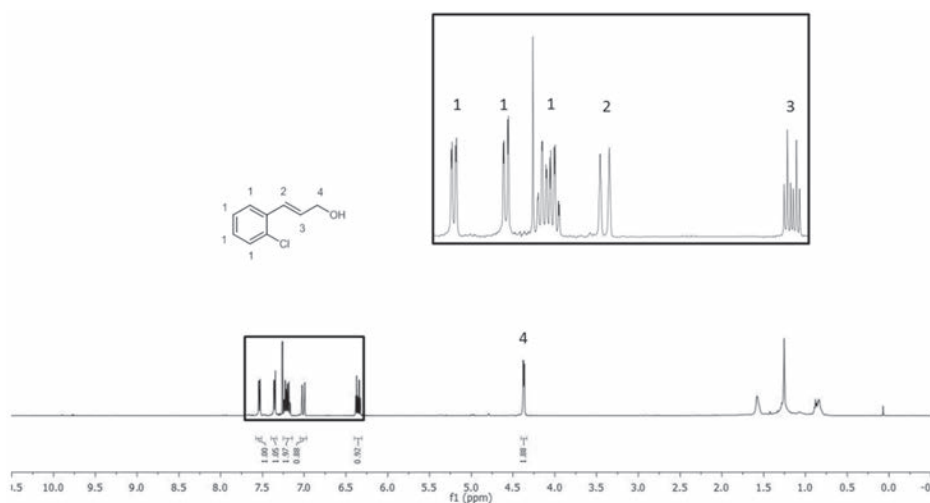
**Supplementary Figure 4.** <sup>1</sup>H NMR spectrum of enzymatically obtained 4-nitro-3-phenylbutanal (7) using *in situ* generated 5 via route II (400 MHz, CDCl<sub>3</sub>).



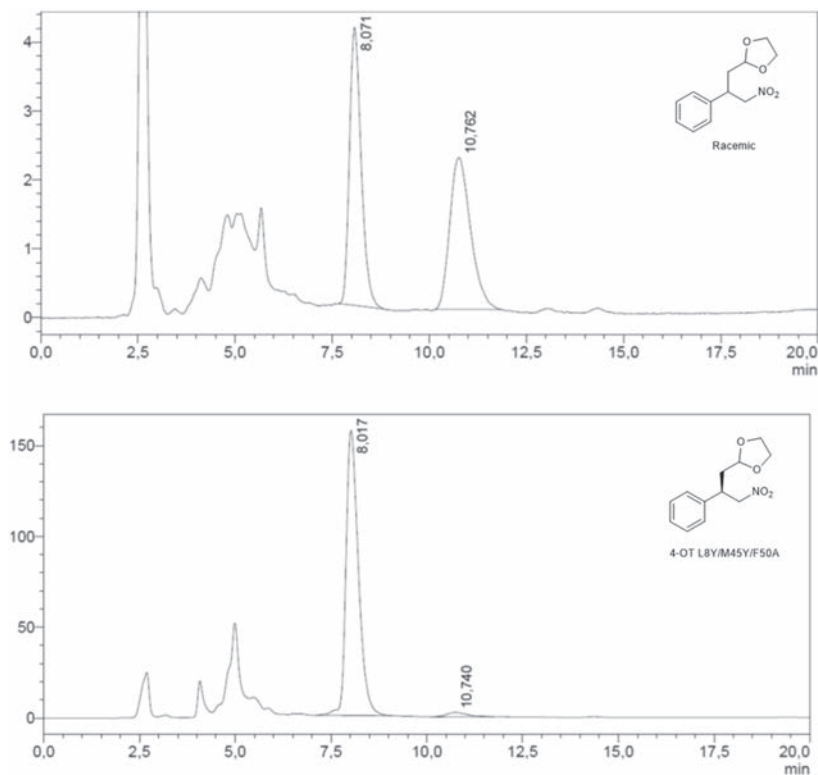
**Supplementary Figure 5.** <sup>1</sup>H NMR spectrum of enzymatically obtained (*E*)-3-(2-chlorophenyl)acrylaldehyde (9) using *in situ* generated 5 via route I (400 MHz, CDCl<sub>3</sub>).



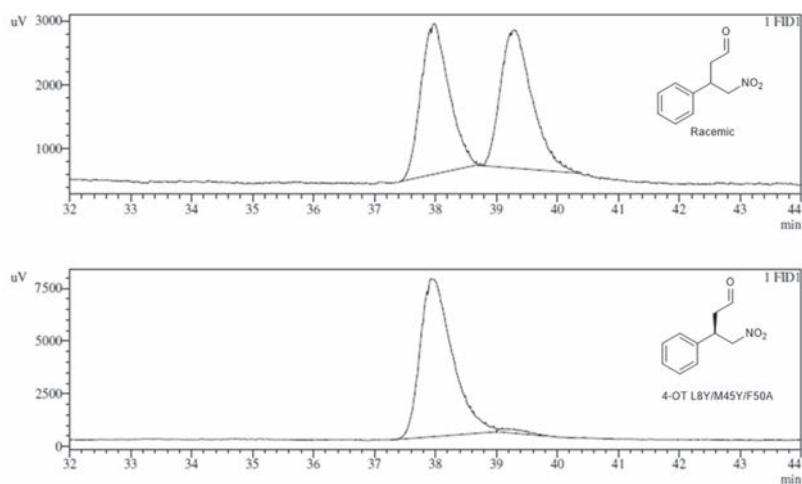
**Supplementary Figure 6.** <sup>1</sup>H NMR spectrum of enzymatically obtained (*E*)-3-(2-chlorophenyl)acrylaldehyde (**9**) using *in situ* generated **5** via route II (400 MHz, CDCl<sub>3</sub>).



**Supplementary Figure 7.** <sup>1</sup>H NMR spectrum of enzymatically obtained (*E*)-3-(2-chlorophenyl)prop-2-en-1-ol (**10**) using *in situ* generated **5** via route II (400 MHz, CDCl<sub>3</sub>).



**Supplementary Figure 8.** HPLC chromatogram of derivatized racemic 4-nitro-3-phenylbutanal (top) and derivatized enzymatically obtained 4-nitro-3-phenylbutanal using *in situ* generated **5** via route I (bottom).



**Supplementary Figure 9.** GC chromatogram of racemic 4-nitro-3-phenylbutanal (top) and enzymatically obtained 4-nitro-3-phenylbutanal using *in situ* generated **5** via route II (bottom).



## References

1. L. Biewenga, T. Saravanan, A. Kunzendorf, J.-Y. van der Meer, T. Pijning, P. G. Tepper, R. van Merkerk, S. J. Charnock, A.-M. W. H. Thunnissen, G. J. Poelarends, *ACS Catal.* 2019, 9, 1503–1513.
2. E. Zandvoort, B.-J. Baas, W. J. Quax, G. J. Poelarends, *Chembiochem* 2011, 12, 602–609.
3. W. J. Waddell, *J. Lab. Clin. Med.* 1956, 48, 311–314.
4. J. P. Huddleston, G. K. Schroeder, K. A. Johnson, C. P. Whitman, *Biochemistry* 2012, 51, 9420–9435.
5. G. J. Poelarends, W. H. Johnson, A. G. Murzin, C. P. Whitman, *J. Biol. Chem.* 2003, 278, 48674–48683.
6. M. Pohl, P. Siegert, K. Mesch, H. Bruhn, J. Grotzinger, *Eur. J. Biochem.* 1998, 257, 538–546.
7. H. Gotoh, H. Ishikawa, Y. Hayashi, *Org. Lett.* 2007, 9, 5307–5309.
8. E. Zandvoort, E. M. Geertsema, B.-J. J. Baas, W. J. Quax, G. J. Poelarends, *Angew. Chem. Int. Ed. Engl.* 2012, 51, 1240–1243.
9. C. Guo, M. Saifuddin, T. Saravanan, M. Sharifi, G. J. Poelarends, *ACS Catal.* 2019, 9, 4369–4373.
10. S. H. Wiedemann, J. A. Ellman, R. G. Bergman, *J. Org. Chem.* 2006, 71, 1969–1976.
11. E. M. Geertsema, Y. Miao, P. G. Tepper, P. de Haan, E. Zandvoort, G. J. Poelarends, *Chem. Eur. J.* 2013, 19, 14407–14410.





# Chapter 6

## *Enantioselective Aldol Addition of Acetaldehyde to Aromatic Aldehydes Catalyzed by Proline-based Carboligases*

---

Mohammad Saifuddin,<sup>a,†</sup> Chao Guo,<sup>a,†</sup> Lieuwe Biewenga,<sup>a</sup> Thangavelu Saravanan,<sup>a</sup> Simon J. Charnock,<sup>b</sup> and Gerrit J. Poelarends<sup>a,\*</sup>

<sup>a</sup>Department of Chemical and Pharmaceutical Biology, Groningen Research Institute of Pharmacy, University of Groningen, Antonius Deusinglaan 1, 9713 AV Groningen, The Netherlands.

<sup>b</sup>Prozomix Ltd., Station Court, Haltwhistle, Northumberland NE49 9HN, U.K.

<sup>†</sup>M.S. and C.G. contributed equally to this work.

\*Corresponding author. Tel.: +31503633354; E-mail: [g.j.poelarends@rug.nl](mailto:g.j.poelarends@rug.nl);

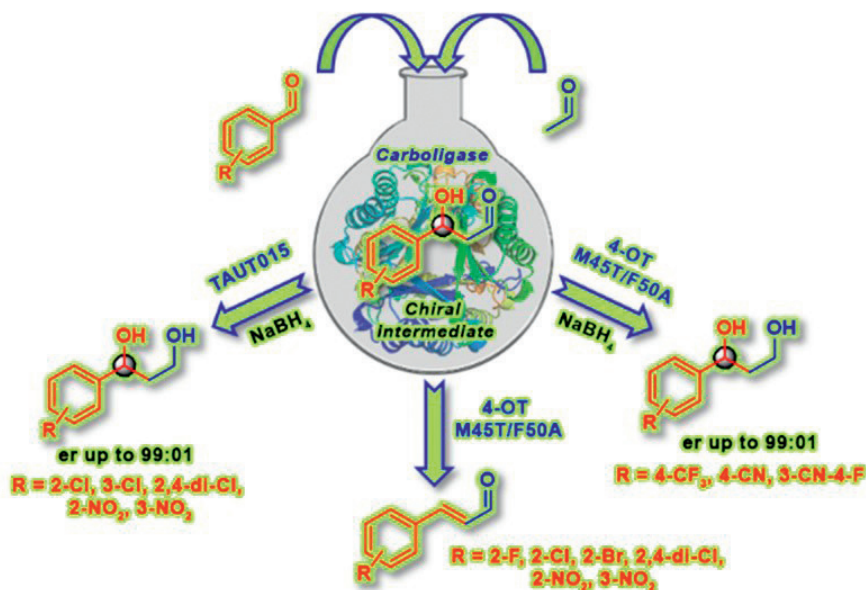
Web: <http://www.rug.nl/staff/g.j.poelarends/>

## Abstract

Aromatic  $\beta$ -hydroxyaldehydes, 1,3-diols, and  $\alpha,\beta$ -unsaturated aldehydes are valuable precursors to biologically active natural products and drug molecules. Herein we report the biocatalytic aldol condensation of acetaldehyde with various aromatic aldehydes to give a number of aromatic  $\alpha,\beta$ -unsaturated aldehydes using a previously engineered variant of 4-oxalocrotonate tautomerase [4-OT(M45T/F50A)] as carbologase. Moreover, an efficient one-pot two-step chemoenzymatic route towards chiral aromatic 1,3-diols has been developed. This one-pot chemoenzymatic strategy successfully combined a highly enantioselective aldol addition step catalyzed by a proline-based carbologase [4-OT(M45T/F50A) or TAUT015] with a chemical reduction step to convert enzymatically prepared aromatic  $\beta$ -hydroxyaldehydes into the corresponding 1,3-diols with high optical purity (e.r. up to >99:1) and in good isolated yield (51-92%). These developed (chemo)enzymatic methodologies offer alternative synthetic choices to prepare a variety of important drug precursors.

## Keywords

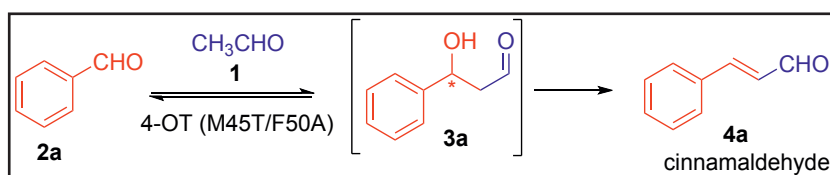
Biocatalysis, aldol reaction, aldolases,  $\beta$ -hydroxyaldehydes, carbologases



## Introduction

The aldol reaction is one of the most important methodologies for the straightforward construction of carbon-carbon bonds in synthetic organic chemistry.<sup>1</sup> This reaction allows complex chiral compounds to be rapidly assembled from simpler building blocks. Enzyme catalysis by using native and engineered aldolases offers a powerful strategy to perform this transformation.<sup>2</sup> With the discovery and engineering of novel aldolase activities, an increasing number of applications of aldolases in stereoselective synthesis have been reported.<sup>2e,2i</sup> Despite the rapidly expanding biocatalytic toolbox, aldolases that can use various aromatic aldehydes as acceptors for acetaldehyde addition are rare.<sup>2f-h</sup> Such aldolases would provide an attractive synthetic tool for the preparation of chiral precursors to various biologically active compounds and drug molecules.

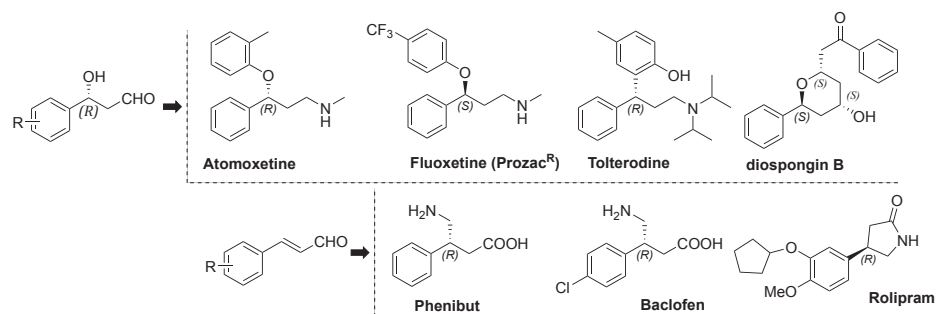
Our group has previously reported that the enzyme 4-oxalocrotonate tautomerase (4-OT) from *Pseudomonas putida* mt-2 promiscuously catalyzes the cross-condensation of acetaldehyde (**1**) with benzaldehyde (**2a**) to give cinnamaldehyde (**4a**) (Scheme 1).<sup>3</sup> It has been shown that 4-OT catalyzes both the initial aldol addition to yield 3-hydroxy-3-phenylpropanal (**3a**) as well as the subsequent rapid dehydration of **3a** to give **4a**, preventing the accumulation of chiral aldol adduct **3a** in the reaction mixture.<sup>3a</sup> Rahimi and co-workers applied mutability landscape-guided protein engineering to further improve this promiscuous aldolase activity of 4-OT. This resulted in the identification of a double mutant, 4-OT M45T/F50A, with a remarkable 3300-fold improved catalytic efficiency (in terms of  $k_{cat}/K_m$ ) over wild-type 4-OT.<sup>3b</sup>



**Scheme 1.** 4-OT catalyzed aldol condensation of acetaldehyde (**1**) with benzaldehyde (**2a**) to yield cinnamaldehyde (**4a**).

Although cinnamaldehydes are valuable synthons in their own right, we were especially interested in the enzymatic preparation of chiral aromatic  $\beta$ -hydroxyaldehydes, and the corresponding 1,3-diols, which are important precursors to various natural products, bioactive compounds and drugs (Figure 1).<sup>4,5</sup> We envisioned that these chiral aldol adducts could be constructed by the enzymatic addition of acetaldehyde to selected aromatic aldehydes, yielding  $\beta$ -hydroxyaldehydes that are not (or slowly) dehydrated,

allowing their accumulation in the reaction mixture. Chemical reduction of the enzymatically prepared aromatic  $\beta$ -hydroxyaldehydes would provide access to the corresponding aromatic 1,3-diols, which are key intermediates in the synthesis of drugs such as Fluoxetine (Figure 1).<sup>5</sup>

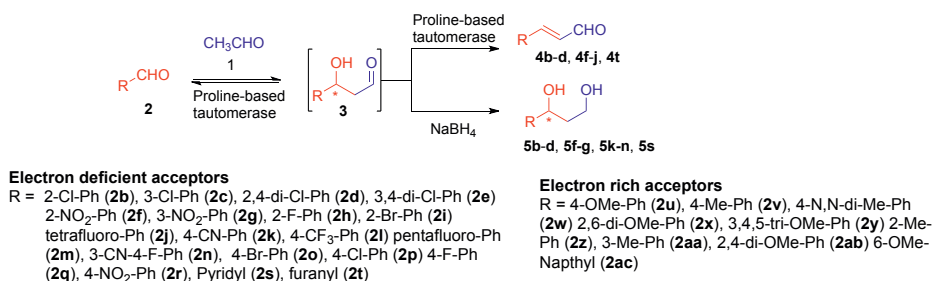


**Figure 1.** Chiral aromatic  $\beta$ -hydroxyaldehydes and cinnamaldehydes are valuable precursors to biologically active compounds and drugs.

Herein we report the 4-OT(M45T/F50A)-catalyzed synthesis of various aromatic  $\alpha,\beta$ -unsaturated aldehydes, starting from acetaldehyde and a number of aromatic aldehydes, in good isolated yields. We also report the one-pot, two-step chemoenzymatic asymmetric synthesis of various aromatic 1,3-diols with high optical purity (e.r. up to >99:1) and in good isolated yield (51-92%). This chemoenzymatic strategy highlights a highly stereoselective aldol addition of acetaldehyde to diverse aromatic aldehydes, catalyzed by either 4-OT(M45T/F50A) or the newly identified carboligase TAUT015, yielding various enantioenriched aromatic  $\beta$ -hydroxyaldehydes, which are chemically reduced into the corresponding 1,3-diols in the same pot. The applied carboligases accept a broad range of aromatic aldehydes for acetaldehyde additions, offering alternative synthetic choices to prepare important drug precursors.

We previously reported that an engineered variant of 4-OT (mutant M45T/F50A) can efficiently catalyze the aldol condensation of acetaldehyde (**1**) with benzaldehyde (**2a**) to yield cinnamaldehyde (**4a**).<sup>3b</sup> This prompted us to start our investigations by testing a large set of twenty-five aromatic aldehydes (**2b-y** and **2ac**) as potential aldol acceptor substrates in the 4-OT(M45T/F50A)-catalyzed aldol reaction with donor substrate **1** (Scheme 2). The results demonstrate that the 4-OT(M45T/F50A) enzyme has a remarkably broad substrate scope, accepting nineteen aromatic aldehydes with electron withdrawing substituents (**2b-t**) as aldol acceptor substrates (Scheme 2, Figures S1-S19). On the contrary, aromatic aldehydes with electron-donating substituents (**2u-y** and **2ac**) were not, or very poorly, accepted by 4-OT(M45T/F50A) (Scheme 2, Figures

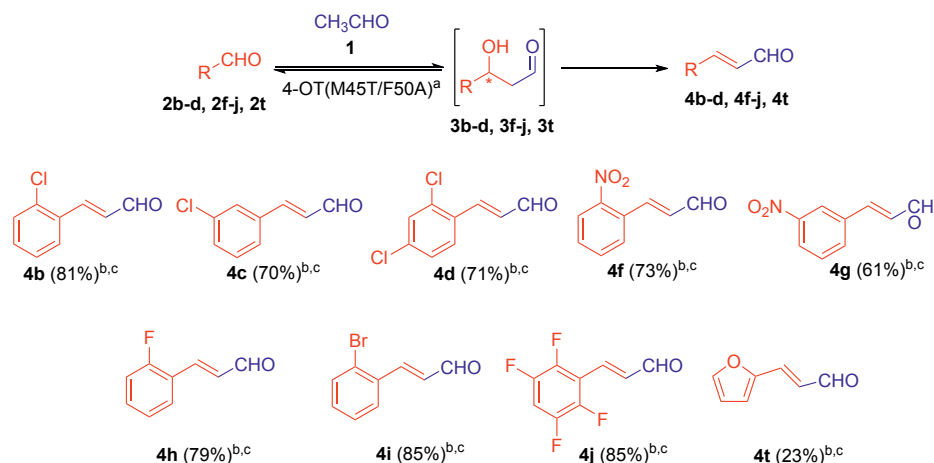
S20-S24, S28). While this may be attributed to unfavorable equilibrium constants for these aromatic aldehydes in the aldol addition with acetaldehyde, the electron donating properties of the substituents on the aromatic ring likely also lead to weak activation of the carbonyl carbon for nucleophilic attack.<sup>2j,2l</sup>



**Scheme 2.** Enzymatic aldol reaction of acetaldehyde (**1**) with various aromatic aldehydes (**2**) using promiscuous proline-based tautomerasases as carboliogases.

Out of the nineteen well-accepted aromatic aldehydes, fourteen substrates undergo an aldol condensation reaction with formation of the corresponding  $\alpha,\beta$ -unsaturated aldehydes, whereas five substrates (**2k-n** and **2s**, see below) undergo conversion without significant formation of the corresponding  $\alpha,\beta$ -unsaturated aldehydes (Figure S1-S19). To demonstrate the synthetic usefulness of 4-OT(M45T/F50A) for the preparation of  $\alpha,\beta$ -unsaturated aldehydes, nine aromatic aldehydes (**2b-d**, **2f-j** and **2t**) that could be efficiently converted by 4-OT(M45T/F50A) were selected and used in semi-preparative scale synthesis (Scheme 3). The corresponding  $\alpha,\beta$ -unsaturated aldehydes (**4b-d**, **4f-j** and **4t**) could be isolated in moderate to good yields (Scheme 3, Figures S30-S47). Notably, *ortho*-substituted benzaldehydes (e.g. **2b**, **2f**, **2h** and **2i**) were better accepted by 4-OT(M45T/F50A) than *para*-substituted benzaldehydes (**2p**, **2r**, **2o** and **2q**).<sup>6</sup> These results underscore the potential of 4-OT(M45T/F50A) for the synthesis of versatile aromatic  $\alpha,\beta$ -unsaturated aldehydes, which are important precursors to various abundantly prescribed drugs (Figure 1).<sup>4f</sup>





**Scheme 3.** Synthesis of aromatic  $\alpha,\beta$ -unsaturated aldehydes from simpler building blocks using the carbiligase 4-OT(M45T/F50A). <sup>a</sup>Reaction conditions: reaction mixtures consisted of 100 mM **1**, 3 mM **2** and 0.1 mg/mL 4-OT (M45T/F50A) in 20 mM sodium phosphate buffer, 5% DMSO v/v, pH 7.3. The reaction volume was 60 mL. <sup>b</sup>Reaction time (48 h, 72 h for **2t**). <sup>c</sup>Purified by silica gel column chromatography.

Interestingly, for the 4-OT(M45T/F50A)-catalyzed aldol addition of **1** to **2k-n** and **2s**, depletion of the starting substrates was observed without significant accumulation of the corresponding  $\alpha,\beta$ -unsaturated aldehydes (Figures S10-S13, S18). These results led us to hypothesize that in these instances the initially formed aldol adducts,  $\beta$ -hydroxyaldehydes **3k-n** and **3s**, were not (or slowly) dehydrated and accumulated in the reaction mixture. To confirm this hypothesis, an analytical scale experiment was performed using 4-OT(M45T/F50A) in NaPi buffer (0.3 mL, pH 7.3, 5% DMSO), containing **1** (50 mM) and **2k** (2 mM). Reaction progress was monitored by following the depletion of **2k** by UV-VIS spectrophotometry, and  $\text{NaBH}_4$  was added to the reaction mixture (after 1 h) to convert the aldol adduct into the more stable 1,3-diol **5k**. Analysis of chemoenzymatic product **5k** by HPLC on a chiral stationary phase, using authentic standards with known absolute configuration, revealed high enantiocontrol at the site of addition with formation of the *R* enantiomer (e.r. = 94:6). These results demonstrate that the enzyme 4-OT(M45T/F50A) can indeed catalyze enantioselective aldol additions.

Having established that 4-OT(M45T/F50A) has potential for the asymmetric synthesis of enantioenriched aromatic  $\beta$ -hydroxyaldehydes, we first optimized the reaction conditions by varying the pH and the concentration of acetaldehyde, buffer and enzyme (Figure S88). Using the optimized conditions, the one-pot two-step chemoenzymatic cascade synthesis of aromatic 1,3-diols **5k-n** and **5s**, via enzymatically prepared  $\beta$ -hydroxyaldehydes **3k-n** and **3s**, was achieved with high enantioselectivity (e.r. up to 99:1) and in good isolated yield (63-92%) (Table 1, Figures S48-S62).

**Table 1.** Asymmetric chemoenzymatic synthesis of aromatic 1,3-diols using 4-OT(M45T/F50A) as carboliase.<sup>a</sup>

Entry	Aldol acceptor	Product	Reaction time [h] <sup>b</sup>	Isolated yield (%) <sup>c</sup>	e.r. <sup>d</sup>	Abs. Conf. <sup>e</sup>
1			6	92	91:9	<i>R</i> <sup>e</sup>
2			7	63	90:10	<i>R</i> <sup>e</sup>
3			2.5	88	90:10	<i>R</i> <sup>e</sup>
4			5	70	92:8	<i>R</i> <sup>f</sup>
5			1	65	99:1	<i>R</i> <sup>f</sup>

<sup>a</sup>Assay conditions: acetaldehyde (**1**, 100 mM), aromatic aldehyde (**2k-n**, **2s**, 2 mM) and purified 4-OT(M45T/F50A) (0.125-0.475 mg/mL) in 20 mM sodium phosphate buffer (MOPS for **2k**), with 5% DMSO (v/v) as cosolvent, at pH 8.0 (pH 7.3 for **2s** to increase its electrophilicity) and room temperature. The reaction volume was 40 mL. To reduce the aldehyde functionality of the enzymatic products **3k-n** and **3s**, NaBH<sub>4</sub> (30 mM) was added to the reaction mixture followed by incubation for 3 h at room temperature. <sup>b</sup>Reaction time of the enzymatic step. <sup>c</sup>Isolated yield of the aromatic 1,3-diol product. <sup>d</sup>Determined by HPLC analysis on a chiral stationary phase using chemically synthesized racemic standards. <sup>e</sup>The absolute configuration was determined by chiral HPLC using chemically prepared authentic standards with known (*R*) or (*S*) configuration. <sup>f</sup>The absolute configuration was assigned based on analogy.

To expand the number of enzymatically accessible  $\beta$ -hydroxyaldehyde products (e.g. **3b-d**, **3f** and **3g**) in principle two different approaches can be used: (i) the use of protein engineering to create new variants of 4-OT(M45T/F50A) that lack dehydration activity and possess unchanged or enhanced aldol addition activity, or (ii) the use of homologous enzymes that possess the desired aldol addition activity but lack dehydration activity. To selectively eliminate the dehydration activity of 4-OT(M45T/F50A), a structure-based approach is favored, which awaits the determination of the enzyme crystal structure complexed with one of the  $\beta$ -hydroxyaldehyde products (work in progress). In this study, we therefore have chosen to screen a panel of fifty commercially available 4-OT homologues for enzymes with the desired aldol addition activity to yield enantioenriched  $\beta$ -hydroxyaldehydes. For initial activity testing, we used donor substrate **1** and aldol acceptor substrate **2b**. Interestingly, out of the fifty tested 4-OT homologues, three enzymes (TAUT015, TAUT021 and TAUT028) were found to show significant activity towards the aldol addition of **1** to **2b** to give the desired product **3b** without promoting significant dehydration leading to formation of **4b** (Figure S29). Notably, the same reaction catalyzed by 4-OT(M45T/F50A) resulted in the formation of  $\alpha,\beta$ -unsaturated aldehyde **4b** as the sole product, providing additional support that the  $\beta$ -hydroxyaldehyde dehydration step is enzyme catalyzed.

We next investigated the substrate scope of the best performing enzyme, TAUT015, which shows 28% overall sequence identity to 4-OT (Figure S89), in additions using **1** as donor substrate and a number of selected benzaldehyde derivatives as potential aldol acceptor substrates. The results show that, like 4-OT(M45T/F50A), TAUT015 does not accept benzaldehydes with electron donating substituents (**2u**, **2v**, **2z**, **2aa** and **2ab**) on the aromatic ring (Figures S25-S27). However, several benzaldehydes with electron withdrawing substituents (**2b-d**, **2f** and **2g**) were well accepted by TAUT015. Encouraged by these findings, semi-preparative scale reactions were carried out under optimized reaction conditions. The reaction progress (**1** with **2b-d**, **2f** and **2g**) was monitored by UV-VIS spectrophotometry and the enzymatically prepared  $\beta$ -hydroxyaldehydes **3b-d**, **3f** and **3g** were reduced to the corresponding aromatic 1,3-diols (**5b-d**, **5f** and **5g**) using  $\text{NaBH}_4$ . These one-pot, two-step chemoenzymatically prepared compounds **5b-d**, **5f** and **5g** were obtained in good isolated yields (up to 72%) and with high e.r. values (e.r. up to >99:1) (Table 2, Figures S63-S87). Note that enzymatic access to these useful chiral synthons requires the use of TAUT015 as carboligase and is not possible with 4-OT(M45T/F50A). Hence, the enzyme TAUT015 nicely supplements the toolbox of biocatalysts for the production of important chiral aromatic  $\beta$ -hydroxyaldehydes and corresponding 1,3-diols.

**Table 2.** Asymmetric chemoenzymatic synthesis of aromatic 1,3-diols using TAUT015 as carboligase.<sup>a</sup>

Entry	Aldol acceptor	Product	Reaction time [h] <sup>b</sup>	Isolated yield (%) <sup>c</sup>	e.r. <sup>d</sup>	Abs. Conf. <sup>e</sup>
1			4.5	72	>99:1	R
2			8.3	51	94:6	R
3			3.3	62	99:1	R
4			6.5	69	99:1	R
5			11	56	89:11	R

<sup>a</sup>Assay conditions: acetaldehyde (**1**, 100 mM), benzaldehyde derivative (**2b-d**, **2f-g**, 2 mM) and enzyme (crude cell extract, 0.33 mg/mL) in 20 mM sodium phosphate buffer, containing 5% EtOH (v/v) as cosolvent, at pH 6.5 and room temperature. The specific enzyme activity was determined to be 0.044 U/mg CFE. Unit definition: 1 U of enzyme converts 1  $\mu$ mol of **2b** in 1 minute; assay conditions: 20 mM NaPi pH 6.5, 5% EtOH, 100 mM **1**, 2 mM **2b**. The reaction volume was 40 mL. To reduce the aldehyde functionality of **3b-d** and **3f-g**, NaBH<sub>4</sub> (30 mM) was added to the reaction mixture followed by incubation for 3 h at room temperature. <sup>b</sup>Reaction time of the enzymatic step. <sup>c</sup>Isolated yield of the aromatic 1,3-diol product. <sup>d</sup>Determined by HPLC analysis on a chiral stationary phase using chemically synthesized racemic standards. <sup>e</sup>The absolute configuration was determined by chiral HPLC using chemically prepared authentic standards with known (R) or (S) configuration.

In conclusion, our results indicate that members of the 4-OT family of enzymes, which possess an uncommon catalytic amino-terminal proline, can function as novel carboligases, promoting the cross-aldol reaction of acetaldehyde with diverse aromatic

aldehydes to give access to a range of important  $\alpha,\beta$ -unsaturated aldehydes as well as enantioenriched  $\beta$ -hydroxyaldehydes. It is feasible that their promiscuous aldol addition activities, which are beyond the synthetic scope of currently known aldolases, can be further optimized by directed evolution or computational redesign. It is important to emphasize that the controlled cross-aldol addition of acetaldehyde is a particularly challenging synthetic aim because acetaldehyde is a relatively reactive and difficult to tame chemical.<sup>5,7</sup> Remarkably, the promiscuous 4-OT enzymes accept acetaldehyde as a nucleophile in aldol additions, which is unparalleled among the aldolases, with the notable exception of 2-deoxy-D-ribose-5-phosphate aldolase (DERA) and d-fructose-6-phosphate aldolase (FSA), which are also able to use acetaldehyde as a nucleophilic substrate.<sup>2g,2k,2o</sup> As such, these proline-based 4-OT enzymes nicely complement the toolbox of biocatalysts for (asymmetric) C-C bond-formation, and open up new opportunities to develop practical enzymatic processes for the more sustainable and step-economic synthesis of valuable drug precursors starting from simple building blocks. Further screening of related proline-based enzymes might prove to be a valuable approach to discover new aldolase activities that could be exploited to develop synthetically useful biocatalysts for carbon-carbon bond formation.

### Author Information

Corresponding Author

\*Corresponding author. Tel.: +31503633354; E-mail: g.j.poelarends@rug.nl; Web: <http://www.rug.nl/staff/g.j.poelarends/>

### Author Contributions

†These authors contributed equally: Chao Guo and Mohammad Saifuddin.

### Associated Content

**Supporting Information.** Additional experimental procedures and characterizations of compounds. This information is available free of charge on the ACS Publications website.

### Acknowledgement

We acknowledge financial support from the Netherlands Organization of Scientific Research (VICI grant 724.016.002), the European Research Council (PoC grant 713483), and the European Union's Horizon 2020 Research and Innovation Programme under grant agreement No 635595 (CarbaZymes).

## References

1. (a) Gröger, H.; Vogl, E. M.; Shibasaki, M. New Catalytic Concepts for the Asymmetric Aldol Reaction. *Chem. Eur. J.* **1998**, *4*, 1137-1141; (b) Palomo, C.; Oiarbide, M.; García, J. M. Current Progress in the Asymmetric Aldol Addition Reaction. *Chem. Soc. Rev.* **2004**, *33*, 65-75; (c) Yamashita, Y.; Yasukawa, T.; Yoo, W. J.; Kitanosono, T.; Kobayashi, S. Catalytic Enantioselective Aldol Reactions. *Chem. Soc. Rev.* **2018**, *47*, 4388-4480.
2. (a) Wong, C. H.; Junceda, E. G.; Chen, L.; Blanco, O.; Gijssen, H. J. M.; Steensma, D. H. Recombinant 2-Deoxyribose-5-phosphate Aldolase in Organic Synthesis: Use of Sequential Two-Substrate and Three-Substrate Aldol Reactions. *J. Am. Chem. Soc.* **1995**, *117*, 3333-3339; (b) Fessner, W. D. Enzyme Mediated C-C Bond Formation. *Curr. Opin. Chem. Biol.* **1998**, *2*, 85-97; (c) Machajewski, T. D.; Wong, C. H. The Catalytic Asymmetric Aldol Reaction. *Angew. Chem. Int. Ed.* **2000**, *39*, 1352-1374; (d) DeSantis, G.; Liu, J.; Clark, D. P.; Heine, A.; Wilson, I. A.; Wong, C. H. Structure-Based Mutagenesis Approaches Toward Expanding the Substrate Specificity of D-2-Deoxyribose-5-phosphate Aldolase. *Bioorg. Med. Chem.* **2003**, *11*, 43-52; (e) Clapés, P.; Garrabou, X. Current Trends in Asymmetric Synthesis with Aldolases. *Adv. Synth. Catal.* **2011**, *353*, 2263-2283; (f) Fesko, K.; Khadjawi, M. G. Biocatalytic Methods for C-C Bond Formation. *ChemCatChem.* **2013**, *5*, 1248-1272; (g) Clapés, P.; Fessner, W. D.; Sprenger, G. A.; Samland, A. K. Recent Progress in Stereoselective Synthesis with Aldolases. *Curr. Opin. Chem. Biol.* **2010**, *14*, 154-167; (h) Chen, Q.; Chen, X.; Cui, Y.; Ren, J.; Lu, W.; Feng, J.; Wu, Q.; Zhu, D. A New D-Threonine Aldolase as a Promising Biocatalyst for Highly Stereoselective Preparation of Chiral Aromatic  $\beta$ -Hydroxy- $\alpha$ -Amino acids. *Catal. Sci. Technol.* **2017**, *7*, 5964-5973; (i) Schmidt, N. G.; Eger, E.; Kroutil, W. Building Bridges: Biocatalytic C-C-Bond Formation toward Multifunctional Products. *ACS Catal.* **2016**, *6*, 4286-4311; (j) Obexer, R.; Godina, A.; Garrabou, X.; Mittl, P. R. E.; Baker, D.; Griffiths, A. D.; Hilvert, D. Emergence of a Catalytic Tetrad during Evolution of a Highly Active Artificial Aldolase. *Nat. Chem.* **2017**, *9*, 50-56; (k) Roldán, R.; Hernandez, K.; Joglar, J.; Bujons, J.; Parella, T.; Sánchez-Moreno, I.; Hélaïne, V.; Lemaire, M.; Guérard-Hélaïne, C.; Fessner, W. D.; Clapés, P. Biocatalytic Aldol Addition of Simple Aliphatic Nucleophiles to Hydroxyaldehydes. *ACS Catal.* **2018**, *8*, 8804-8809; (l) Xie, Z. B.; Wang, N.; Jiang, G. F.; Yu, X. Q. Biocatalytic Asymmetric Aldol Reaction in Buffer Solution. *Tetrahedron Lett.* **2013**, *54*, 945-948; (m) Roldán, R.; Hernández, K.; Joglar, J.; Bujons, J.; Parella, T.; Fessner, W. D.; Clapés, P. Aldolase-Catalyzed Asymmetric Synthesis of N-Heterocycles by Addition of Simple Aliphatic Nucleophiles to Aminoaldehydes. *Adv. Synth. Catal.* **2019**, *361*, 2673-2687; (n) Junker, S.; Roldan, R.; Joosten, H. J.; Clapés, P.; Fessner, W. D. Complete Switch of Reaction Specificity of an Aldolase by Directed Evolution In Vitro: Synthesis of Generic Aliphatic Aldol Products. *Angew. Chem. Int. Ed.* **2018**, *57*, 10153-10157; (o) Roldán, R.; Sanchez-Moreno, I.; Scheidt, T.; Hélaïne, V.; Lemaire, M.; Parella, T.; Clapés, P.; Fessner, W. D.; Guérard-Hélaïne, C. Breaking the Dogma of Aldolase Specificity: Simple Aliphatic Ketones and Aldehydes are Nucleophiles for Fructose-6-phosphate Aldolase. *Chem. Eur. J.* **2017**, *23*, 5005-5009.

3. (a) Zandvoort, E.; Geertsema, E. M.; Quax, W. J.; Poelarends, G. J. Enhancement of the Promiscuous Aldolase and Dehydration Activities of 4-Oxalocrotonate Tautomerase by Protein Engineering. *ChemBioChem* **2012**, *13*, 1274-1277; (b) Rahimi, M.; van der Meer, J. Y.; Geertsema, E. M.; Poddar, H.; Baas, B. J.; Poelarends, G. J. Mutations Closer to the Active Site Improve the Promiscuous Aldolase Activity of 4-Oxalocrotonate Tautomerase More Effectively than Distant Mutations. *ChemBioChem* **2016**, *17*, 1225-1228; (c) Rahimi, M.; van der Meer, J. Y.; Geertsema, E. M.; Poelarends, G. J. Engineering a Promiscuous Tautomerase into a More Efficient Aldolase for Self-Condensations of Linear Aliphatic Aldehydes. *ChemBioChem* **2017**, *18*, 1435-1441; (d) Rahimi, M.; Geertsema, E. M.; Miao, Y.; van der Meer, J. Y.; van den Bosch, T.; de Haan, P.; Zandvoort, E.; Poelarends, G. J. Inter- and Intramolecular Aldol Reactions Promiscuously Catalyzed by a Proline-based Tautomerase. *Org. Biomol. Chem* **2017**, *15*, 2809-2816.
4. (a) Rej, R. K.; Das, T.; Hazra, S.; Nanda, S. Chemoenzymatic Asymmetric Synthesis of Fluoxetine, Atomoxetine, Nisoxetine, and Duloxetine. *Tetrahedron: Asymmetry* **2013**, *24*, 913-918; (b) Wenthur, C. J.; Bennett, M. R.; Lindsley, C. W. Classics in Chemical Neuroscience: Fluoxetine (Prozac). *ACS Chem. Neurosci.* **2014**, *5*, 14-23; (c) de Fátima, Â.; Lapis, A. A. M.; Pilli, R. A. A Concise Total Synthesis of (R)-Fluoxetine, a Potent and Selective Serotonin Reuptake Inhibitor. *J. Braz. Chem. Soc.* **2005**, *16*, 495-499; (d) Roesner, S.; Aggarwal, V. K. Enantioselective synthesis of (R)-tolterodine using lithiation/borylation-protodeboronation methodology. *Can. J. Chem.* **2012**, *90*, 1-10; (e) Sabitha, G. Padmaja, P.; Yadav, J. S. A Concise Total Synthesis of Diospongins A and B. *Helvetica Chimica Acta.* **2008**, *91*, 2235-2239; (f) Guo, C.; Saifuddin, M.; Saravanan, T.; Sharifi, M.; Poelarends, G. J. Biocatalytic Asymmetric Michael Additions of Nitromethane to  $\alpha,\beta$ -Unsaturated Aldehydes via Enzyme-bound Iminium Ion Intermediates. *ACS Catal.* **2019**, *9*, 4369-4373; (g) Fan, X.; Rodriguez-Esrich, C.; Wang, S.; Sayalero, S.; Pericas, M. A. Highly Enantioselective Cross-Aldol Reactions of Acetaldehyde Mediated by a Dual Catalytic System Operating under Site Isolation. *Chem. Eur. J.* **2014**, *20*, 13089-13093.
5. Hayashi, Y.; Itoh, T.; Aratake, S.; Ishikawa, H. A Diarylprolinol in an Asymmetric, Catalytic, and Direct Crossed- Aldol Reaction of Acetaldehyde. *Angew. Chem. Int. Ed.* **2008**, *47*, 2082-2084.
6. Martínez, A.; Van Gemmeren, M.; List, B. Unexpected Beneficial Effect of ortho-Substituents on the (S)-Proline-Catalyzed Asymmetric Aldol Reaction of Acetone with Aromatic Aldehydes. *Synlett.* **2014**, *25*, 961-964.
7. Yang, J. W.; Chandler, C.; Stadler, M.; Kampen, D.; List, B. Proline-Catalysed Mannich Reactions of Acetaldehyde. *Nature* **2008**, *452*, 453-455.

## Supporting Information

### 1. Materials and Methods

#### 1.1 Materials

Chemicals were purchased from Sigma Aldrich, Across, Merck or Fluka (unless noted otherwise) and were used without further purification. The  $\alpha,\beta$ -unsaturated aldehydes **2o-r** were prepared using a previously reported method.<sup>[1a]</sup> Synthesis of racemic and enantiopure reference compounds, required for chiral analysis of enzymatic products, was done according to published protocols.<sup>[2]</sup>

#### 1.2 General methods

NMR spectra were recorded on a Bruker 500 MHz spectrometer. Enzymatic assays were performed on a V-650 or V-660 spectrophotometer from Jasco (IJsselstein, The Netherlands). Reverse phase HPLC was carried out using an in-house analytical HPLC equipped with a Shimadzu LC-10 AT pump and a Shimadzu SPD-M10A diode array detector.

#### 1.3 Protein expression and purification

The expression and purification of 4-OT(M45T/F50A) mutant was based on a protocol described elsewhere.<sup>[3]</sup> A sample of purified protein was analyzed by ESI-MS to confirm that the protein had been processed correctly and the initiating methionine had been removed. The purified protein was flash frozen in liquid nitrogen and stored at -80 °C until further use.

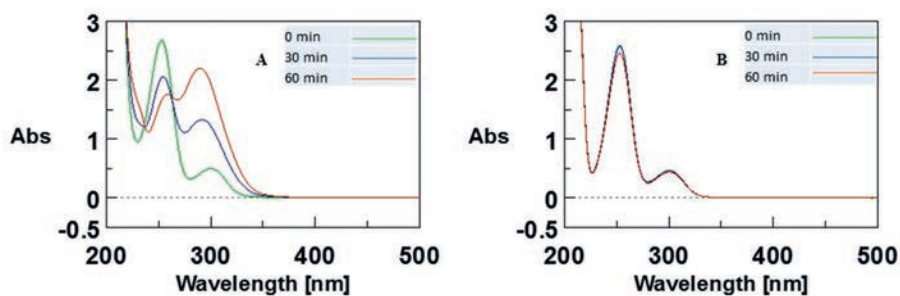
#### 1.4 Chiral analysis

The enantiomeric ratio and absolute configuration of the chemoenzymatically synthesized compounds **5b-d**, **5f-g** and **5k-m** was analyzed by chiral HPLC using chemically synthesized racemic 1,3-diols and enantiopure 1,3-diols as reference compounds.<sup>[2a-b]</sup> For compounds **5n** and **5s**, the enantiomeric ratio was analyzed by chiral HPLC using chemically synthesized racemic 1,3-diols<sup>[2a]</sup> and the absolute configuration was based on analogy.

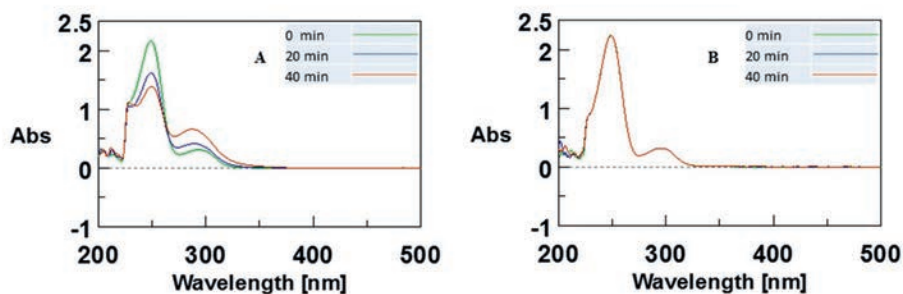
#### 1.5 Analytical scale reactions

Enzymatic and non-enzymatic reactions between substrates **1** and **2** were monitored by following the depletion in the absorbance corresponding to the consumption of **2**. Reaction conditions are given below each figure.

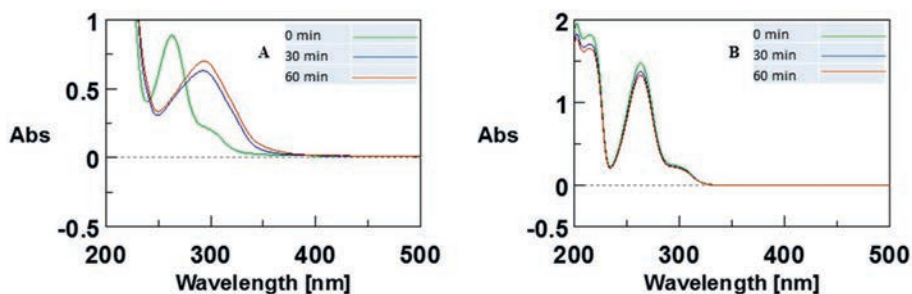




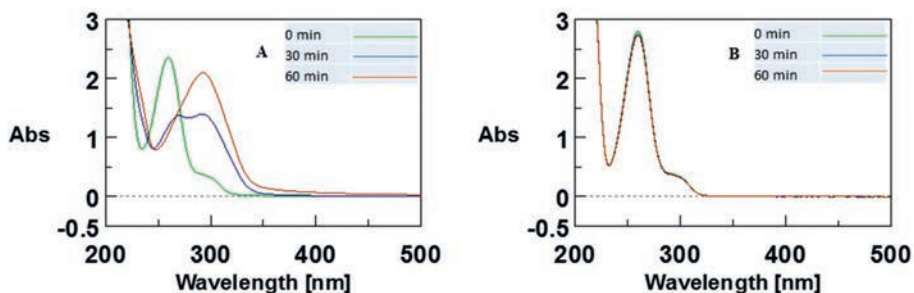
**Figure S1.** UV spectra monitoring the aldol reaction of **1** (50 mM) with **2b** (2 mM) catalyzed by 4-OT(M45T/F50A) (1 mg/mL) in buffer [20 mM NaPi/5% (v/v) EtOH, 0.3 mL] at pH 7.3. A. Enzymatic reaction. B. Non-enzymatic reaction.



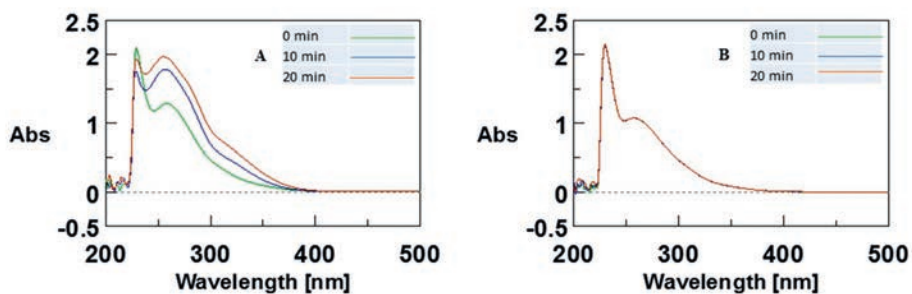
**Figure S2.** UV spectra monitoring the aldol reaction of **1** (50 mM) with **2c** (2 mM) catalyzed by 4-OT(M45T/F50A) (1 mg/mL) in buffer [20 mM NaPi/5% (v/v) DMSO, 0.3 mL] at pH 7.3. A. Enzymatic reaction. B. Non-enzymatic reaction.



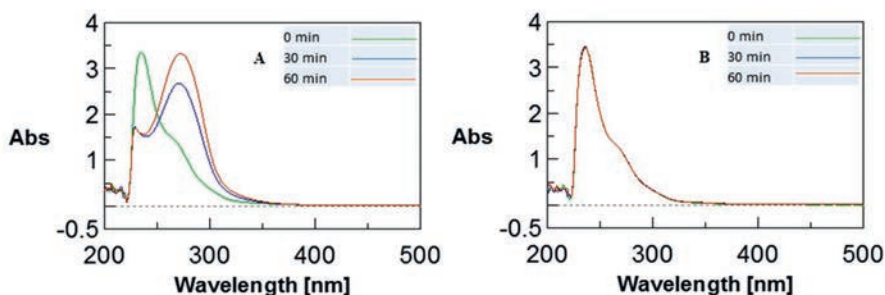
**Figure S3.** UV spectra monitoring the aldol reaction of **1** (50 mM) with **2d** (2 mM) catalyzed by 4-OT(M45T/F50A) (1 mg/mL) in buffer [20 mM NaPi/5% (v/v) EtOH, 0.3 mL] at pH 7.3. A. Enzymatic reaction. B. Non-enzymatic reaction.



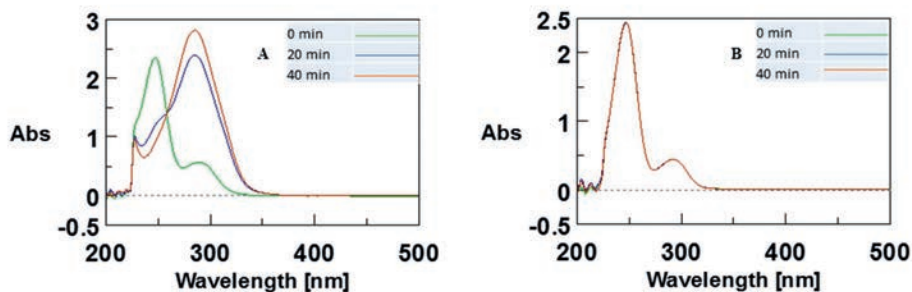
**Figure S4.** UV spectra monitoring the aldol reaction of **1** (50 mM) with **2e** (2 mM) catalyzed by 4-OT(M45T/F50A) (1 mg/mL) in buffer [20 mM NaPi/5% (v/v) EtOH, 0.3 mL] at pH 7.3. **A.** Enzymatic reaction. **B.** Non-enzymatic reaction.



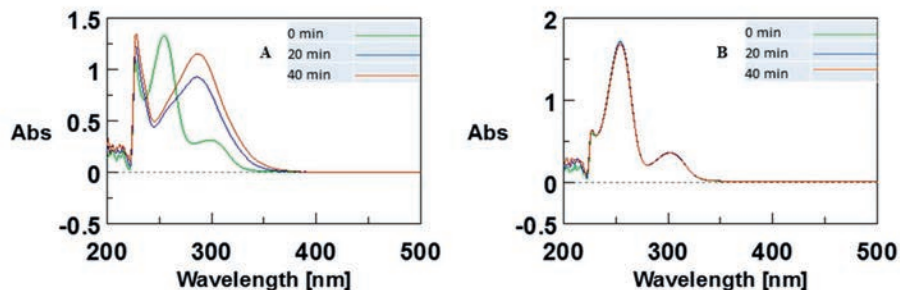
**Figure S5.** UV spectra monitoring the aldol reaction of **1** (50 mM) with **2f** (2 mM) catalyzed by 4-OT(M45T/F50A) (1 mg/mL) in buffer [20 mM NaPi/5% (v/v) DMSO, 0.3 mL] at pH 7.3. **A.** Enzymatic reaction. **B.** Non-enzymatic reaction.



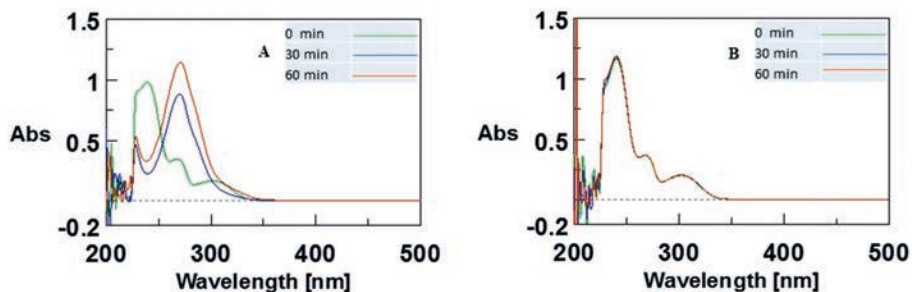
**Figure S6.** UV spectra monitoring the aldol reaction of **1** (50 mM) with **2g** (2 mM) catalyzed by 4-OT(M45T/F50A) (1 mg/mL) in buffer [20 mM NaPi/5% (v/v) DMSO, 0.3 mL] at pH 7.3. **A.** Enzymatic reaction. **B.** Non-enzymatic reaction.



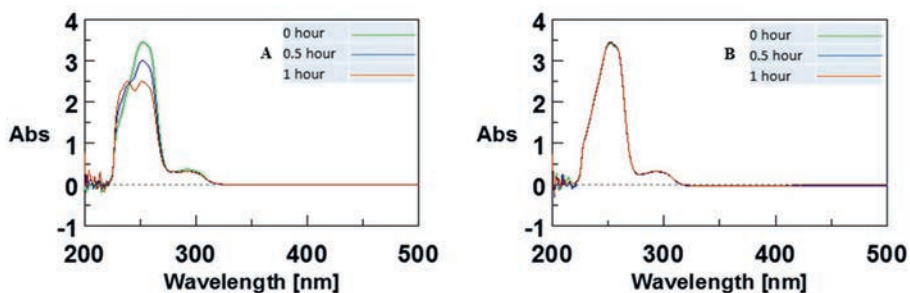
**Figure S7.** UV spectra monitoring the aldol reaction of **1** (50 mM) with **2h** (2 mM) catalyzed by 4-OT(M45T/F50A) (1 mg/mL) in buffer [20 mM NaPi/5% (v/v) DMSO, 0.3 mL] at pH 7.3. **A.** Enzymatic reaction. **B.** Non-enzymatic reaction.



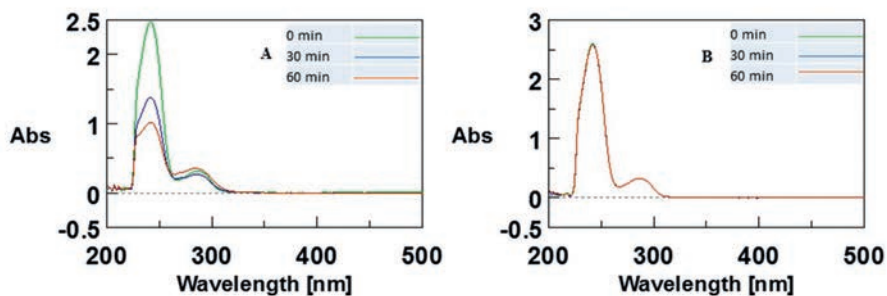
**Figure S8.** UV spectra monitoring the aldol reaction of **1** (50 mM) with **2i** (2 mM) catalyzed by 4-OT(M45T/F50A) (1 mg/mL) in buffer [20 mM NaPi/5% (v/v) DMSO, 0.3 mL] at pH 7.3. **A.** Enzymatic reaction. **B.** Non-enzymatic reaction.



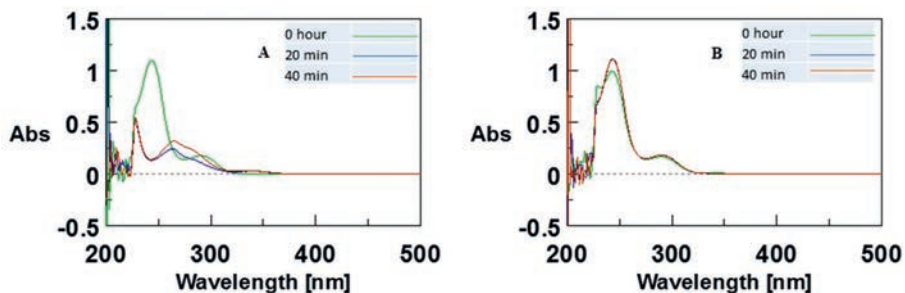
**Figure S9.** UV spectra monitoring the aldol reaction of **1** (50 mM) with **2j** (2 mM) catalyzed by 4-OT(M45T/F50A) (1 mg/mL) in buffer [20 mM NaPi/5% (v/v) DMSO, 0.3 mL] at pH 7.3. **A.** Enzymatic reaction. **B.** Non-enzymatic reaction.



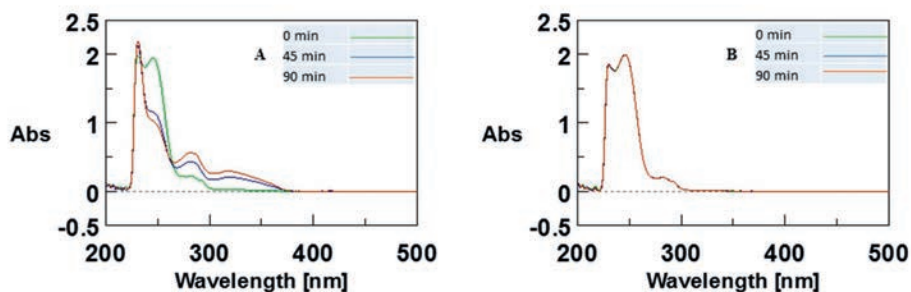
**Figure S10.** UV spectra monitoring the aldol reaction of **1** (50 mM) with **2k** (2 mM) catalyzed by 4-OT(M45T/F50A) (1 mg/mL) in buffer [20 mM NaPi/5% (v/v) DMSO, 0.3 mL] at pH 7.3. **A.** Enzymatic reaction. **B.** Non-enzymatic reaction.



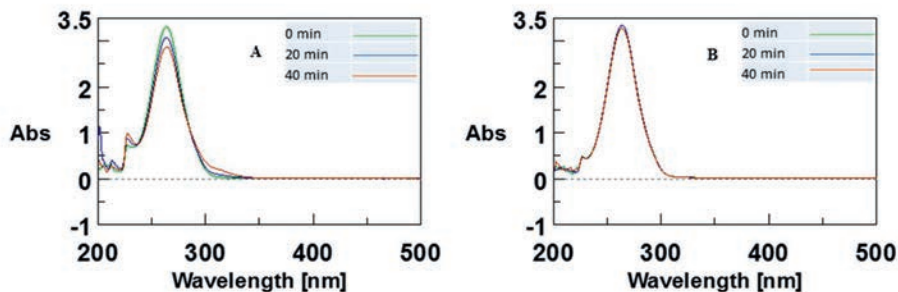
**Figure S11.** UV spectra monitoring the aldol reaction of **1** (50 mM) with **2l** (2 mM) catalyzed by 4-OT(M45T/F50A) (1 mg/mL) in buffer [20 mM NaPi/5% (v/v) DMSO, 0.3 mL] at pH 7.3. **A.** Enzymatic reaction. **B.** Non-enzymatic reaction.



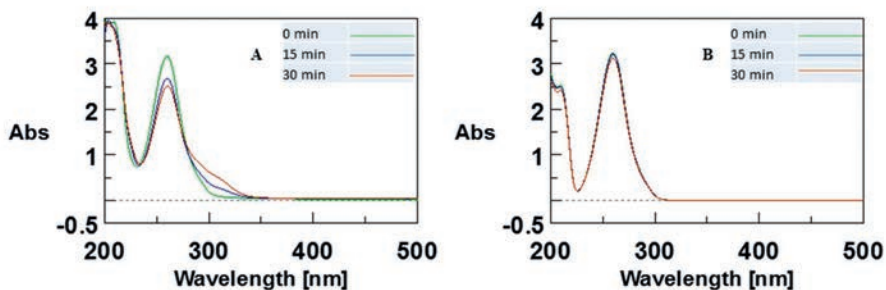
**Figure S12.** UV spectra monitoring the aldol reaction of **1** (50 mM) with **2m** (2 mM) catalyzed by 4-OT(M45T/F50A) (1 mg/mL) in buffer [20 mM NaPi/5% (v/v) DMSO, 0.3 mL] at pH 7.3. **A.** Enzymatic reaction. **B.** Non-enzymatic reaction.



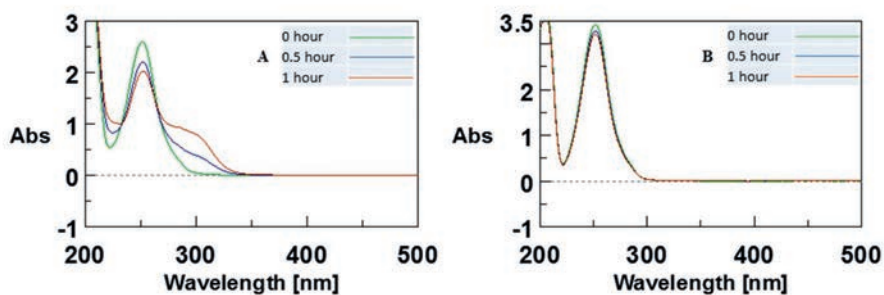
**Figure S13.** UV spectra monitoring the aldol reaction of **1** (50 mM) with **2n** (2 mM) catalyzed by 4-OT(M45T/F50A) (1 mg/mL) in buffer [20 mM NaPi/5% (v/v) DMSO, 0.3 mL] at pH 7.3. **A.** Enzymatic reaction. **B.** Non-enzymatic reaction.



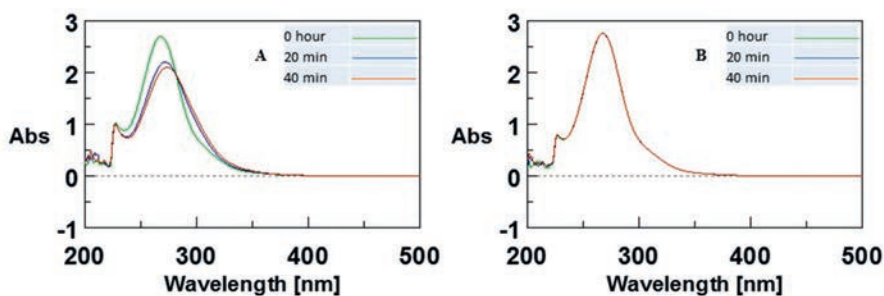
**Figure S14.** UV spectra monitoring the aldol reaction of **1** (50 mM) with **2o** (2 mM) catalyzed by 4-OT(M45T/F50A) (1 mg/mL) in buffer [20 mM NaPi/5% (v/v) DMSO, 0.3 mL] at pH 7.3. **A.** Enzymatic reaction. **B.** Non-enzymatic reaction.



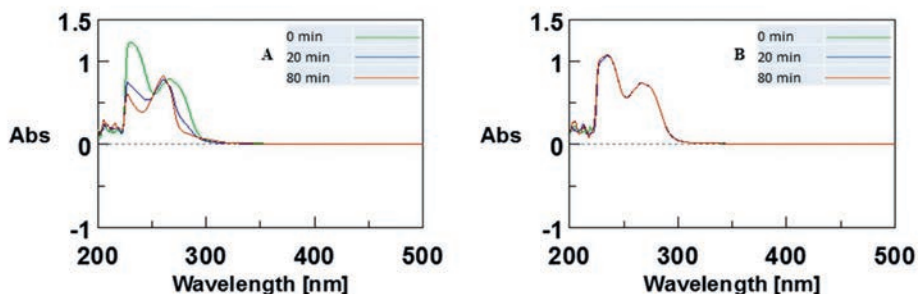
**Figure S15.** UV spectra monitoring the aldol reaction of **1** (50 mM) with **2p** (2 mM) catalyzed by 4-OT(M45T/F50A) (1 mg/mL) in buffer [20 mM NaPi/5% (v/v) EtOH, 0.3 mL] at pH 7.3. **A.** Enzymatic reaction. **B.** Non-enzymatic reaction.



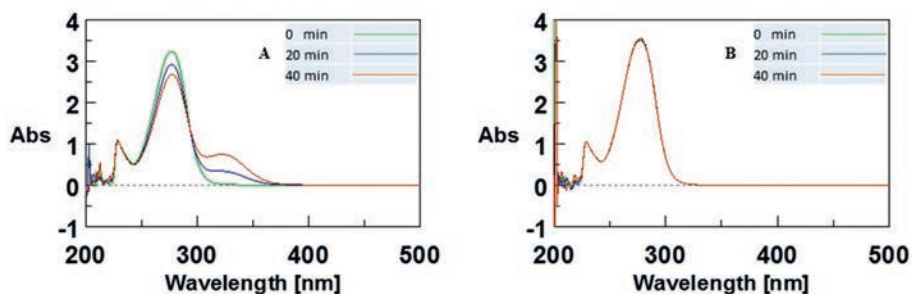
**Figure S16.** UV spectra monitoring the aldol reaction of **1** (50 mM) with **2q** (2 mM) catalyzed by 4-OT(M45T/F50A) (1 mg/mL) in buffer [20 mM NaPi/5% (v/v) EtOH, 0.3 mL] at pH 7.3. **A.** Enzymatic reaction. **B.** Non-enzymatic reaction.



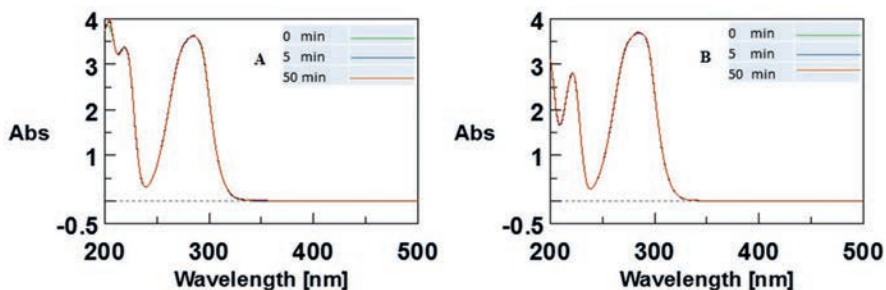
**Figure S17.** UV spectra monitoring the aldol reaction of **1** (50 mM) with **2r** (2 mM) catalyzed by 4-OT(M45T/F50A) (1 mg/mL) in buffer [20 mM NaPi/5% (v/v) DMSO, 0.3 mL] at pH 7.3. **A.** Enzymatic reaction. **B.** Non-enzymatic reaction.



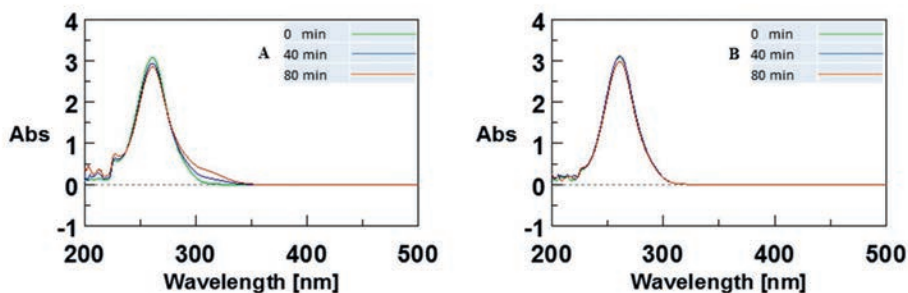
**Figure S18.** UV spectra monitoring the aldol reaction of **1** (50 mM) with **2s** (2 mM) catalyzed by 4-OT(M45T/F50A) (1 mg/mL) in buffer [20 mM NaPi/5% (v/v) DMSO, 0.3 mL] at pH 7.3. **A.** Enzymatic reaction. **B.** Non-enzymatic reaction.



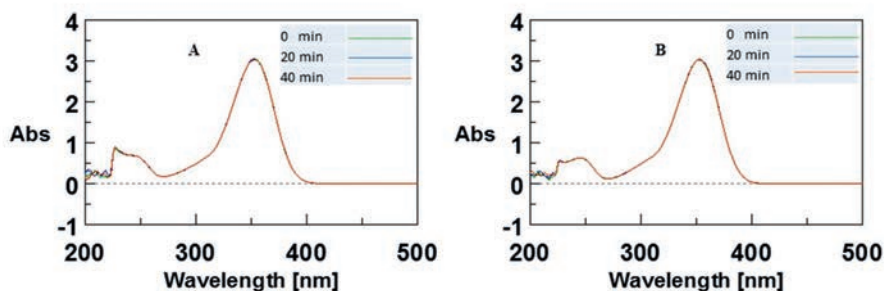
**Figure S19.** UV spectra monitoring the aldol reaction of **1** (50 mM) with **2t** (2 mM) catalyzed by 4-OT(M45T/F50A) (1 mg/mL) in buffer [20 mM NaPi/5% (v/v) DMSO, 0.3 mL] at pH 7.3. A. Enzymatic reaction. B. Non-enzymatic reaction.



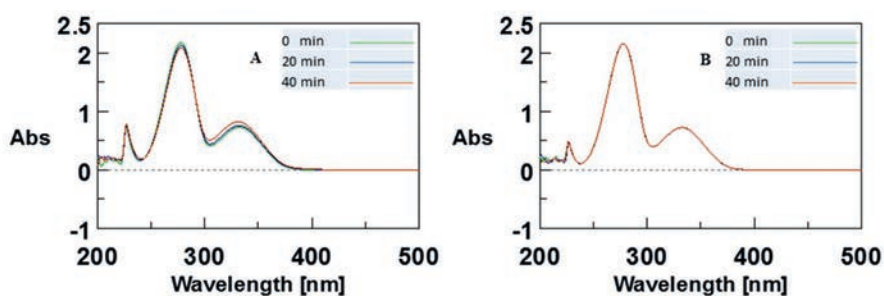
**Figure S20.** UV spectra monitoring the aldol reaction of **1** (50 mM) with **2u** (2 mM) catalyzed by 4-OT(M45T/F50A) (1 mg/mL) in buffer [20 mM NaPi/5% (v/v) EtOH, 0.3 mL] at pH 7.3. A. Enzymatic reaction. B. Non-enzymatic reaction.



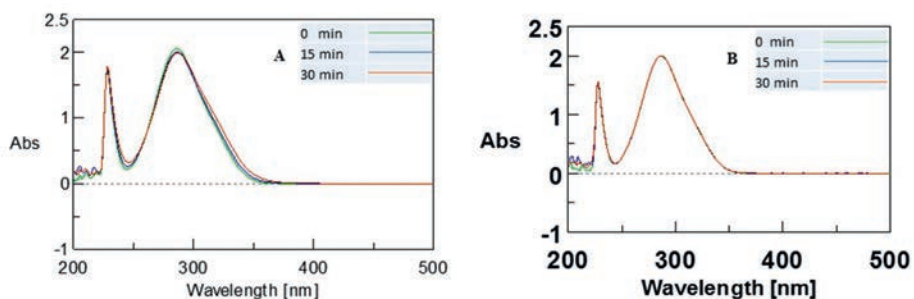
**Figure S21.** UV spectra monitoring the aldol reaction of **1** (50 mM) with **2v** (2 mM) catalyzed by 4-OT(M45T/F50A) (1 mg/mL) in buffer [20 mM NaPi/5% (v/v) DMSO, 0.3 mL] at pH 7.3. A. Enzymatic reaction. B. Non-enzymatic reaction.



**Figure S22.** UV spectra monitoring the aldol reaction of **1** (50 mM) with **2w** (2 mM) catalyzed by 4-OT(M45T/F50A) (1 mg/mL) in buffer [20 mM NaPi/5% (v/v) DMSO, 0.3 mL] at pH 7.3. **A.** Enzymatic reaction. **B.** Non-enzymatic reaction.

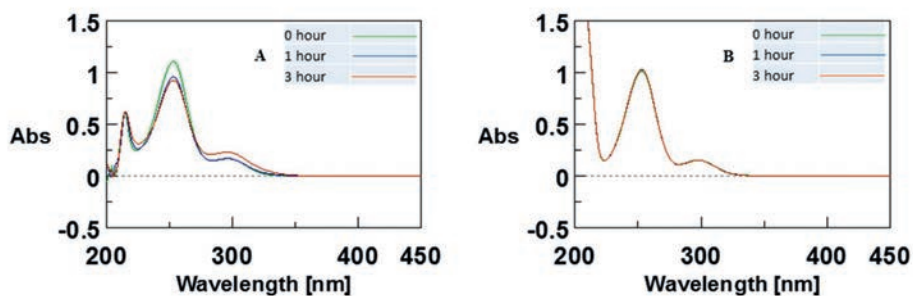


**Figure S23.** UV spectra monitoring the aldol reaction of **1** (50 mM) with **2x** (2 mM) catalyzed by 4-OT(M45T/F50A) (1 mg/mL) in buffer [20 mM NaPi/5% (v/v) DMSO, 0.3 mL] at pH 7.3. **A.** Enzymatic reaction. **B.** Non-enzymatic reaction.

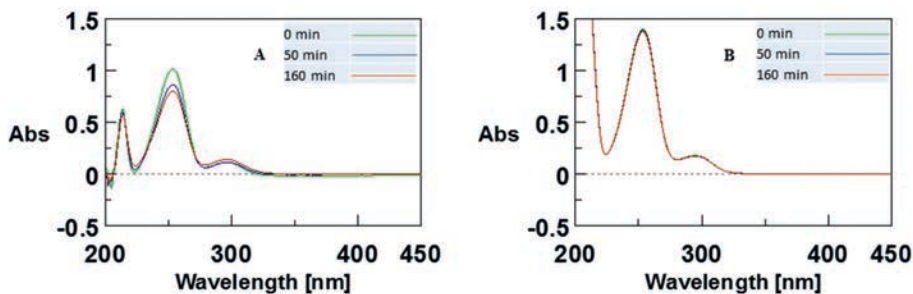


**Figure S24.** UV spectra monitoring the aldol reaction of **1** (50 mM) with **2y** (2 mM) catalyzed by 4-OT(M45T/F50A) (1 mg/mL) in buffer [20 mM NaPi/5% (v/v) DMSO, 0.3 mL] at pH 7.3. **A.** Enzymatic reaction. **B.** Non-enzymatic reaction.

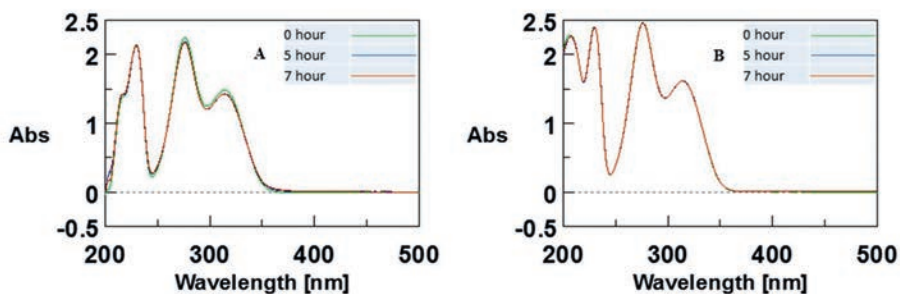




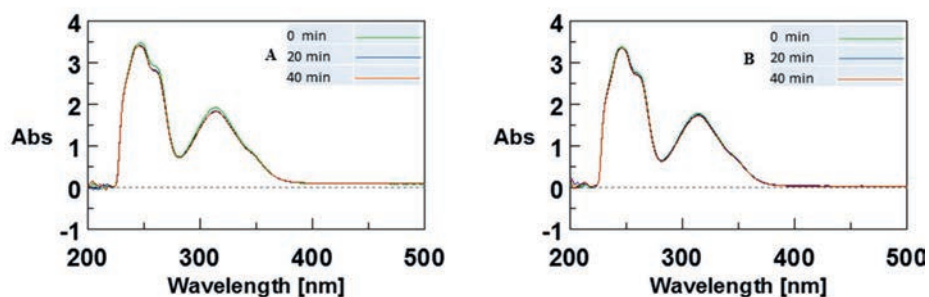
**Figure S25.** UV spectra monitoring the aldol reaction of **1** (50 mM) with **2z** (2 mM) catalyzed by TAUT015 (1 mg/mL) in buffer [20 mM NaPi/5% (v/v) EtOH, 0.3 mL] at pH 6.5. A. Enzymatic reaction. B. Non-enzymatic reaction.



**Figure S26.** UV spectra monitoring the aldol reaction of **1** (50 mM) with **2aa** (2 mM) catalyzed by TAUT015 (1 mg/mL) in buffer [20 mM NaPi/5% (v/v) EtOH, 0.3 mL] at pH 6.5. A. Enzymatic reaction. B. Non-enzymatic reaction.



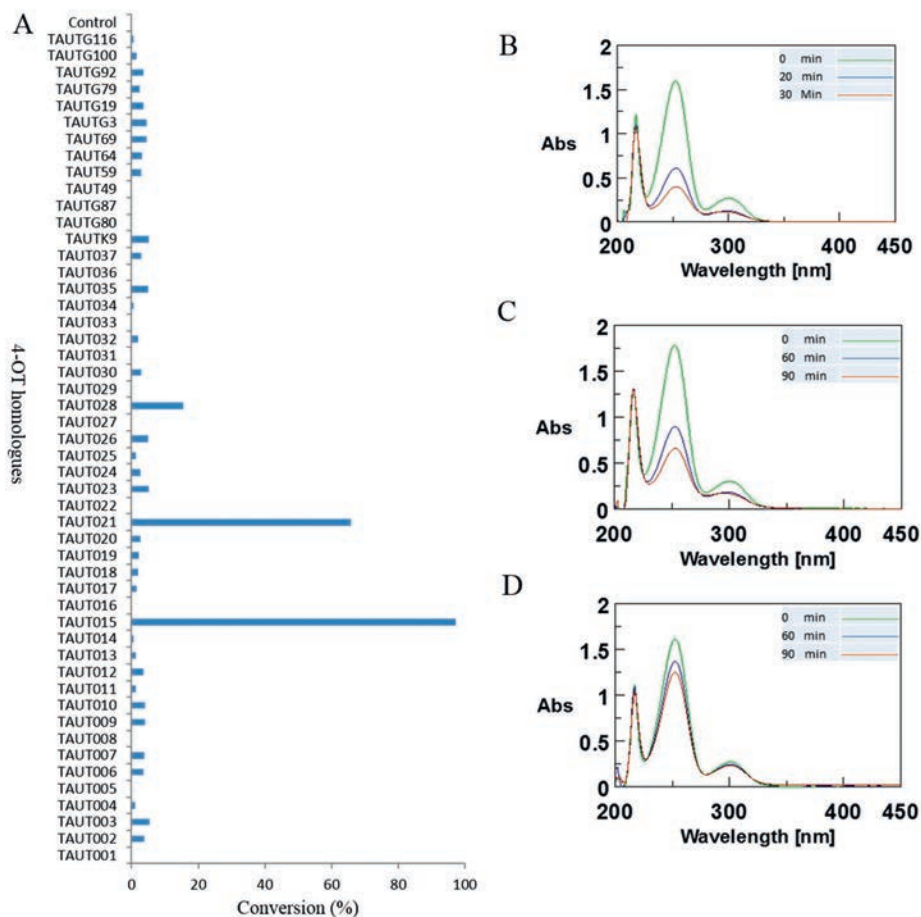
**Figure S27.** UV spectra monitoring the aldol reaction of **1** (50 mM) with **2ab** (2 mM) catalyzed by TAUT015 (1 mg/mL) in buffer [20 mM NaPi/5% (v/v) EtOH, 0.3 mL] at pH 6.5. A. Enzymatic reaction. B. Non-enzymatic reaction.



**Figure S28.** UV spectra monitoring the aldol reaction of **1** (50 mM) with **2ac** (2 mM) catalyzed by 4-OT(M45T/F50A) (1 mg/mL) in buffer [20 mM NaPi/5% (v/v) DMSO, 0.3 mL] at pH 7.3. **A.** Enzymatic reaction. **B.** Non-enzymatic reaction.

## 2. Screening of 4-OT homologues for aldol addition activity

Fifty 4-OT homologues (selected by Prozomix from in-house metagenomes) were provided as crude cell-free extracts (CFEs) by Prozomix Ltd (Station Court, Haltwhistle, Northumberland, UK), and used without any further purification. The aldol addition activity of the 4-OT homologues was determined using the donor substrate **1** and aldol acceptor substrate **2b**. Briefly, the reaction mixtures (300  $\mu$ L) consisted of CFE (1 mg/mL), **1** (50 mM), **2b** (2 mM) and ethanol (5% v/v) in 20 mM NaPi buffer (pH 6.5). The reaction was initiated by the addition of 15  $\mu$ L from a stock solution of **2b** (20 mM in ethanol). The progress of the reaction was monitored by following the depletion in absorbance at 254 nm, which corresponds to the consumption of **2b** (Figure S29).



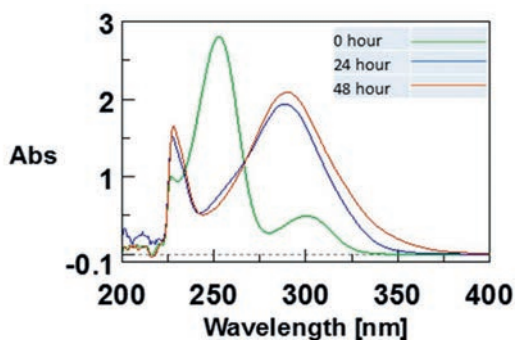
**Figure S29.** Screening of 4-OT homologues for the aldol addition of **1** (50 mM) to **2b** (2 mM). **A**. General screening results. The conversion (%) was calculated based on the amount of **2b** consumed. **B**, **C** and **D**. UV spectra monitoring the aldol addition of **1** (50 mM) to **2b** (2 mM) catalyzed by TAUT015 (**B**), TAUT021 (**C**) and TAUT028 (**D**).

### 3. Biocatalytic synthesis of $\alpha,\beta$ -unsaturated aldehydes (**4b-d**, **4f-j** and **4t**) using 4-OT(M45T/F50A)

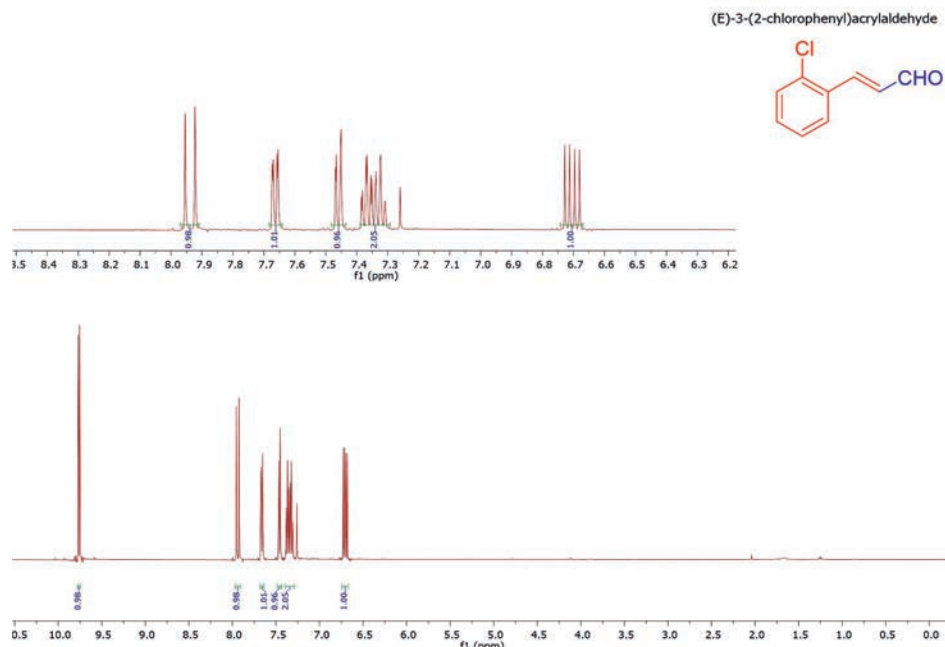
Reaction mixtures (60 mL) contained acetaldehyde (**1**, 100 mM), aromatic aldehydes (**2b-d**, **2f-j** and **2t**, 3 mM) and 4-OT(M45T/F50A) enzyme (0.1 mg/mL) in NaPi buffer (20 mM, pH 7.3). Stock solutions of **2b-d**, **2f-j** and **2t** were prepared in DMSO. The reaction mixture was incubated at room temperature and reaction progress was monitored by UV-vis spectrophotometry. After the completion of the reaction, the reaction mixture

was extracted with ethyl acetate ( $3 \times 40$  mL). The combined organic layers were washed with brine, dried over anhydrous  $\text{Na}_2\text{SO}_4$ , and concentrated under *vacuo*. The crude products **4b-d**, **4f-j** and **4t** were further purified by silica gel column chromatography (petroleum ether/ethyl acetate from 95:5 to 50:50).

**(E)-3-(2-chlorophenyl)acrylaldehyde (4b)**.<sup>[1a]</sup> White solid; yield = 81% (24 mg).  $^1\text{H NMR}$  (500 MHz,  $\text{CDCl}_3$ )  $\delta$  9.76 (d,  $J = 7.7$  Hz, 1H), 7.94 (d,  $J = 16.0$  Hz, 1H), 7.66 (dd,  $J = 7.7$ , 1.7 Hz, 1H), 7.46 (dd,  $J = 7.9$ , 1.3 Hz, 1H), 7.39 – 7.29 (m, 2H), 6.70 (dd,  $J = 16.0$ , 7.7 Hz, 1H). The  $^1\text{H NMR}$  data is in agreement with published data.<sup>[1a]</sup>

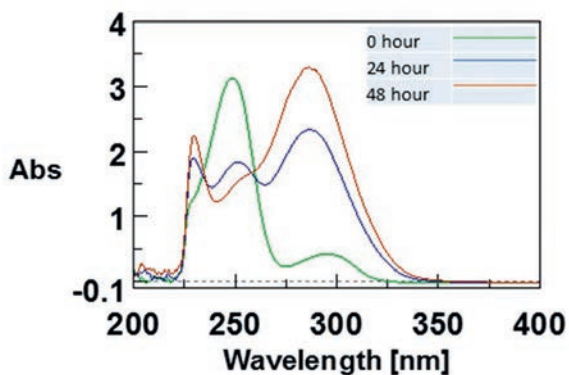


**Figure S30.** UV spectra monitoring the aldol condensation of **1** (100 mM) with **2b** (3 mM) catalyzed by 4-OT(M45T/F50A) (0.1 mg/mL) in buffer [20 mM NaPi/5% (v/v) DMSO] at pH 7.3.



**Figure S31.**  $^1\text{H}$  NMR spectrum of enzymatic compound **4b**.

(*E*)-3-(3-chlorophenyl)acrylaldehyde (**4c**).<sup>[1a]</sup> White solid; yield = 70% (21 mg).  $^1\text{H}$  NMR (500 MHz,  $\text{CDCl}_3$ )  $\delta$  9.72 (d,  $J$  = 7.6 Hz, 1H), 7.55 (t,  $J$  = 1.7 Hz, 1H), 7.48 – 7.42 (m, 2H), 7.42 – 7.36 (m, 2H), 6.71 (dd,  $J$  = 16.0, 7.6 Hz, 1H). The  $^1\text{H}$  NMR data is in agreement with published data.<sup>[1a]</sup>



**Figure S32.** UV spectra monitoring the aldol condensation of **1** (100 mM) with **2c** (3 mM) catalyzed by 4-OT(M45T/F50A) (0.1 mg/mL) in buffer [20 mM NaPi/5% (v/v) DMSO] at pH 7.3.

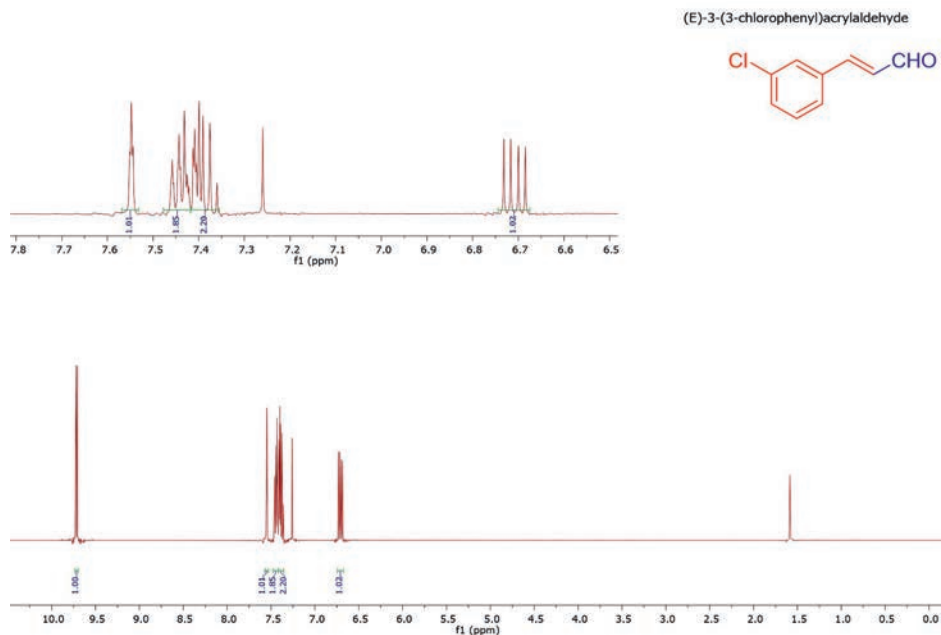


Figure S33.  $^1\text{H}$  NMR spectrum of enzymatic compound **4c**.

(*E*)-3-(2,4-dichlorophenyl)acrylaldehyde (**4d**).<sup>[1b]</sup> White solid; yield = 71% (25 mg).  $^1\text{H}$  NMR (500 MHz,  $\text{CDCl}_3$ )  $\delta$  9.76 (d,  $J = 7.6$  Hz, 1H), 7.86 (d,  $J = 16.0$  Hz, 1H), 7.60 (d,  $J = 8.5$  Hz, 1H), 7.49 (d,  $J = 2.1$  Hz, 1H), 7.32 (dd,  $J = 8.5, 1.8$  Hz, 1H), 6.69 (dd,  $J = 16.0, 7.6$  Hz, 1H). The  $^1\text{H}$  NMR data is in agreement with published data.<sup>[1b]</sup>

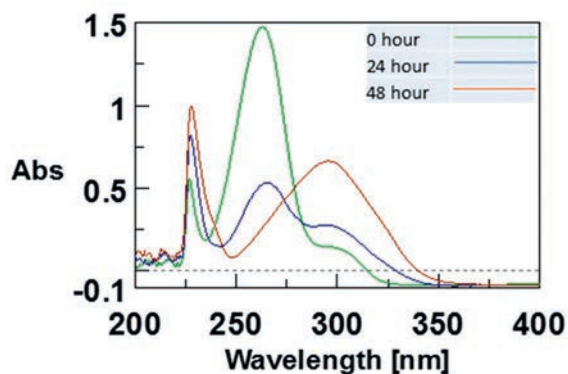
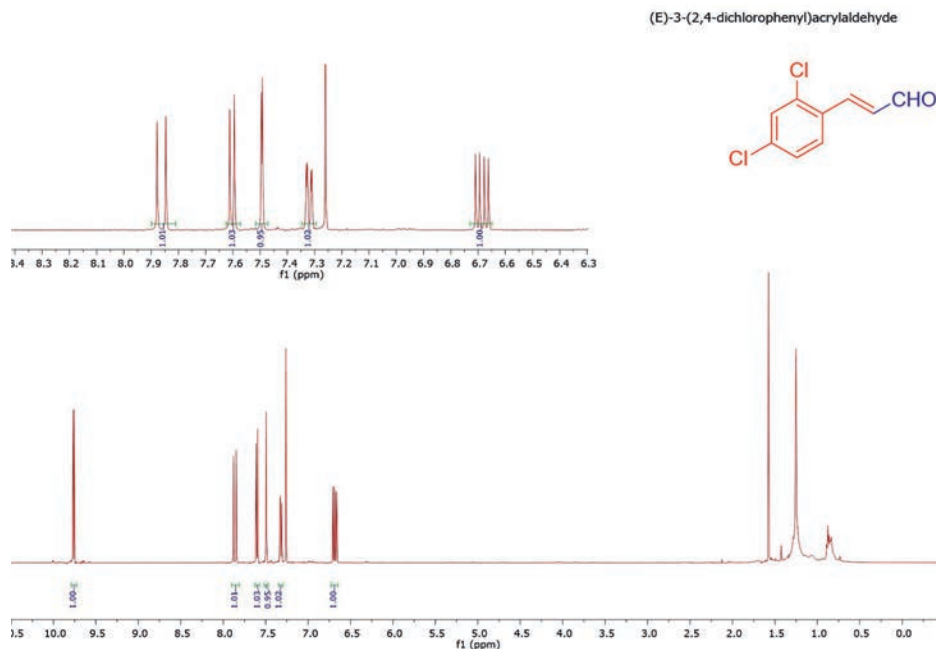
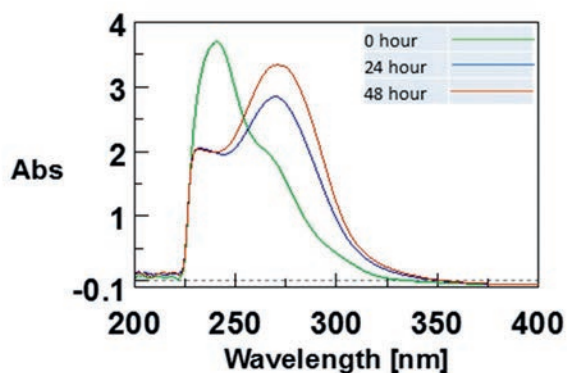


Figure S34. UV spectra monitoring the aldol condensation of **1** (100 mM) with **2d** (3 mM) catalyzed by 4-OT(M45T/F50A) (0.1 mg/mL) in buffer [20 mM NaPi/5% (v/v) DMSO] at pH 7.3.



**Figure S35.**  $^1\text{H}$  NMR spectrum of enzymatic compound **4d**.

(*E*)-3-(2-nitrophenyl)acrylaldehyde (**4f**).<sup>[1c]</sup> White solid; yield = 73% (23 mg).  $^1\text{H}$  NMR (500 MHz,  $\text{CDCl}_3$ )  $\delta$  9.78 (d,  $J = 7.4$  Hz, 1H), 8.42 (t,  $J = 1.9$  Hz, 1H), 8.30 (ddd,  $J = 8.2$ , 2.2, 0.9 Hz, 1H), 7.90 (d,  $J = 7.7$  Hz, 1H), 7.65 (t,  $J = 8.0$  Hz, 1H), 7.54 (d,  $J = 16.1$  Hz, 1H), 6.82 (dd,  $J = 16.1$ , 7.4 Hz, 1H). The  $^1\text{H}$  NMR data is in agreement with published data.<sup>[1c]</sup>



**Figure S36.** UV spectra monitoring the aldol condensation of **1** (100 mM) with **2f** (3 mM) catalyzed by 4-OT(M45T/F50A) (0.1 mg/mL) in buffer [20 mM NaPi/5% (v/v) DMSO] at pH 7.3.

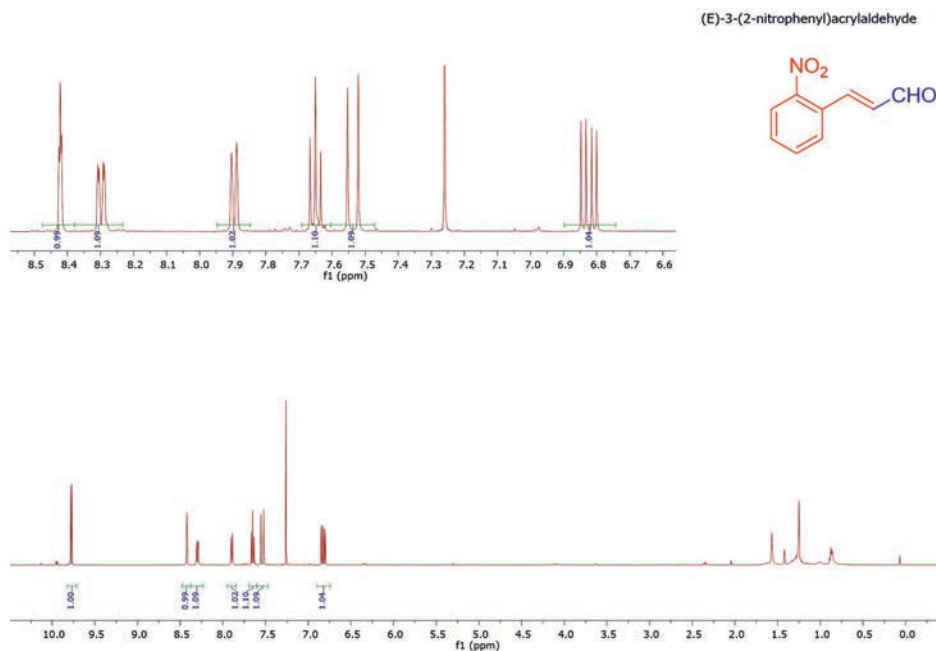


Figure S37.  $^1\text{H}$  NMR spectrum of enzymatic compound **4f**.

(*E*)-3-(3-nitrophenyl)acrylaldehyde (**4g**).<sup>[1d]</sup> White solid; yield = 61% (16 mg).  $^1\text{H}$  NMR (500 MHz,  $\text{CDCl}_3$ )  $\delta$  9.79 (d,  $J = 7.7$  Hz, 1H), 8.12 (d,  $J = 8.2$  Hz, 1H), 8.05 (d,  $J = 15.9$  Hz, 1H), 7.74 – 7.66 (m, 2H), 7.65 – 7.58 (m, 1H), 6.64 (dd,  $J = 15.8, 7.7$  Hz, 1H). The  $^1\text{H}$  NMR data is in agreement with published data.<sup>[1d]</sup>

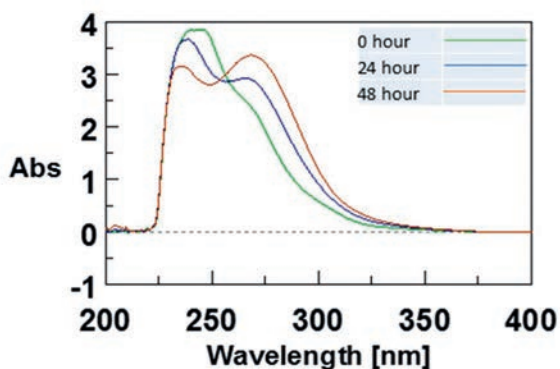
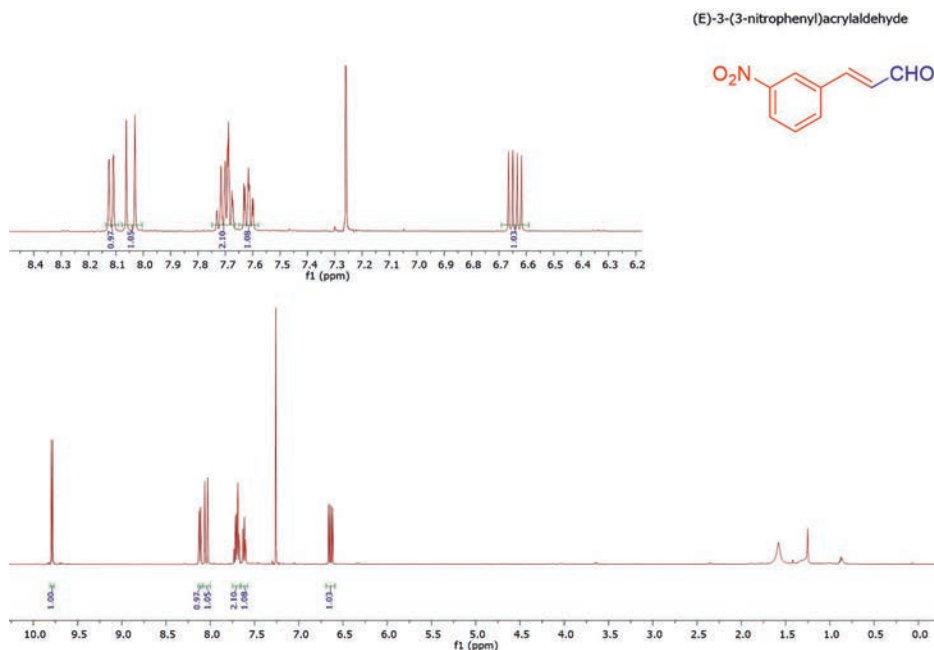


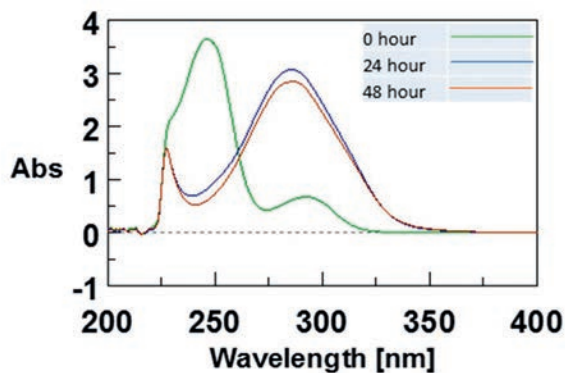
Figure S38. UV spectra monitoring the aldol condensation of **1** (100 mM) with **2g** (3 mM) catalyzed by 4-OT(M45T/F50A) (0.1 mg/mL) in buffer [20 mM NaPi/5% (v/v) DMSO] at pH 7.3.





**Figure S39.**  $^1\text{H}$  NMR spectrum of enzymatic compound **4g**.

(*E*)-3-(2-fluorophenyl)acrylaldehyde (**4h**).<sup>[1b]</sup> White solid; yield = 79% (21 mg).  $^1\text{H}$  NMR (500 MHz,  $\text{CDCl}_3$ )  $\delta$  9.72 (d,  $J = 7.7$  Hz, 1H), 7.66 (d,  $J = 16.2$  Hz, 1H), 7.59 (td,  $J = 7.6, 1.7$  Hz, 1H), 7.43 (dddd,  $J = 8.3, 7.2, 5.3, 1.7$  Hz, 1H), 7.25 – 7.19 (m, 1H), 7.15 (ddd,  $J = 10.5, 8.3, 1.0$  Hz, 1H), 6.79 (dd,  $J = 16.2, 7.7$  Hz, 1H). The  $^1\text{H}$  NMR data is in agreement with published data.<sup>[1b]</sup>



**Figure S40.** UV spectra monitoring the aldol condensation of **1** (100 mM) with **2h** (3 mM) catalyzed by 4-OT(M45T/F50A) (0.1 mg/mL) in buffer [20 mM NaPi/5% (v/v) DMSO] at pH 7.3.

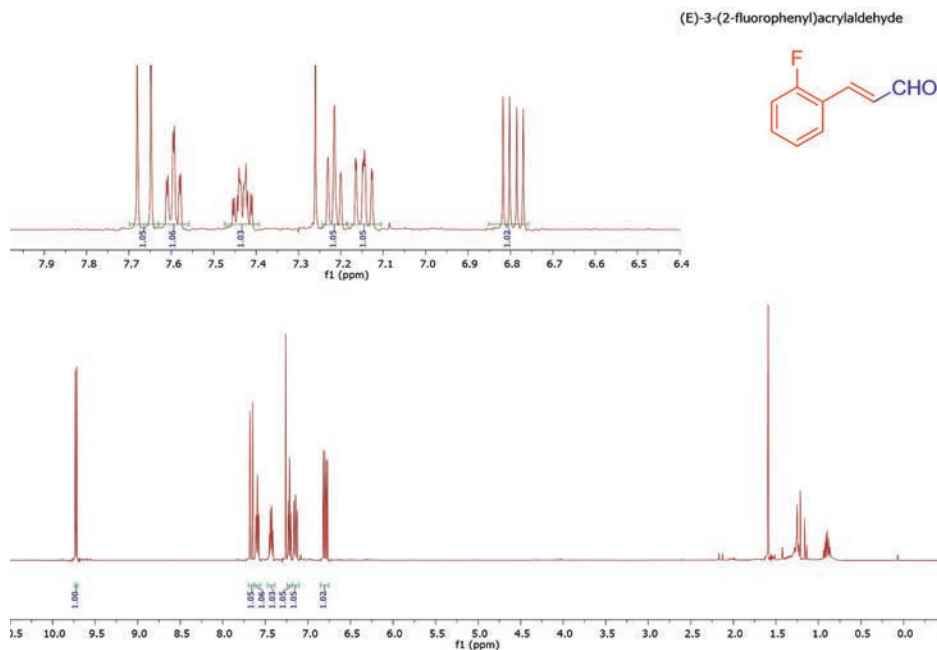


Figure S41.  $^1\text{H}$  NMR spectrum of enzymatic compound 4h.

(E)-3-(2-bromophenyl)acrylaldehyde (4i).<sup>[1c]</sup> White solid; yield = 85% (32 mg).  $^1\text{H}$  NMR (500 MHz,  $\text{CDCl}_3$ )  $\delta$  9.78 (d,  $J = 7.7$  Hz, 1H), 7.91 (d,  $J = 15.9$  Hz, 1H), 7.66 (dt,  $J = 7.7, 1.2$  Hz, 2H), 7.41 – 7.35 (m, 1H), 7.29 (td,  $J = 7.8, 1.6$  Hz, 1H), 6.68 (dd,  $J = 15.9, 7.7$  Hz, 1H). The  $^1\text{H}$  NMR data is in agreement with published data.<sup>[1c]</sup>

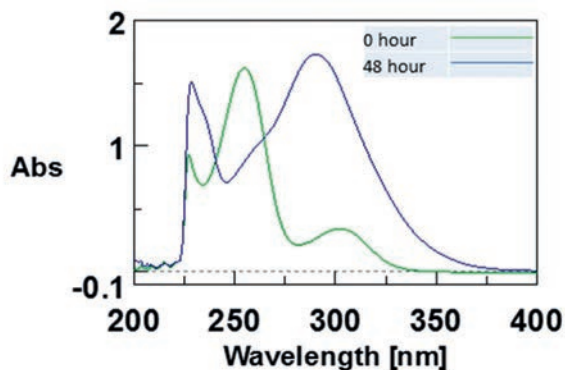
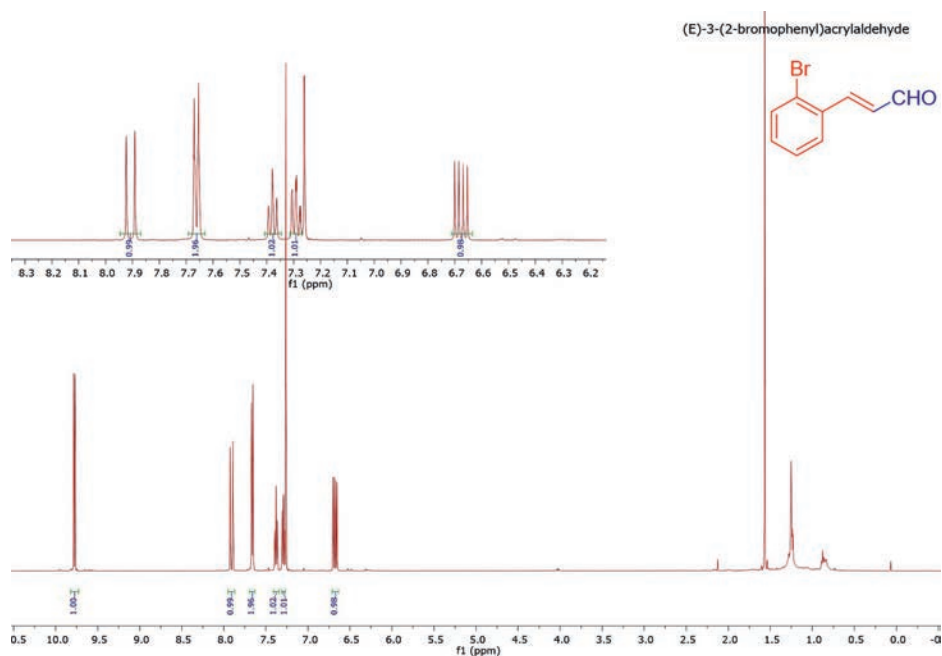
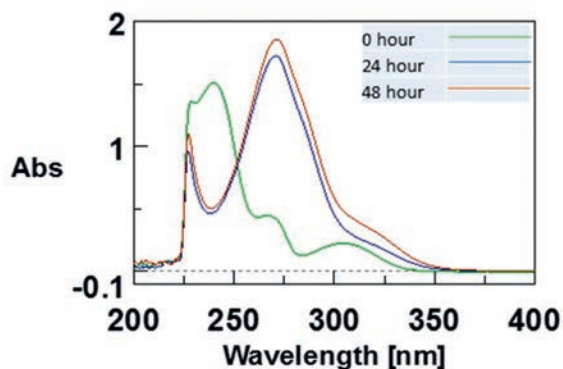


Figure S42. UV spectra monitoring the aldol condensation of 1 (100 mM) with 2i (3 mM) catalyzed by 4-OT(M45T/F50A) (0.1 mg/mL) in buffer [20 mM NaPi/5% (v/v) DMSO] at pH 7.3.



**Figure S43.**  $^1\text{H}$  NMR spectrum of enzymatic compound **4i**.

(*E*)-3-(2,3,5,6-tetrafluorophenyl)acrylaldehyde (**4j**). White solid; yield = 85% (31 mg).  $^1\text{H}$  NMR (500 MHz,  $\text{CDCl}_3$ )  $\delta$  9.75 (d,  $J = 7.5$  Hz, 1H), 7.52 (d,  $J = 16.5$  Hz, 1H), 7.23 – 7.11 (m, 1H), 7.02 (dd,  $J = 16.5, 7.4$  Hz, 1H).  $^{13}\text{C}$  NMR (126 MHz,  $\text{CDCl}_3$ )  $\delta$  193.61, 147.19, 146.00, 145.21, 143.97, 136.27, 135.72, 135.66, 135.59, 114.86, 108.21, 108.03, 107.85. HRMS (ESI+): calcd. for  $\text{C}_9\text{H}_5\text{F}_4\text{O}$   $[\text{M}+\text{H}]^+$ : 205.0277, found: 205.0268.



**Figure S44.** UV spectra monitoring the aldol condensation of **1** (100 mM) with **2j** (3 mM) catalyzed by 4-OT(M45T/F50A) (0.1 mg/mL) in buffer [20 mM NaPi/5% (v/v) DMSO] at pH 7.3.

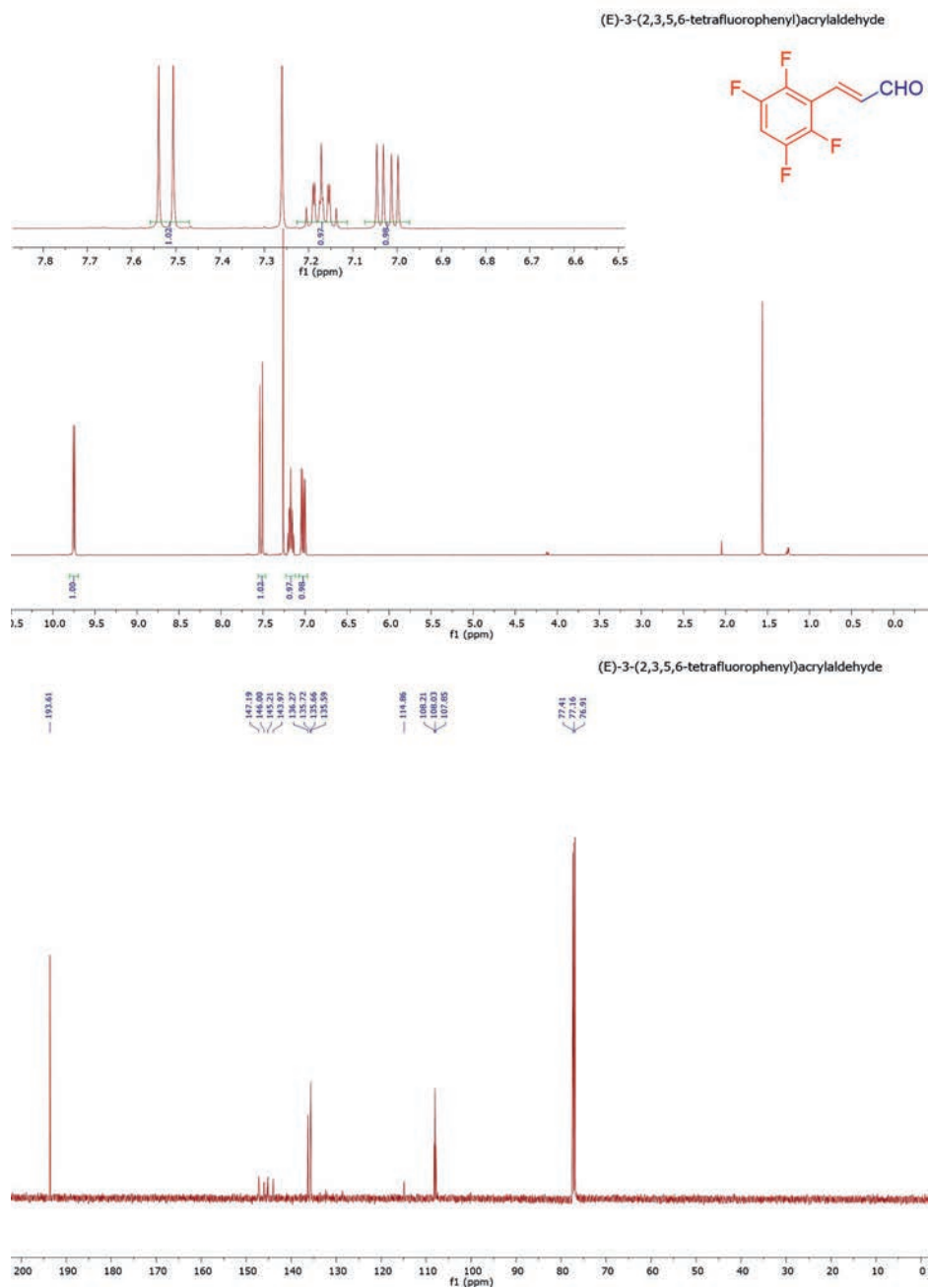
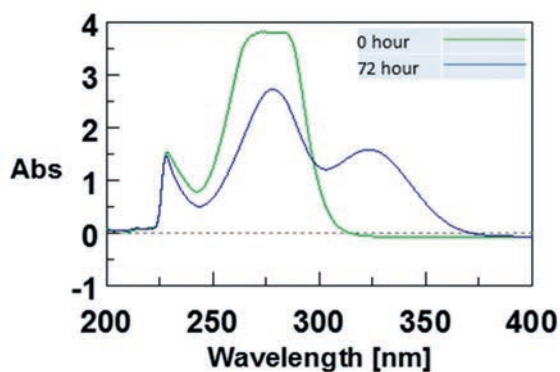
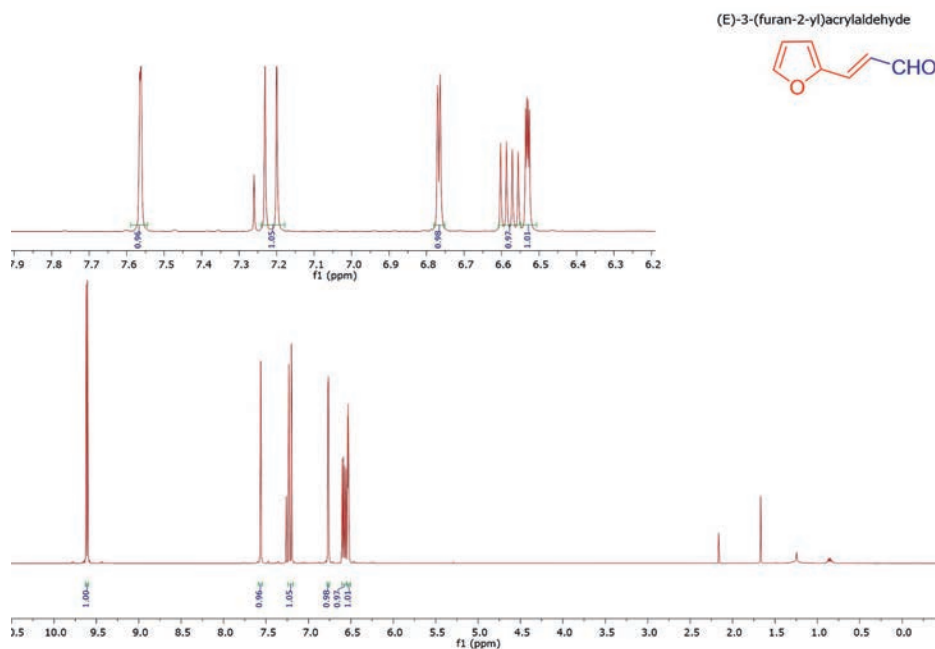


Figure S45. <sup>1</sup>H NMR (top) and <sup>13</sup>C NMR (bottom) spectra of enzymatic compound 4j.

(*E*)-3-(furan-2-yl)acrylaldehyde (**4t**).<sup>[1e]</sup> White solid; yield = 23% (5 mg). <sup>1</sup>H NMR (500 MHz, CDCl<sub>3</sub>) δ 9.61 (d, *J* = 7.9 Hz, 1H), 7.56 (d, *J* = 1.5 Hz, 1H), 7.22 (d, *J* = 15.7 Hz, 1H), 6.77 (d, *J* = 3.4 Hz, 1H), 6.58 (dd, *J* = 15.7, 7.9 Hz, 1H), 6.53 (dd, *J* = 3.4, 1.8 Hz, 1H). The <sup>1</sup>H NMR data is in agreement with published data.<sup>[1e]</sup>



**Figure S46.** UV spectra monitoring the aldol condensation of **1** (100 mM) with **2t** (3 mM) catalyzed by 4-OT(M45T/F50A) (0.1 mg/mL) in buffer [20 mM NaPi/5% (v/v) DMSO] at pH 7.3.

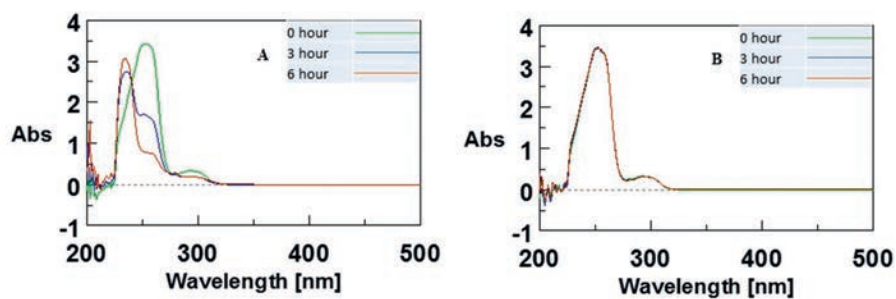


**Figure S47.** <sup>1</sup>H NMR spectrum of enzymatic compound **4t**.

#### 4. Chemoenzymatic synthesis of 1,3-diols (**5k-n** and **5s**) using 4-OT(M45T/F50A)

The reaction was performed in NaPi buffer (MOPS buffer for **2k**) [20 mM, pH 8.0 (pH 7.3 for **2s**), 40 mL]. Stock solutions of **2k-n** and **2s** (40 mM) were prepared in DMSO. From the stock solution of aldol acceptor substrate, 2 mL (2 mM final concentration) was added to buffer containing acetaldehyde (**1**, 100 mM) and an appropriate amount of purified enzyme 4-OT(M45T/F50A) (0.125-0.475 mg/mL). The reaction mixture was incubated at room temperature and the progress of the reaction was monitored by UV-vis spectrophotometry, following the consumption of **2k-n** and **2s**. When there was no further depletion of the substrates **2k-n** and **2s**, NaBH<sub>4</sub> was added into the reaction mixture (final concentration 30 mM). After 3 h of incubation at rt, the reaction mixture was extracted with ethyl acetate (3 × 40 mL). The combined organic layers were washed with brine, dried over anhydrous Na<sub>2</sub>SO<sub>4</sub>, and concentrated under *vacuo*. The crude products **5k-n** and **5s** were further purified by silica gel column chromatography (using petroleum ether/ethyl acetate from 95:5 to 50:50). The aromatic aldehydes **2k-n** and **2s** were used at a concentration of 2 mM to allow for easy monitoring of their consumption by UV-vis spectrophotometry. Notably, slightly higher reaction rates can be achieved when using these aldol acceptor substrates at a concentration of 10 mM.

**(R)-4-(1,3-dihydroxypropyl)benzotrile (5k)**.<sup>[4]</sup> Clear oil; yield = 92% (13 mg). <sup>1</sup>H NMR (500 MHz, CDCl<sub>3</sub>) δ 7.64 (d, *J* = 8.2 Hz, 2H), 7.49 (d, *J* = 8.1 Hz, 2H), 5.09 – 4.97 (m, 1H), 3.89 (q, *J* = 4.7 Hz, 2H), 3.51 (d, *J* = 3.0 Hz, 1H), 2.31 (s, 1H), 1.94 (q, *J* = 5.6 Hz, 2H). <sup>13</sup>C NMR (126 MHz, CDCl<sub>3</sub>) δ 149.86, 132.45 (2xC), 126.47 (2xC), 118.97, 111.25, 73.67, 61.45, 40.35. The NMR data are in agreement with published data.<sup>[4]</sup> The enantiomeric ratio of chemoenzymatic product **5k** was determined using reverse phase HPLC on a Chiralpak<sup>®</sup> ID column (150 mm × 4.6 mm, Daicel) (MeCN/water = 7:93, 20°C) at a flow rate of 1 mL/min. UV detection at 210 nm: *t*<sub>R</sub>:(major) = 11.6 min, (minor) = 16.8 min. The absolute configuration of **5k** was assigned by chiral HPLC using chemically prepared authentic standards with known configuration.



**Figure S48.** UV spectra monitoring the aldol addition of **1** (100 mM) to **2k** (2 mM) catalyzed by 4-OT(M45T/F50A) (0.475 mg/mL) in buffer [20 mM MOPS/5% (v/v) DMSO] at pH 8.0. **A.** Enzymatic reaction. **B.** Non-enzymatic reaction.

Enantioselective Aldol Addition of Acetaldehyde to Aromatic Aldehydes

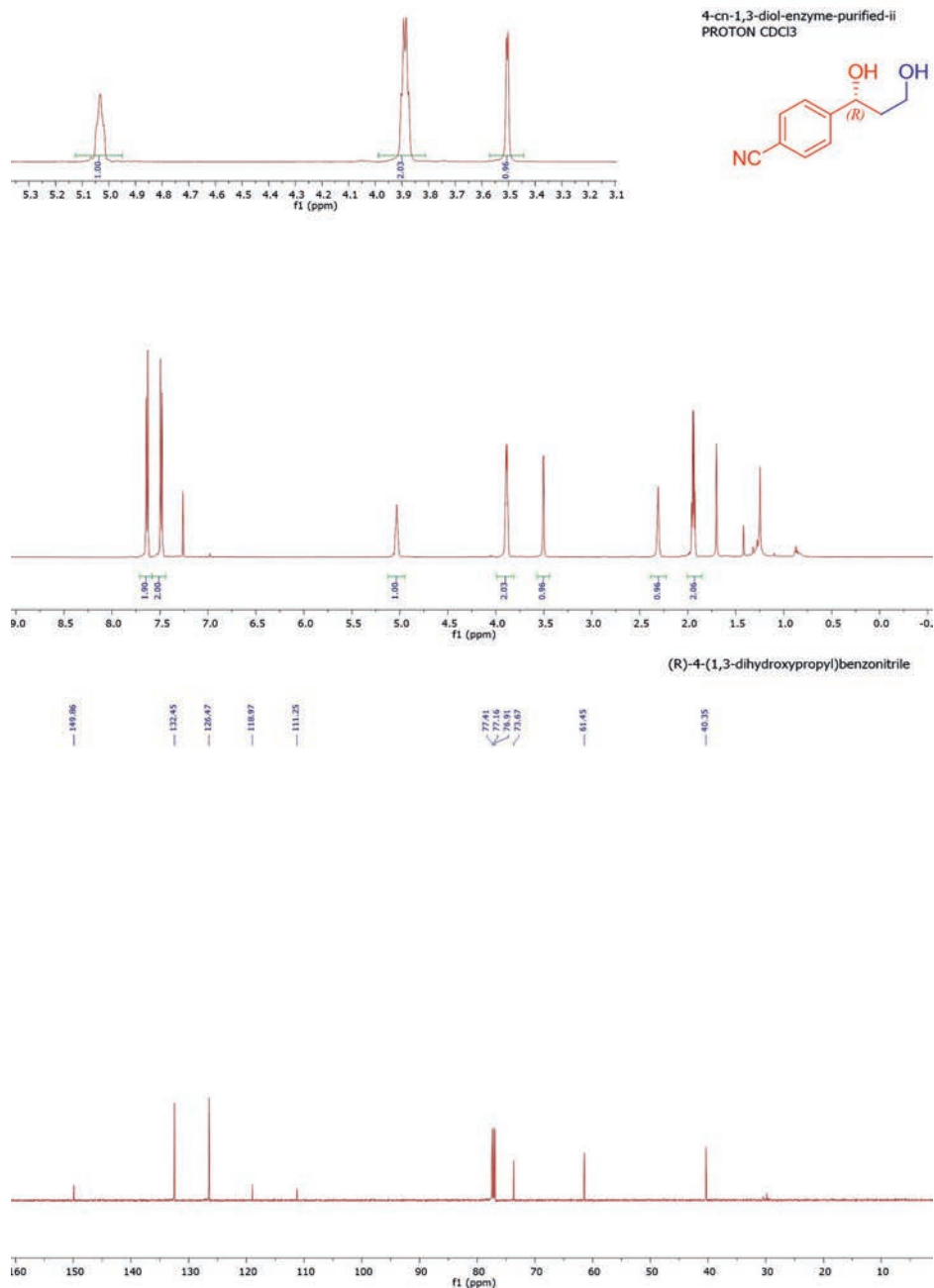


Figure S49. <sup>1</sup>H NMR (top) and <sup>13</sup>C NMR (bottom) spectra of chemoenzymatic product **5k**.



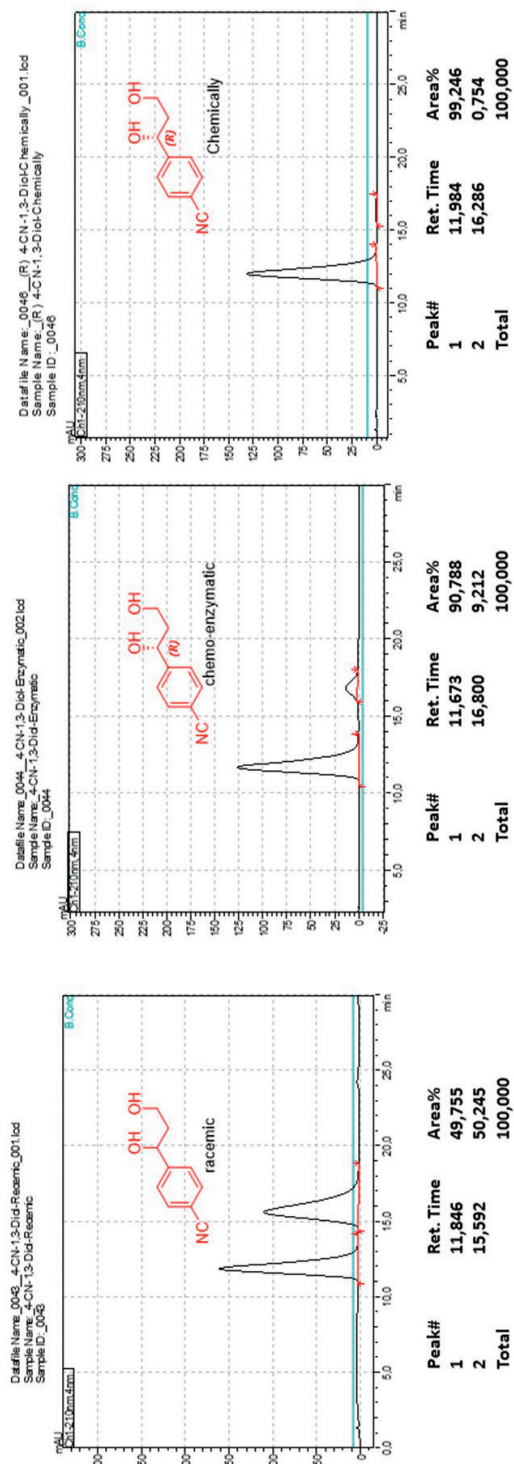
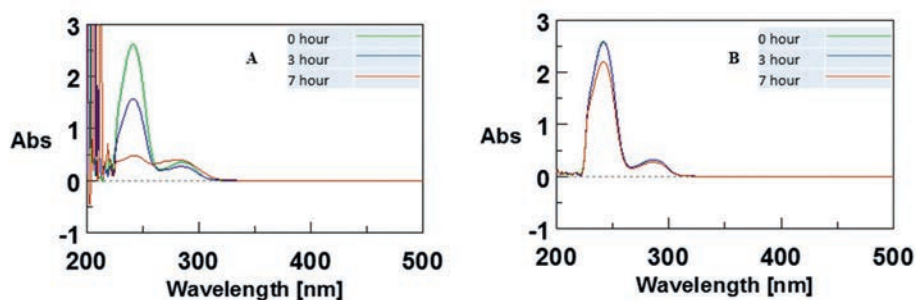


Figure S50. Chiral HPLC analysis of chemoenzymatic product 5k.

**(R)-1-(4-(trifluoromethyl)phenyl)propane-1,3-diol (51).**<sup>[2b]</sup> White solid; yield = 63% (11 mg). <sup>1</sup>H NMR (500 MHz, CDCl<sub>3</sub>) δ 7.61 (d, *J* = 8.2 Hz, 2H), 7.49 (d, *J* = 8.1 Hz, 2H), 5.10 – 5.00 (m, 1H), 3.90 (q, *J* = 4.9 Hz, 2H), 3.22 (d, *J* = 2.9 Hz, 1H), 2.22 (t, *J* = 4.1 Hz, 1H), 2.06 – 1.90 (m, 2H). <sup>13</sup>C NMR (126 MHz, CDCl<sub>3</sub>) δ 148.39, 129.95, 129.69, 126.05, 125.61, 125.58, 125.55, 125.52, 73.84, 61.52, 40.47. The NMR data are in agreement with published data.<sup>[2b]</sup> The enantiomeric ratio of chemoenzymatic product **51** was determined using normal phase HPLC on a Chiralpak<sup>†</sup> IC column (150 mm × 4.6 mm, Daicel) (Heptane/IPA = 95:5, 30°C) at a flow rate of 0.5 mL/min. UV detection at 210 nm: *t<sub>r</sub>*: (minor) = 13.09 min, (major) = 14.73 min. The absolute configuration of chemoenzymatically prepared **51** was assigned by chiral HPLC using chemically prepared authentic standards with known configuration.



**Figure S51.** UV spectra monitoring the aldol addition of **1** (100 mM) to **2I** (2 mM) catalyzed by 4-OT(M45T/F50A) (0.125 mg/mL) in buffer [20 mM NaPi/5% (v/v) DMSO] at pH 8.0. **A.** Enzymatic reaction. **B.** Non-enzymatic reaction.

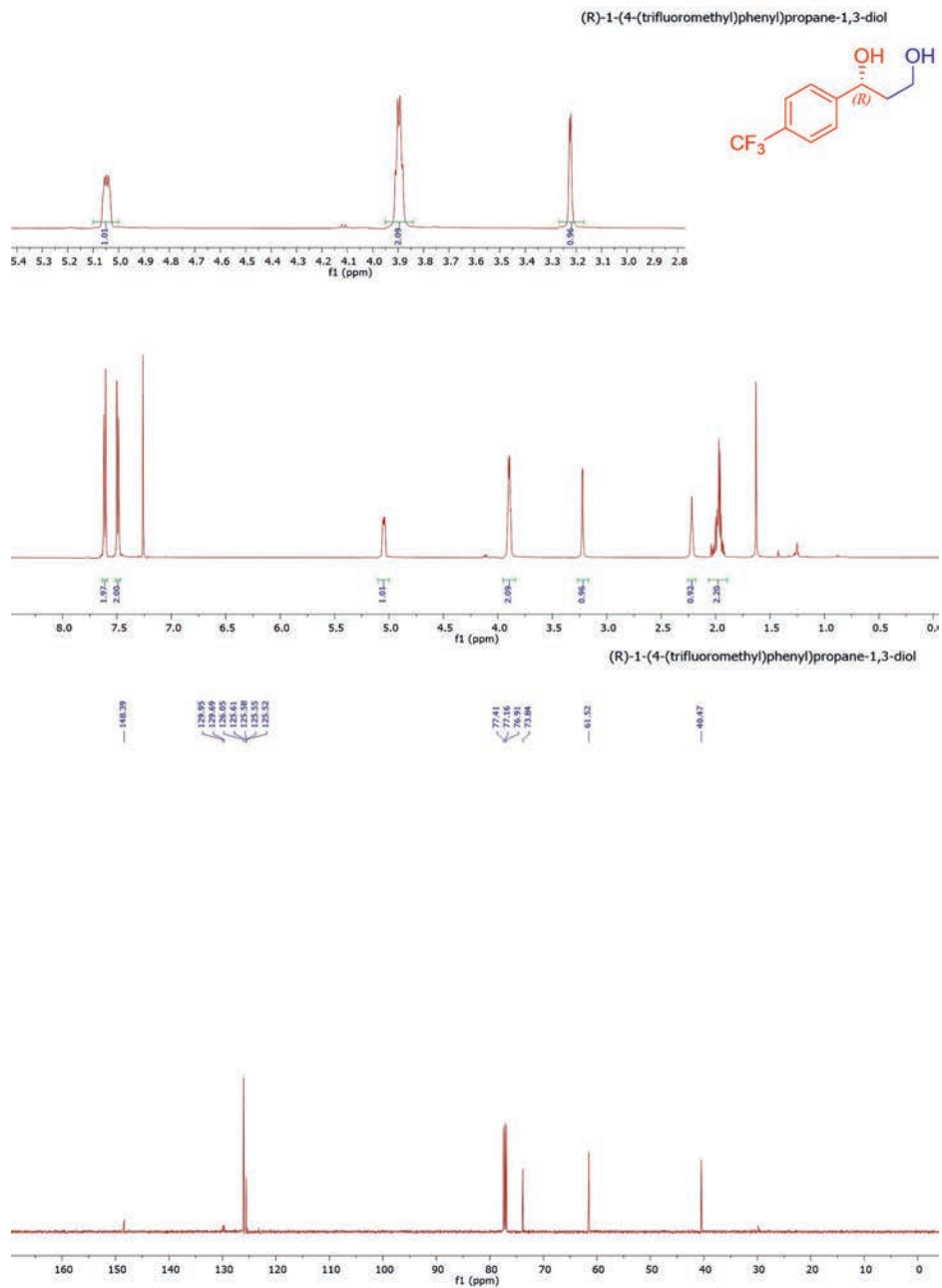


Figure S52.  $^1\text{H}$  NMR (top) and  $^{13}\text{C}$  NMR (bottom) spectra of chemoenzymatic product **5l**.

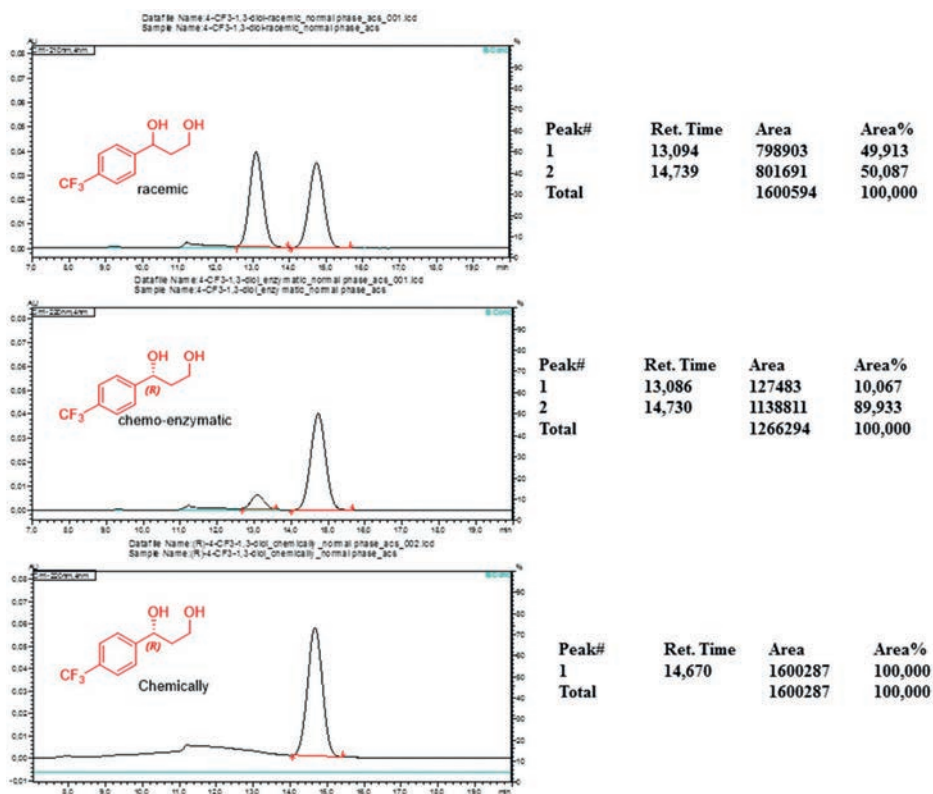
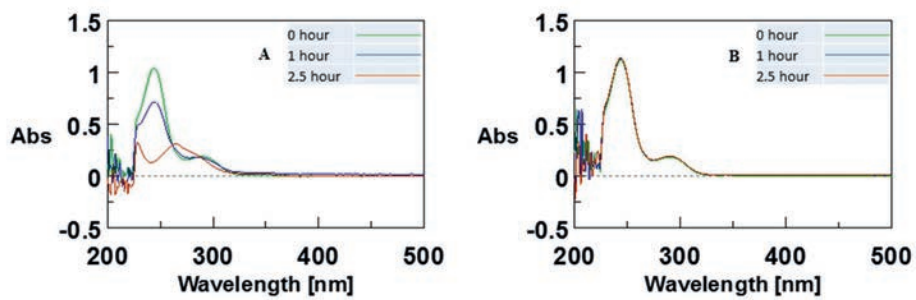


Figure S53. Chiral HPLC analysis of chemoenzymatic product **5l**.

**(R)-1-(perfluorophenyl)propane-1,3-diol (5m)**.<sup>[2b]</sup> White solid; yield = 88% (17 mg). <sup>1</sup>H NMR (500 MHz, CDCl<sub>3</sub>) δ 5.33 (d, *J* = 8.7 Hz, 1H), 3.98 – 3.81 (m, 2H), 3.37 (s, 1H), 2.38 – 2.30 (m, 1H), 2.27 (s, 1H), 2.01 – 1.89 (m, 1H). <sup>13</sup>C NMR (126 MHz, CDCl<sub>3</sub>) δ 145.93, 143.97, 141.77, 139.75, 138.72, 136.71, 116.84, 65.43, 61.03, 38.08. The NMR data are in agreement with published data.<sup>[2b]</sup> The enantiomeric ratio of chemoenzymatic product **5m** was determined using reverse phase HPLC on a Chiralpak<sup>®</sup> ID column (150 mm × 4.6 mm, Daicel) (MeCN/water = 7:93, 20°C) at a flow rate of 1 mL/min. UV detection at 210 nm: *t*<sub>R</sub>:(major) = 14.7 min, (minor) = 23.8 min. The absolute configuration of chemoenzymatically prepared **5m** was assigned by chiral HPLC using chemically prepared authentic standards with known configuration.



**Figure S54.** UV spectra monitoring the aldol addition of **1** (100 mM) to **2m** (2 mM) catalyzed by 4-OT(M45T/F50A) (0.25 mg/mL) in buffer [20 mM NaPi/5% (v/v) DMSO] at pH 8.0. **A.** Enzymatic reaction. **B.** Non-enzymatic reaction.

Enantioselective Aldol Addition of Acetaldehyde to Aromatic Aldehydes

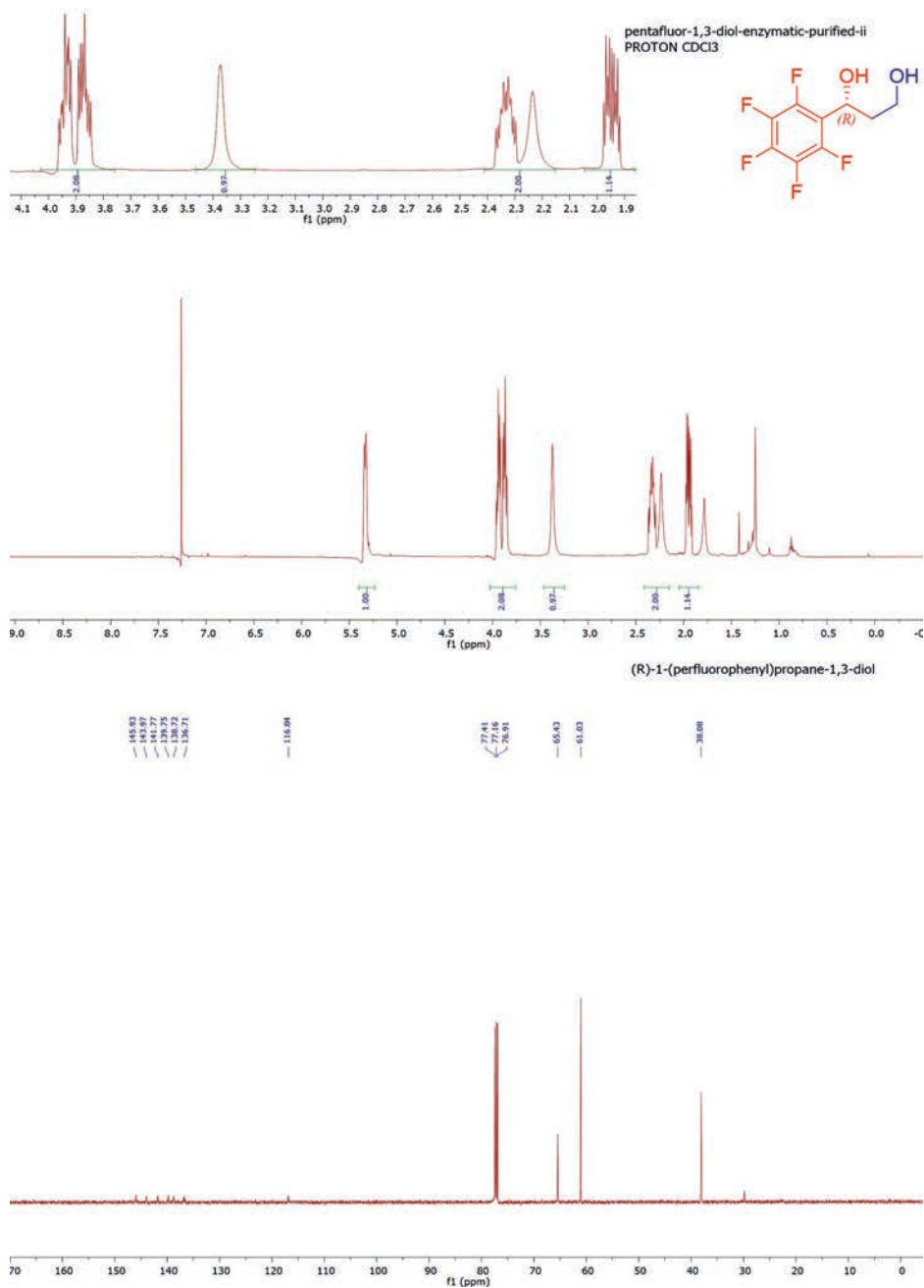


Figure S55. <sup>1</sup>H NMR (top) and <sup>13</sup>C NMR (bottom) spectra of chemoenzymatic product **5m**.

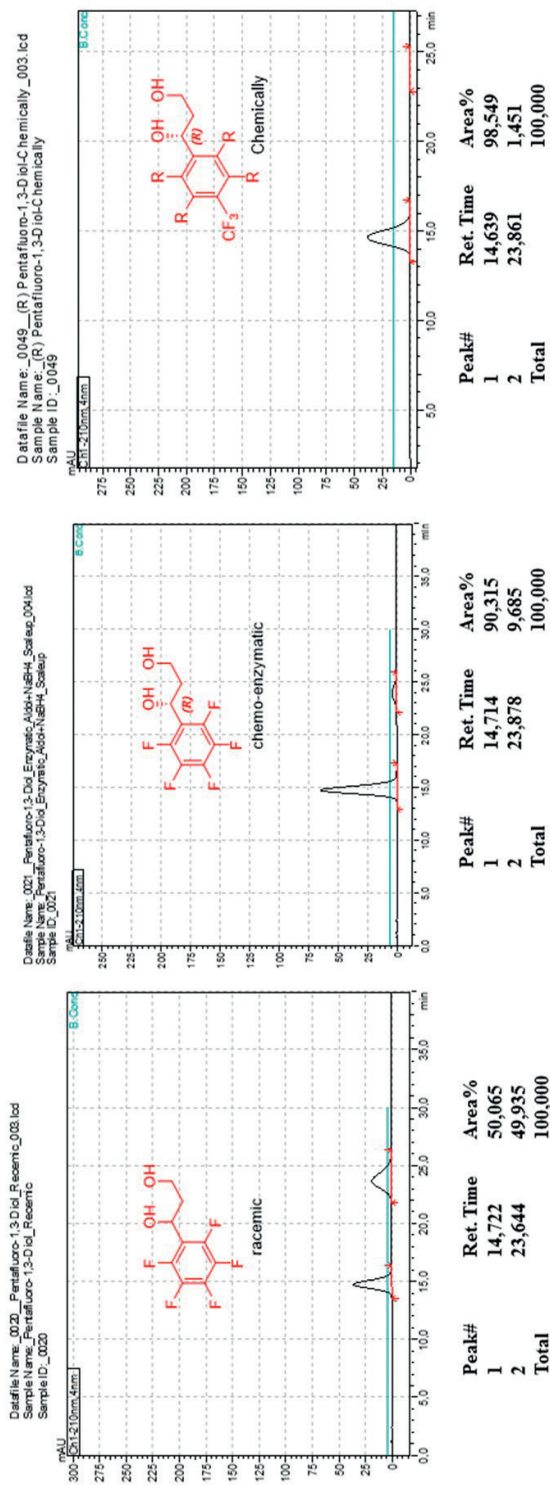
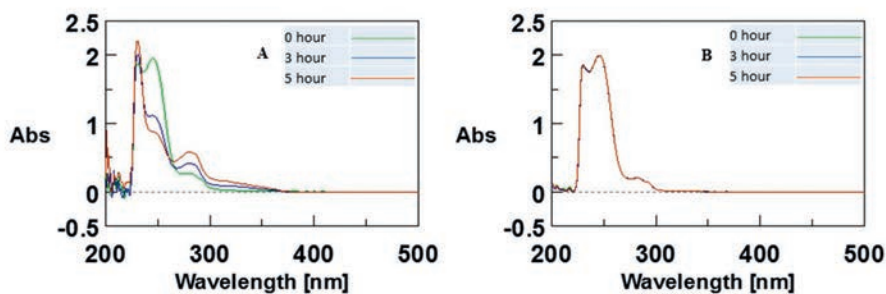


Figure S56. Chiral HPLC analysis of chemoenzymatic product 5m.

**(R)-5-(1,3-dihydroxypropyl)-2-fluorobenzonitrile (5n).** White solid; yield = 70% (11 mg).  $^1\text{H NMR}$  (500 MHz,  $\text{CDCl}_3$ )  $\delta$  7.67 (dd,  $J = 6.0, 2.2$  Hz, 1H), 7.62 (ddd,  $J = 7.5, 5.1, 2.3$  Hz, 1H), 7.20 (t,  $J = 8.7$  Hz, 1H), 5.06 – 4.98 (m, 1H), 3.98 – 3.88 (m, 2H), 3.47 (d,  $J = 2.6$  Hz, 1H), 2.04 (s, 1H), 2.02 – 1.88 (m, 2H).  $^{13}\text{C NMR}$  (126 MHz,  $\text{CDCl}_3$ )  $\delta$  163.47, 161.41, 141.72, 132.54, 132.47, 130.79, 116.65, 116.49, 114.14, 72.97, 61.55, 40.44. HRMS (ESI+): calcd. for  $\text{C}_{10}\text{H}_{11}\text{FNO}_2$   $[\text{M}+\text{H}]^+$ : 196.0774, found: 196.0766. The enantiomeric ratio of chemoenzymatic product **5n** was determined using reverse phase HPLC on a Chiralpak ID column (150 mm  $\times$  4.6 mm, Daicel) (MeCN/water = 7:93, 20°C) at a flow rate of 1 mL/min. UV detection at 210 nm:  $t_{\text{R}}$ :(major) = 13.7 min, (minor) = 27.1 min. The absolute configuration of chemoenzymatically prepared **5n** was tentatively assigned based on analogy.



**Figure S57.** UV spectra monitoring the aldol addition of **1** (100 mM) to **2n** (2 mM) catalyzed by 4-OT(M45T/F50A) (0.25 mg/mL) in buffer [20 mM NaPi/5% (v/v) DMSO] at pH 8.0. **A.** Enzymatic reaction. **B.** Non-enzymatic reaction.



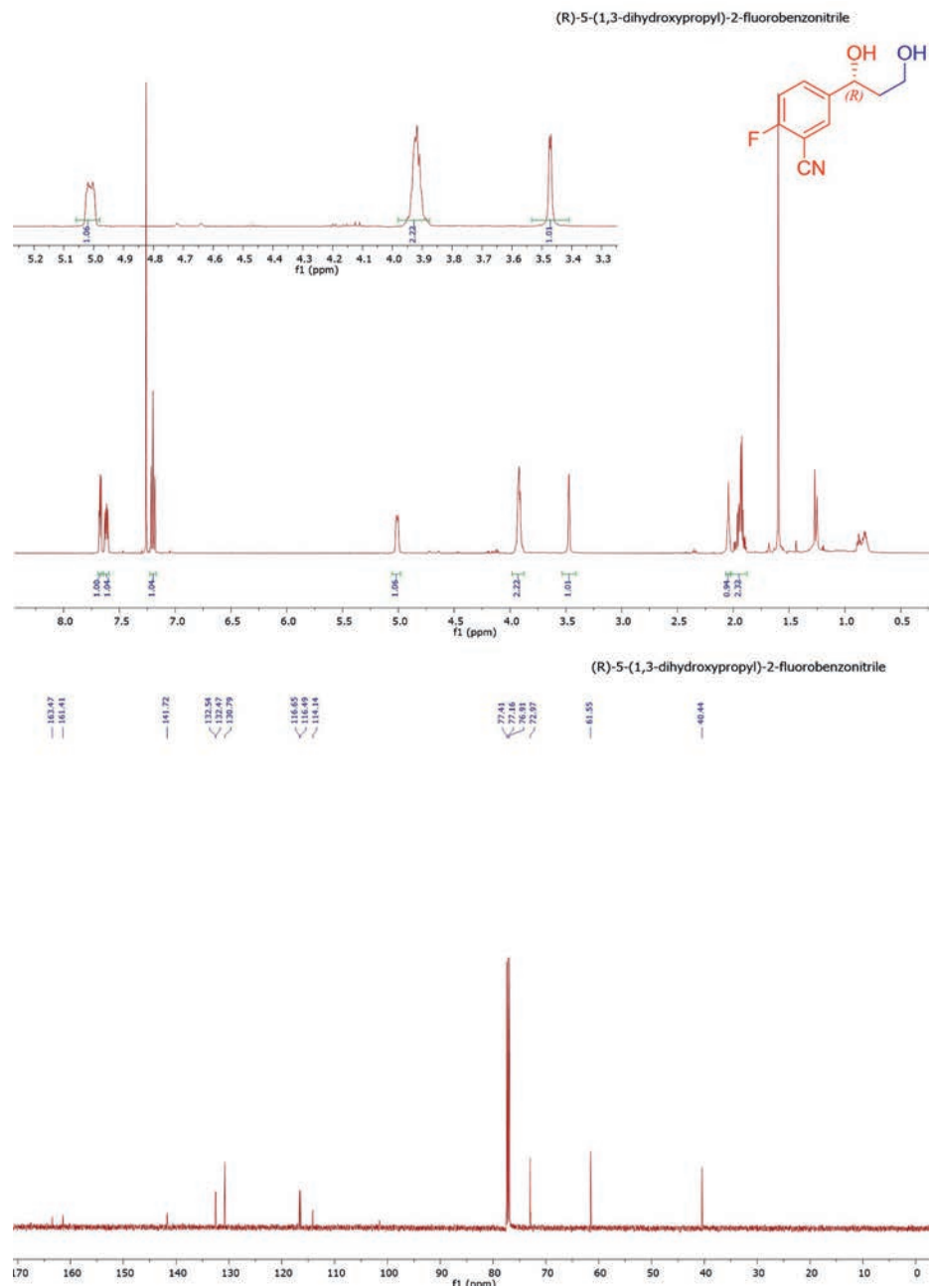


Figure S58. <sup>1</sup>H NMR (top) and <sup>13</sup>C NMR (bottom) spectra of chemoenzymatic product **5n**.

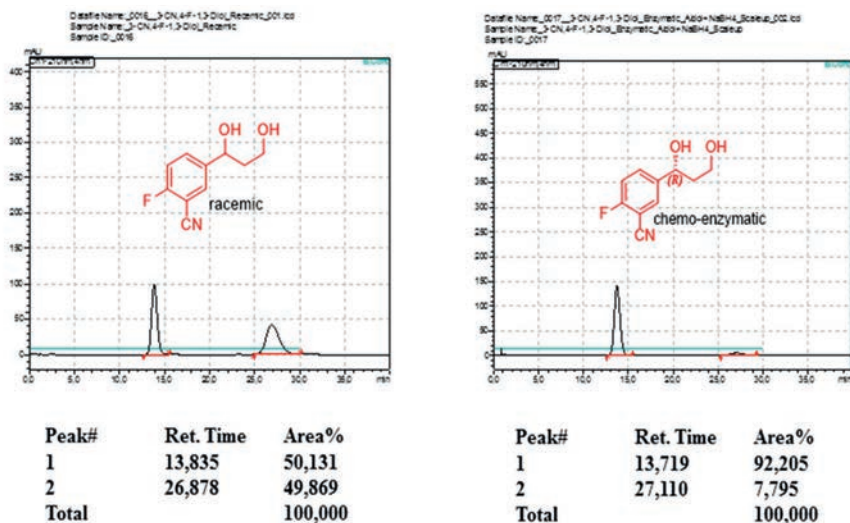


Figure S59. Chiral HPLC analysis of chemoenzymatic product **5n**.

**(R)-1-(pyridin-2-yl)propane-1,3-diol (5s)**.<sup>[5]</sup> Clear oil; yield = 65% (8 mg). <sup>1</sup>H NMR (500 MHz, CDCl<sub>3</sub>) δ 8.55 (d, *J* = 4.7 Hz, 1H), 7.72 (td, *J* = 7.6, 1.4 Hz, 1H), 7.32 (d, *J* = 7.9 Hz, 1H), 7.23 (dd, *J* = 7.4, 5.1 Hz, 1H), 5.00 (dd, *J* = 8.4, 3.2 Hz, 1H), 4.59 (s, 1H), 3.99 – 3.81 (m, 2H), 2.87 (s, 1H), 2.17 – 2.08 (m, 1H), 1.98 – 1.87 (m, 1H). <sup>13</sup>C NMR (126 MHz, CDCl<sub>3</sub>) δ 161.52, 148.30, 137.11, 122.60, 120.24, 72.78, 61.26, 39.99. The NMR data are in agreement with published data.<sup>[5]</sup> The enantiomeric ratio of chemoenzymatic product **5s** was determined using reverse phase HPLC on a Chiralpak<sup>®</sup> ID column (150 mm × 4.6 mm, Daicel) (MeCN/water = 7:93, 20°C) at a flow rate of 0.5 mL/min. UV detection at 210 nm: *t*<sub>R</sub>:(major) = 7.3 min, (minor) = 9.2 min. The absolute configuration of chemoenzymatically prepared **5s** was tentatively assigned based on analogy.

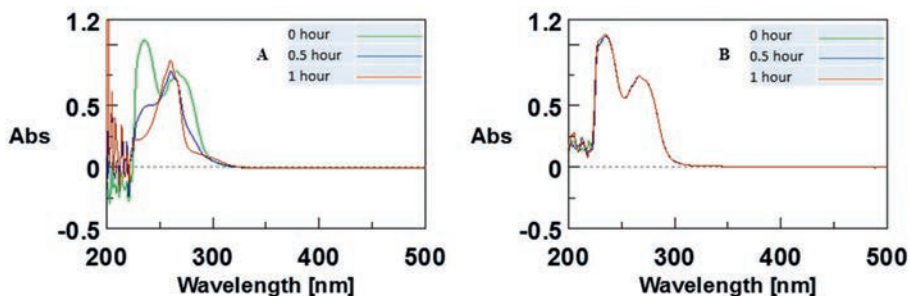


Figure S60. UV spectra monitoring the aldol addition of **1** (100 mM) to **2s** (2 mM) catalyzed by 4-OT(M45T/F50A) (0.125 mg/mL) in buffer [20 mM NaPi/5% (v/v) DMSO] at pH 7.3. **A**. Enzymatic reaction. **B**. Non-enzymatic reaction.

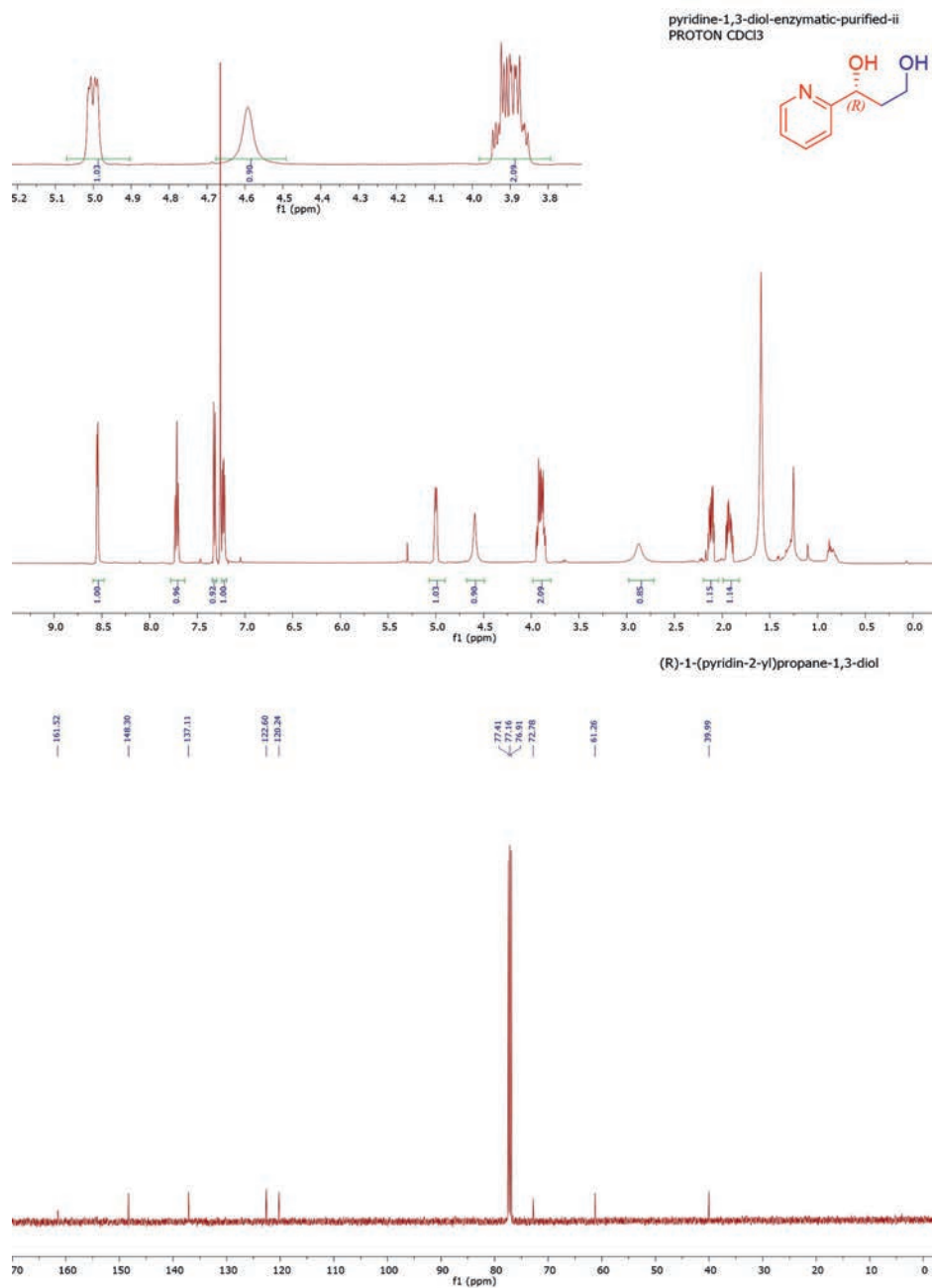


Figure S61. <sup>1</sup>H NMR (top) and <sup>13</sup>C NMR (bottom) spectra of chemoenzymatic product 5s.

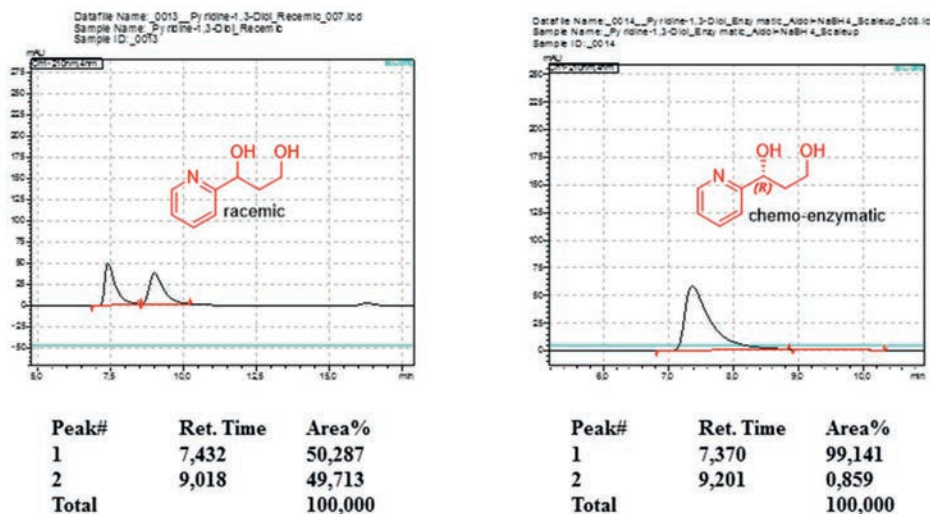


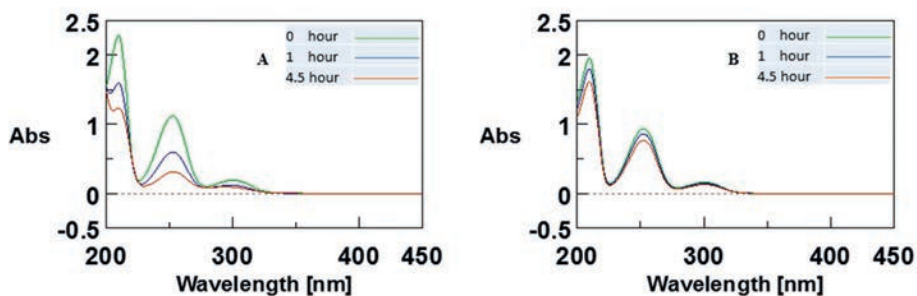
Figure S62. Chiral HPLC analysis of chemoenzymatic product **5s**.

## 5. Chemoenzymatic synthesis of 1,3-diols (**5b-d** and **5f-g**) using TAUT015

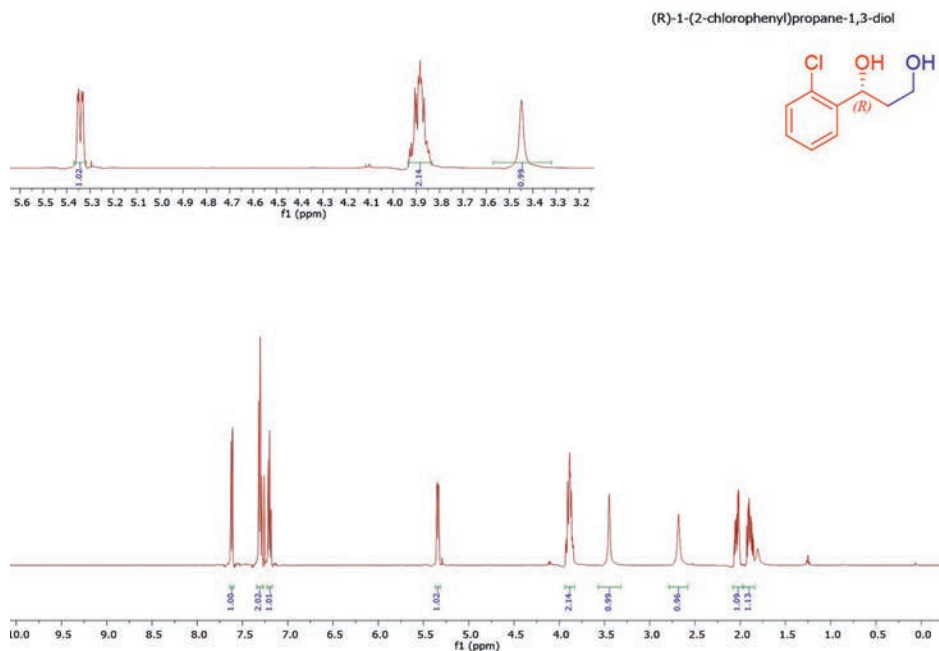
The reaction was performed in NaPi buffer (20 mM, pH 6.5, 40 mL). Stock solutions of **2b-d** and **2f-g** (40 mM) were prepared in absolute ethanol. From the stock solution, 2 mL (2 mM final concentration of the aldol acceptor substrate) was added to the NaPi buffer containing acetaldehyde (100 mM) and an appropriate amount of TAUT015 (crude extract, 0.33 mg/mL). The reaction mixture was incubated at room temperature and the progress of the reaction was monitored by UV-vis spectrophotometry, following the consumption of **2b-d** and **2f-g**. When there was no further depletion of the substrates **2b-d** and **2f-g**, NaBH<sub>4</sub> was added into the reaction mixture (final concentration 30 mM). After 3 h of incubation at rt, the reaction mixture was extracted with ethyl acetate (3 × 40 mL). The combined organic layers were washed with brine, dried over anhydrous Na<sub>2</sub>SO<sub>4</sub>, and concentrated under *vacuo*. The crude products **5b-d** and **5f-g** were further purified by silica gel column chromatography (using petroleum ether/ethyl acetate from 95:5 to 50:50).

(**R**)-1-(2-chlorophenyl)propane-1,3-diol (**5b**).<sup>[2b]</sup> yellowish oil; yield = 72% (11 mg). <sup>1</sup>H NMR (500 MHz, CDCl<sub>3</sub>) δ 7.62 (dd, *J* = 7.7, 1.4 Hz, 1H), 7.33 – 7.28 (m, 2H), 7.20 (td, *J* = 7.6, 1.7 Hz, 1H), 5.34 (dd, *J* = 8.8, 2.6 Hz, 1H), 3.93 – 3.83 (m, 2H), 3.45 (s, 1H), 2.68 (s, 1H), 2.09 – 1.99 (m, 1H), 1.94 – 1.84 (m, 1H). The <sup>1</sup>H NMR data is in agreement

with published data.<sup>[2b]</sup> The enantiomeric ratio (e.r.) of chemoenzymatic product **5b** was determined by reverse phase HPLC using a Chiralpak<sup>®</sup> ID column (150 mm × 4.6 mm, Daicel) (MeCN/water = 5:95, 25°C) at a flow rate of 1 mL/min. UV detection at 220 nm:  $t_R$ :(major) = 11.1 min, (minor) = 12.5 min.



**Figure S63.** UV spectra monitoring the aldol addition of **1** (100 mM) to **2b** (2 mM) catalyzed by TAUT015 (0.33 mg/mL) in NaPi buffer [20 mM/5% (v/v) EtOH] at pH 6.5. **A.** Enzymatic reaction. **B.** Non-enzymatic reaction.



**Figure S64.** <sup>1</sup>H NMR spectrum of chemoenzymatic product **5b**.

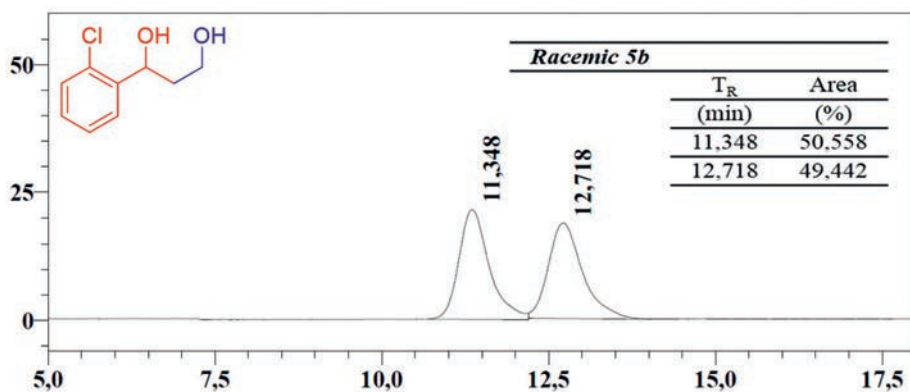


Figure S65. HPLC analysis of racemic 5b.

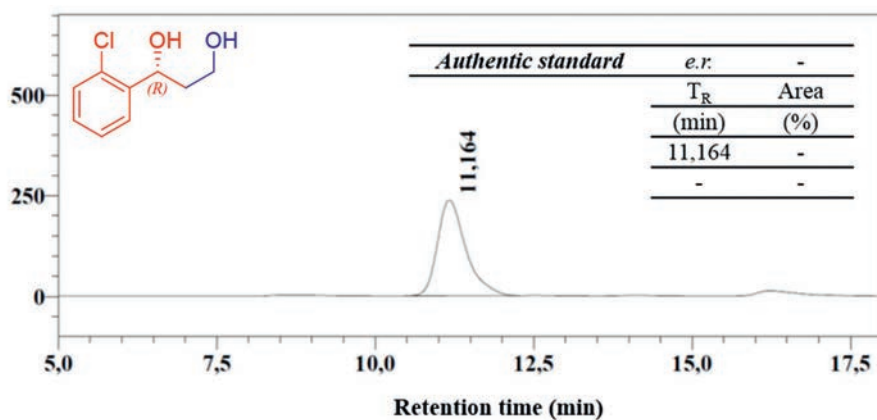


Figure S66. HPLC analysis of authentic standard (*R*)-5b

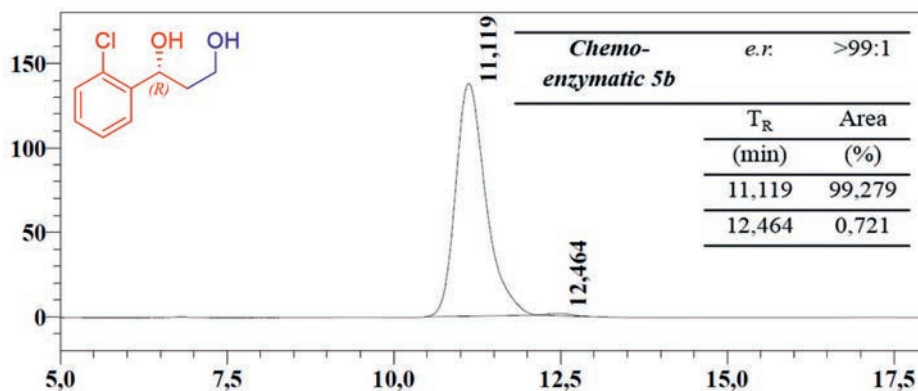
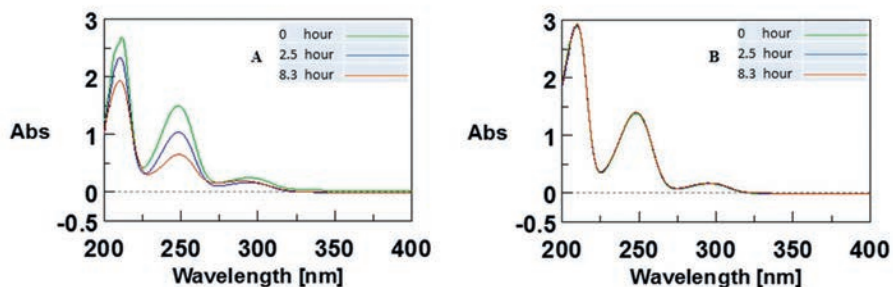
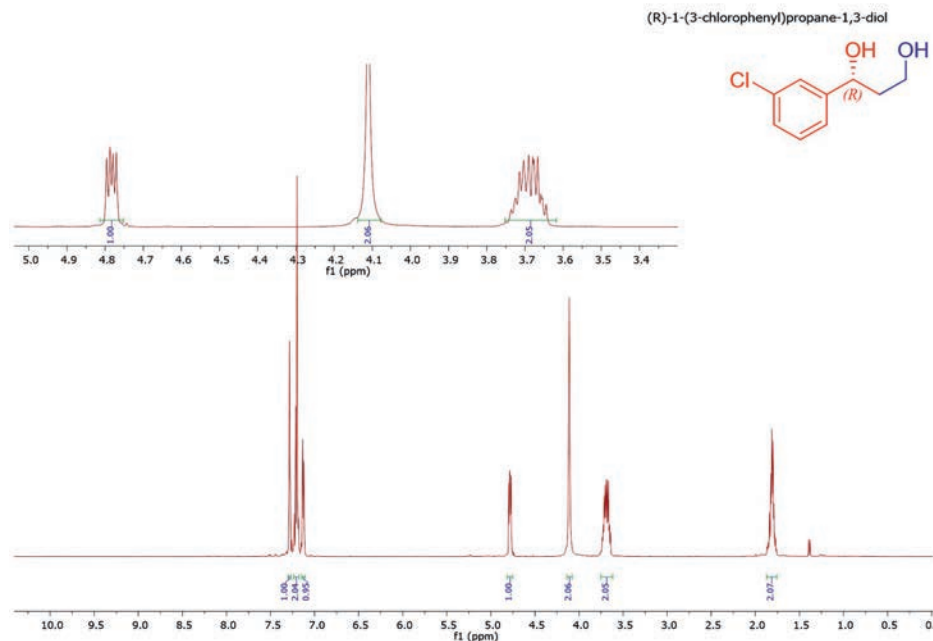


Figure S67. HPLC analysis of chemoenzymatic product 5b.

(*R*)-1-(3-chlorophenyl)propane-1,3-diol (**5c**).<sup>[6]</sup> yellowish oil; yield = 51% (8 mg). <sup>1</sup>H NMR (500 MHz, CDCl<sub>3</sub>) δ 7.29 (t, *J* = 2.1 Hz, 1H), 7.24 – 7.18 (m, 2H), 7.14 – 7.11 (m, 1H), 4.78 (dd, *J* = 8.3, 4.4 Hz, 1H), 4.11 (s, 2H), 3.75 – 3.62 (m, 2H), 1.87 – 1.76 (m, 2H). The <sup>1</sup>H NMR data is in agreement with published data.<sup>[6]</sup> The enantiomeric ratio of chemoenzymatic product **5c** was determined by reverse phase HPLC using a Chiralpak<sup>®</sup> ID column (150 mm × 4.6 mm, Daicel) (MeCN/water = 10:90, 25°C) at a flow rate of 1 mL/min. UV detection at 220 nm: *t*<sub>R</sub>:(major) = 9.7 min, (minor) = 14.1 min.



**Figure S68.** UV spectra monitoring the aldol addition of **1** (100 mM) to **2c** (2 mM) catalyzed by TAUT015 (0.33 mg/mL) in NaPi buffer [20 mM/5% (v/v) EtOH] at pH 6.5. A. Enzymatic reaction. B. Non-enzymatic reaction.



**Figure S69.** <sup>1</sup>H NMR spectrum of chemoenzymatic product **5c**.

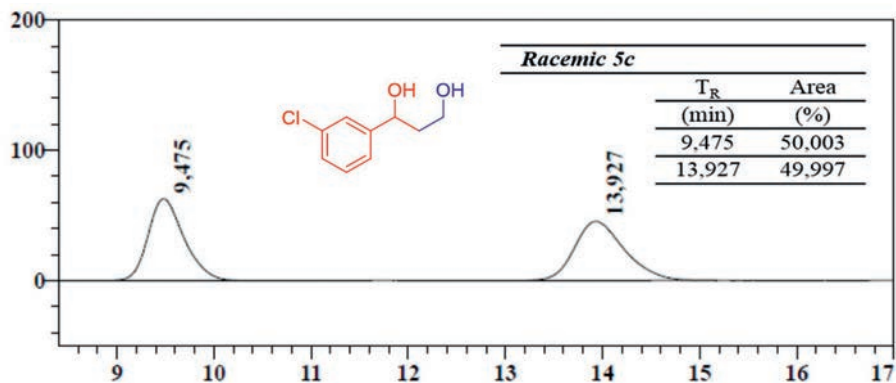


Figure S70. HPLC analysis of racemic 5c.

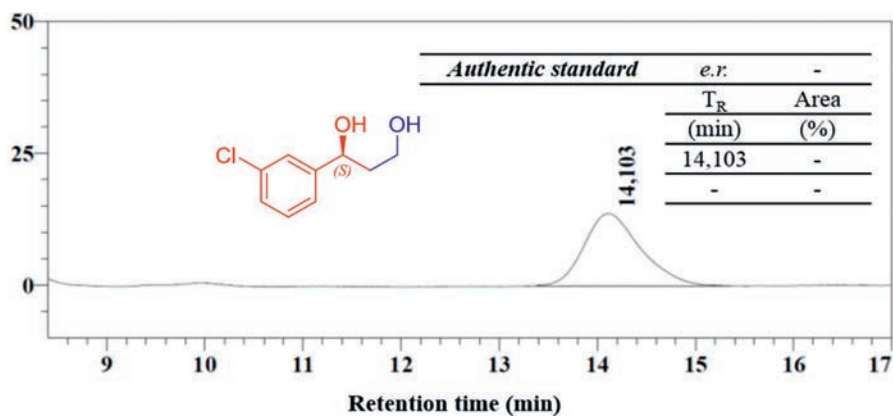


Figure S71. HPLC analysis of authentic standard (S)-5c.

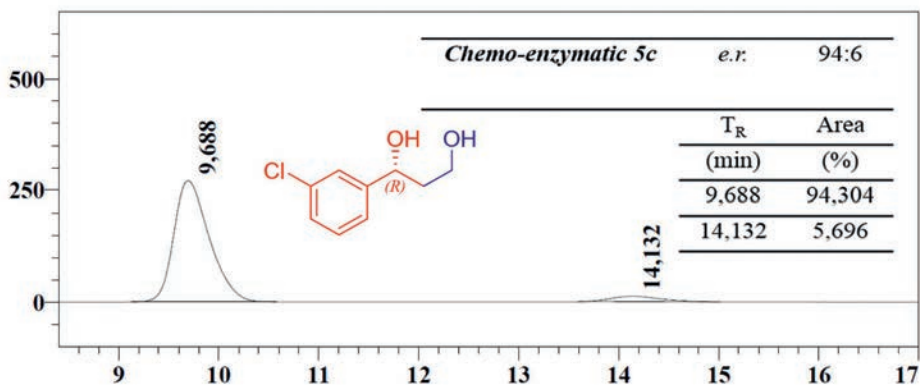
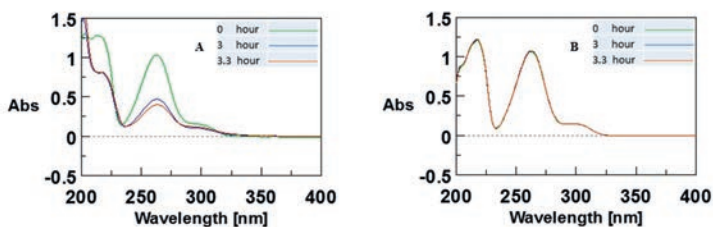


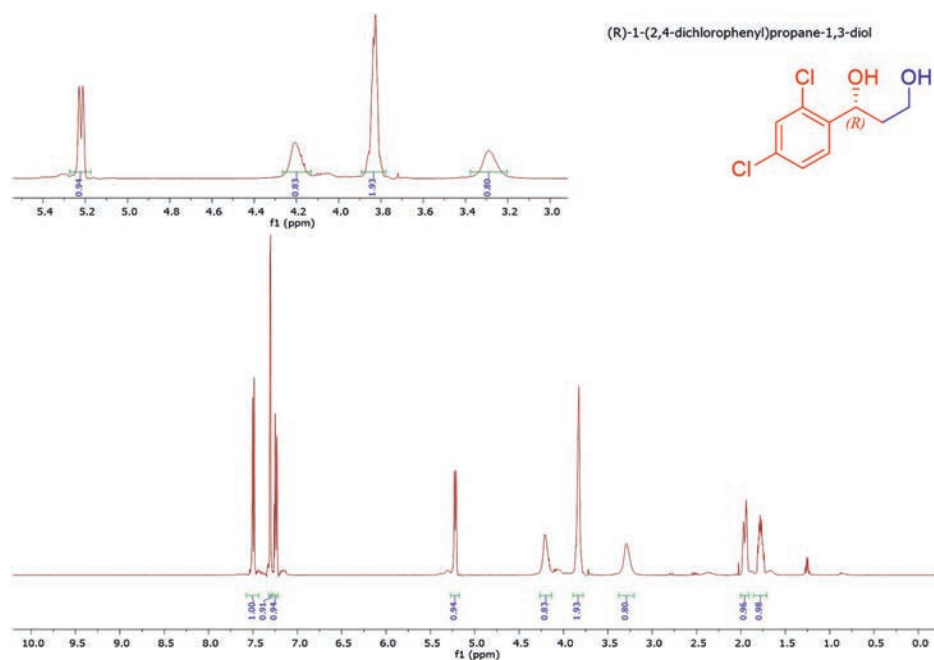
Figure S72. HPLC analysis of chemoenzymatic product 5c.



**(R)-1-(2,4-dichlorophenyl)propane-1,3-diol (5d).**<sup>[7]</sup> yellowish oil; yield = 62% (11 mg). <sup>1</sup>H NMR (500 MHz, CDCl<sub>3</sub>) δ 7.49 (d, *J* = 8.4 Hz, 1H), 7.30 (d, *J* = 2.0 Hz, 1H), 7.24 (dd, *J* = 8.4, 1.8 Hz, 1H), 5.22 (dd, *J* = 8.7 Hz, 1H), 4.21 (s, 1H), 3.88 – 3.77 (m, 2H), 3.29 (s, 1H), 2.01 – 1.89 (m, 1H), 1.86 – 1.72 (m, 1H). The <sup>1</sup>H NMR data is in agreement with published data.<sup>[7]</sup> The enantiomeric ratio of chemoenzymatic product **5d** was determined by reverse phase HPLC using a Chiralpak<sup>®</sup> ID column (150 mm × 4.6 mm, Daicel) (MeCN/water = 10:90, 25°C) at a flow rate of 1 mL/min. UV detection at 220 nm; *t<sub>R</sub>*:(major) = 17.9 min, (minor) = 19.1 min.



**Figure S73.** UV spectra monitoring the aldol addition of **1** (100 mM) to **2d** (2 mM) catalyzed by TAUT015 (0.33 mg/mL) in NaPi buffer [20 mM/5% (v/v) EtOH] at pH 6.5. **A.** Enzymatic reaction. **B.** Non-enzymatic reaction.



**Figure S74.** <sup>1</sup>H NMR spectrum of chemoenzymatic product **5d**.

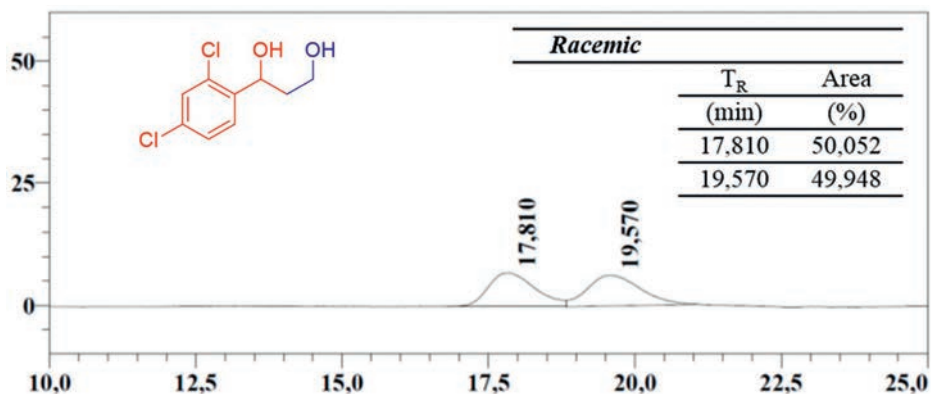


Figure S75. HPLC analysis of racemic 5d.

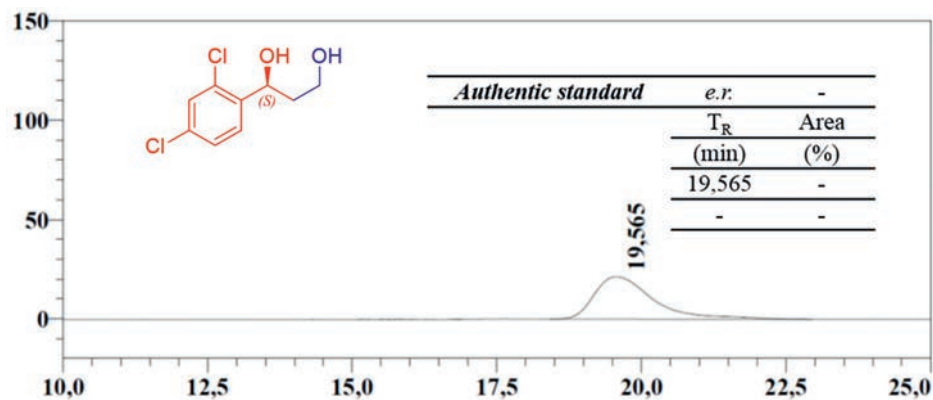


Figure S76. HPLC analysis of authentic standard (S)-5d.

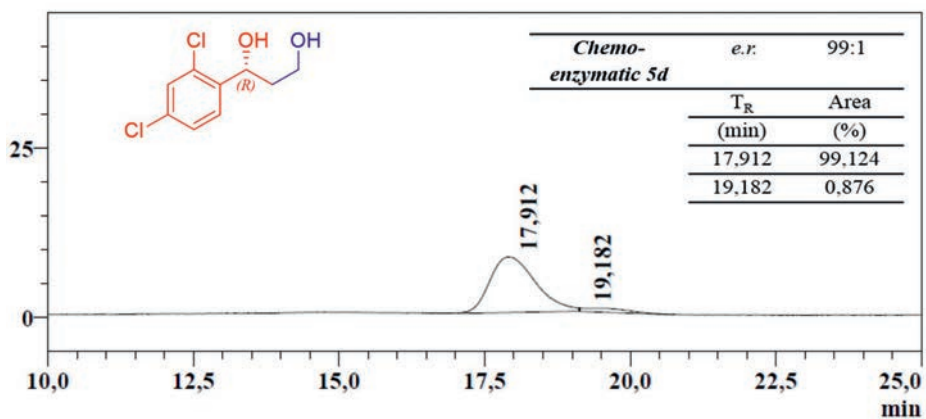
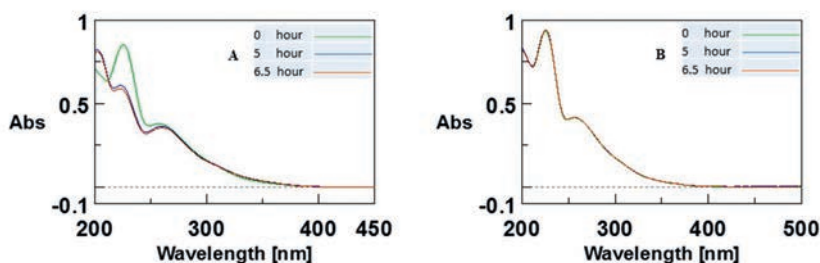
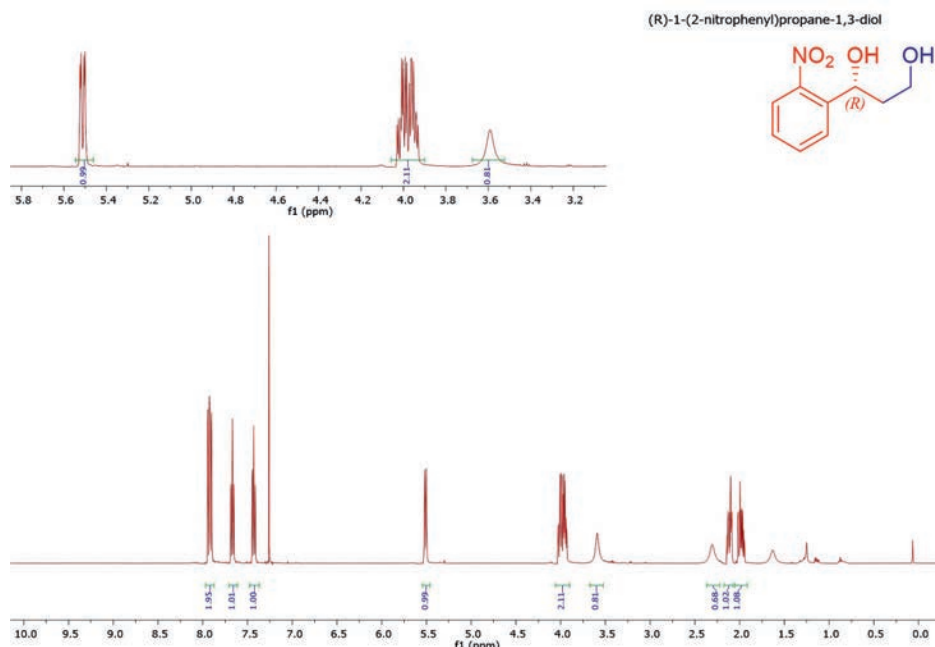


Figure S77. HPLC analysis of chemoenzymatic product 5d.

**(R)-1-(2-nitrophenyl)propane-1,3-diol (5f)**.<sup>[2b]</sup> yellowish oil; yield = 69% (11 mg). <sup>1</sup>H NMR (500 MHz, CDCl<sub>3</sub>) δ 7.92 (ddd, *J* = 11.9, 8.1, 1.2 Hz, 2H), 7.67 (td, *J* = 7.6, 1.1 Hz, 1H), 7.43 (td, *J* = 8.0, 7.5, 1.4 Hz, 1H), 5.51 (dd, *J* = 9.0, 2.5 Hz, 1H), 4.04 – 3.92 (m, 2H), 3.59 (s, 1H), 2.31 (s, 1H), 2.15 – 2.07 (m, 1H), 2.03 – 1.94 (m, 1H). The <sup>1</sup>H NMR data is in agreement with published data.<sup>[2b]</sup> The enantiomeric ratio of chemoenzymatic product **5f** was determined by reverse phase HPLC using a Chiralpak<sup>®</sup> ID column (150 mm × 4.6 mm, Daicel) (MeCN/water = 10:90, 25°C) at a flow rate of 1 mL/min. UV detection at 220 nm: *t*<sub>R</sub>:(major) = 5.8 min, (minor) = 6.8 min.



**Figure S78.** UV spectra monitoring the aldol addition of **1** (100 mM) to **2f** (2 mM) catalyzed by TAUT015 (0.33 mg/mL) in NaPi buffer [20 mM/5% (v/v) EtOH] at pH 6.5. A. Enzymatic reaction. B. Non-enzymatic reaction.



**Figure S79.** <sup>1</sup>H NMR spectrum of chemoenzymatic product **5f**.

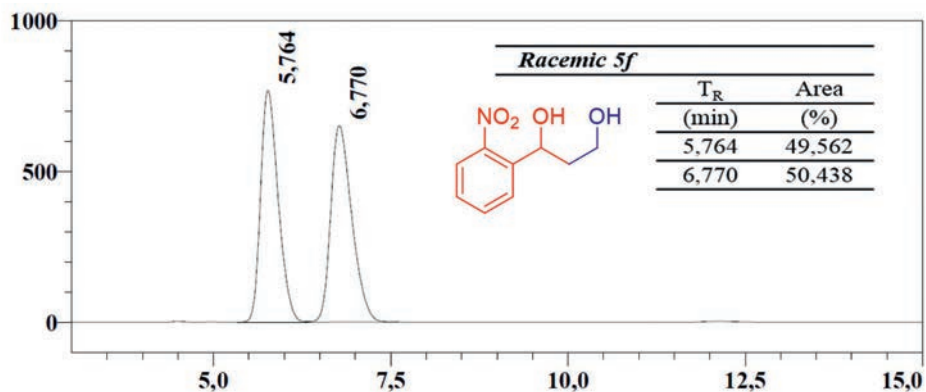


Figure S80. HPLC analysis of racemic 5f.

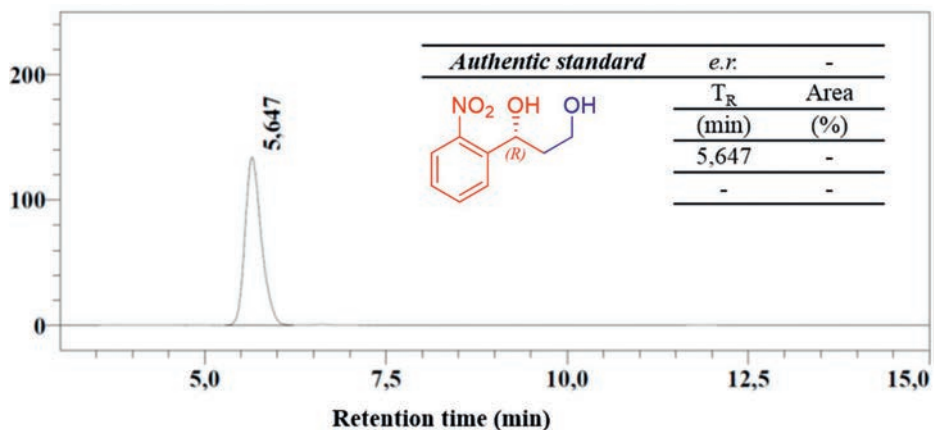


Figure S81. HPLC analysis of authentic standard (*R*)-5f.

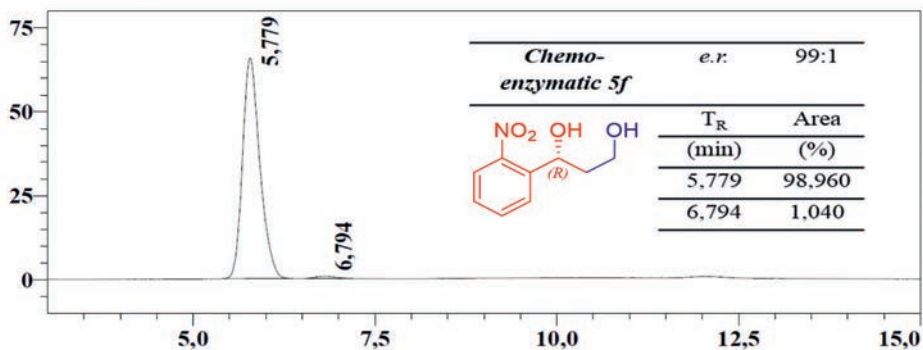
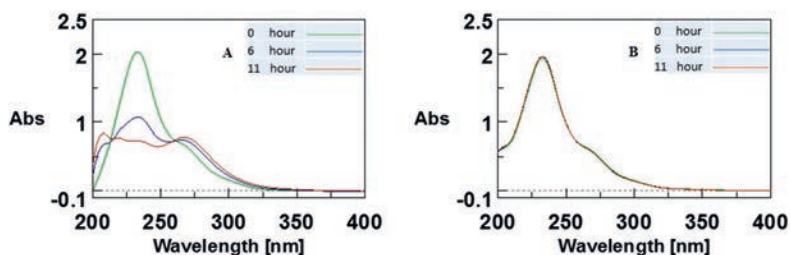
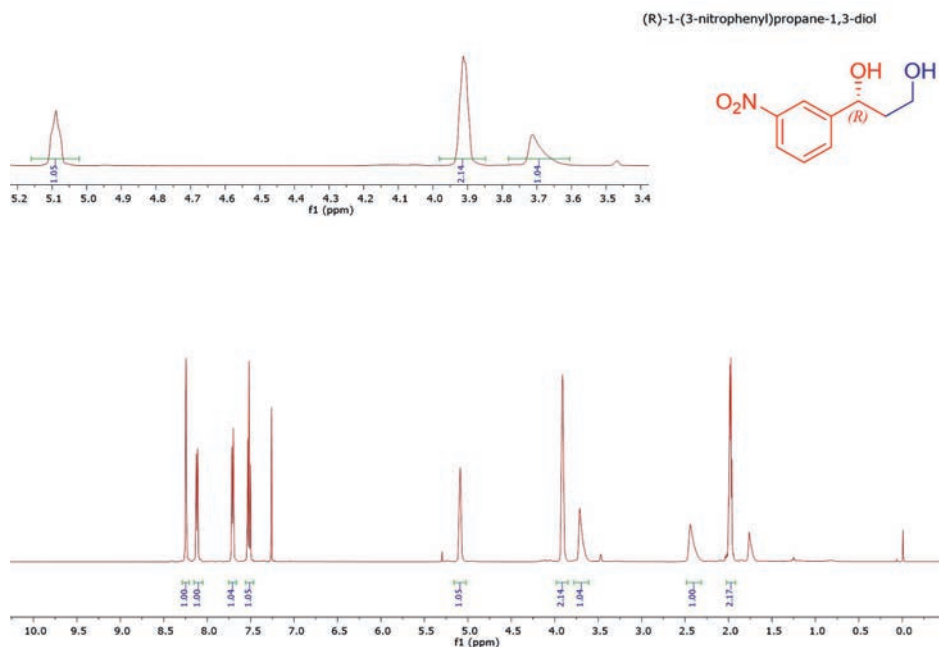


Figure S82. HPLC analysis of chemoenzymatic product 5f.

**(R)-1-(3-nitrophenyl)propane-1,3-diol (5g).**<sup>[2b]</sup> yellowish oil; yield = 56% (9 mg). <sup>1</sup>H NMR (500 MHz, CDCl<sub>3</sub>) δ 8.24 (t, *J* = 2.0 Hz, 1H), 8.12 (dd, *J* = 8.3, 2.2 Hz, 1H), 7.72-7.70 (m, 1H), 7.52 (t, *J* = 8.0, 1.7 Hz, 1H), 5.16 – 5.02 (m, 1H), 3.98 – 3.85 (m, 2H), 3.71 (s, 1H), 2.44 (s, 1H), 1.98 (ddd, *J* = 10.2, 6.2, 4.0 Hz, 2H). The <sup>1</sup>H NMR data is in agreement with published data.<sup>[2b]</sup> The enantiomeric ratio of enzymatic product **5g** was determined by reverse phase HPLC using a Chiralpak<sup>®</sup> ID column (150 mm × 4.6 mm, Daicel) (MeCN/water = 10:90, 25°C) at a flow rate of 1 mL/min. UV detection at 220 nm: *t<sub>R</sub>*:(major) = 5.8 min, (minor) = 8.8 min.



**Figure S83.** UV spectra monitoring the aldol addition of **1** (100 mM) to **2g** (2 mM) catalyzed by TAUT015 (0.33 mg/mL) in NaPi buffer [20 mM/5% (v/v) EtOH] at pH 6.5. **A.** Enzymatic reaction. **B.** Non-enzymatic reaction.



**Figure S84.** <sup>1</sup>H NMR spectrum of chemoenzymatic product **5g**.

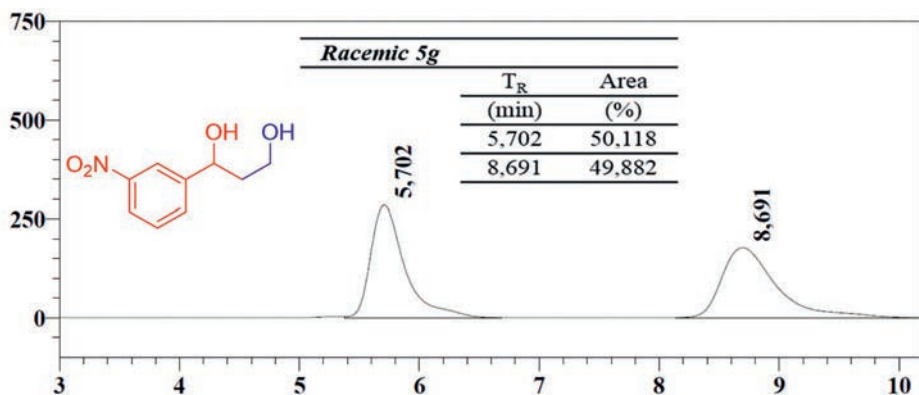


Figure S85. HPLC analysis of racemic

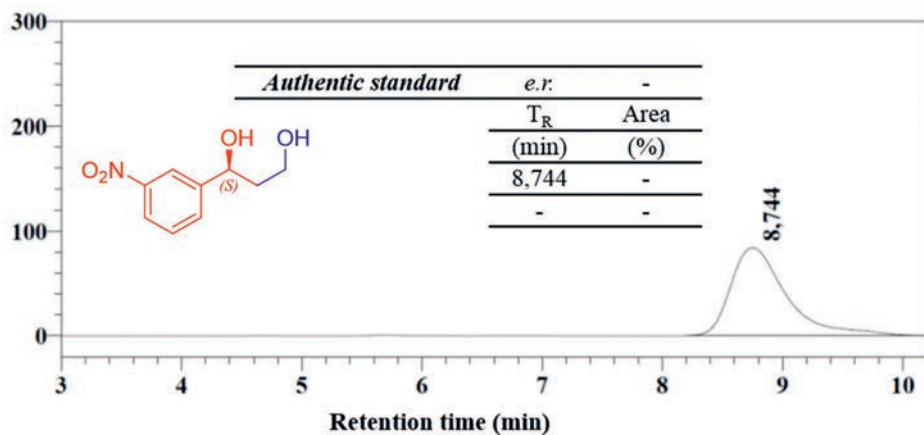


Figure S86. HPLC analysis of authentic standard (S)-5g.

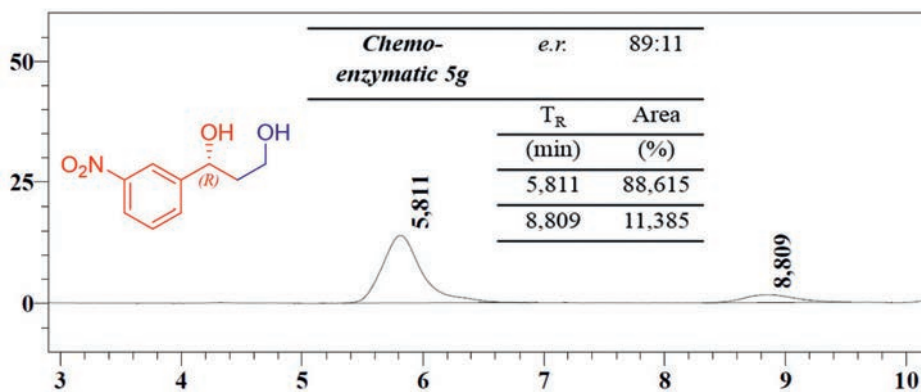
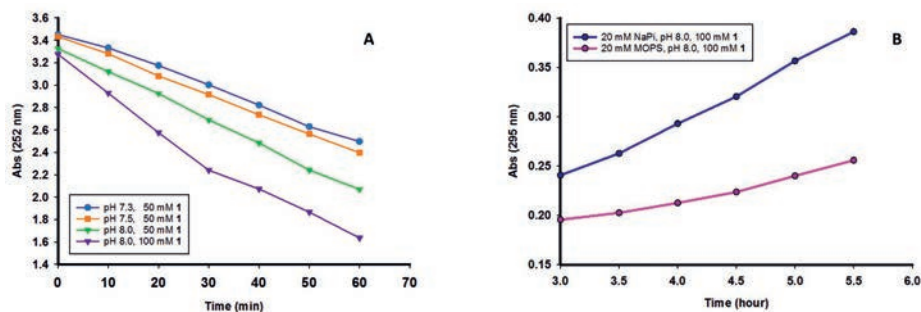


Figure S87. HPLC analysis of chemoenzymatic product 5g.



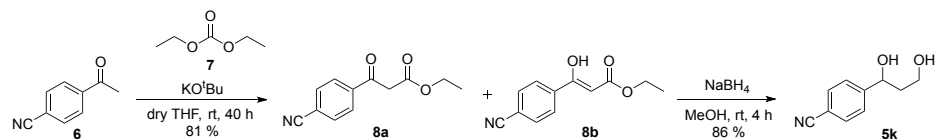
**Figure S88.** Effect of pH, acetaldehyde (**1**) concentration, and buffer on the aldol addition of **1** to **2k**. **A.** Conversion of **2k** in the aldol addition of **1** (50 mM or 100 mM) to **2k** (2 mM) catalyzed by 4-OT(M45T/F50A) (1 mg/mL) in buffer [20 mM NaPi/5% (v/v) DMSO, 0.3 mL] at pH 7.3, 7.5 or 8.0. **B.** Formation of corresponding cinnamaldehyde in the aldol addition of **1** (100 mM) with **2k** (2 mM) catalyzed by 4-OT(M45T/F50A) (1 mg/mL) in buffer [20 mM NaPi or MOPS/5% (v/v) DMSO, 0.3 mL] at pH 8.0.



**Figure S89.** Sequence alignment of 4-OT from *P. putida* mt-2 and TAUT015. Indicated with a star (\*) are positions 45 and 50 that were mutated to give 4-OT M45T/F50A. The sequence alignment was generated using the online Constraint-based Multiple Alignment Tool (COBALT, NCBI <https://www.ncbi.nlm.nih.gov/tools/cobalt/>) using default parameters.

## 6. Chemical synthesis of racemic 1,3-diols

### 6.1. Synthesis of compound 5k



Compound **5k** was prepared according to the route shown above, which involves the preparation of a mixture of compounds **8a** and **8b**.

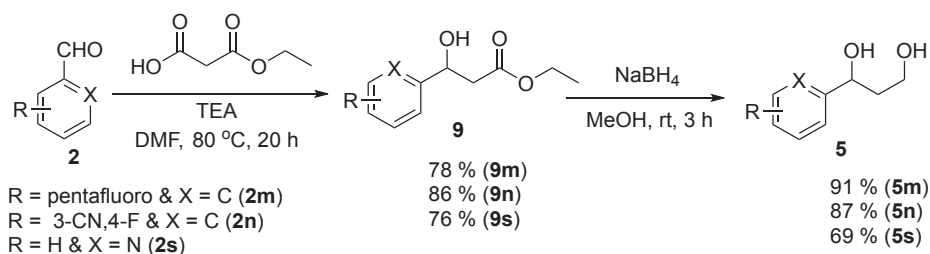
#### 6.1.1. Synthesis of compound 8

Compound **8**<sup>[8a]</sup> was prepared according to a literature procedure and its <sup>1</sup>H NMR spectrum matches with earlier reported NMR data.<sup>[8b]</sup>

### 6.1.2. Synthesis of compound 5k

The reported literature procedure was slightly modified.<sup>[9a]</sup> To a stirred solution of a mixture of **8a** and **8b** (1.0 equiv.) in dry MeOH (10 mL) at 0 °C was added NaBH<sub>4</sub> (3.0 equiv.) portion wise, and the reaction was run at room temp for 4 h. After completion of the reaction, as monitored by TLC (KMnO<sub>4</sub> test), the reaction mixture was extracted three times with EtOAc. The organic layers were combined, dried over Na<sub>2</sub>SO<sub>4</sub> and evaporated under *vacuo*. The crude product was purified by silica gel column chromatography using petroleum ether/ethyl acetate (90:10) as an eluent to give the corresponding racemic 1,3-diol **5k**. Colorless liquid; yield = 86% (140 mg, starting from 200 mg of **8**). The <sup>1</sup>H NMR data of compound **5k**<sup>[9b]</sup> match with earlier reported NMR data. <sup>1</sup>H NMR (500 MHz, CDCl<sub>3</sub>) δ 7.66 – 7.63 (m, 2H), 7.50 (d, *J* = 8.2 Hz, 2H), 5.05 (t, *J* = 6.1 Hz, 1H), 3.90 (t, *J* = 5.5 Hz, 2H), 3.35 (s, 1H), 2.15 – 2.07 (m, 1H), 1.99 – 1.90 (m, 2H).

### 6.2. Synthesis of compounds 5m-n and 5s



Compounds **5m-n** and **5s** were prepared according to the route shown above, which involves the preparation of compounds **9m-n** and **9s**.

#### 6.2.1. Synthesis of compounds 9m-n and 9s

Compounds **9m-n** and **9s** were prepared according to a literature procedure.<sup>[2a]</sup>

**Ethyl 3-hydroxy-3-(perfluorophenyl)propanoate (9m)**. <sup>1</sup>H NMR (500 MHz, CDCl<sub>3</sub>) δ 5.59 – 5.46 (m, 1H), 4.26 – 4.17 (m, 2H), 3.27 (d, *J* = 5.3 Hz, 1H), 3.13 (dd, *J* = 16.7, 9.4 Hz, 1H), 2.75 (dd, *J* = 16.7, 4.1 Hz, 1H), 1.28 (t, *J* = 7.1 Hz, 3H). <sup>13</sup>C NMR (126 MHz, CDCl<sub>3</sub>) δ 171.37, 146.12, 144.13, 142.08, 140.05, 138.74, 136.72, 115.25, 62.15, 61.42, 40.41, 14.12. HRMS (ESI+): calcd. for C<sub>11</sub>H<sub>8</sub>F<sub>5</sub>O<sub>2</sub> [MH-H<sub>2</sub>O]<sup>+</sup>: 267.0444, found: 267.0436.

**Ethyl 3-(3-cyano-4-fluorophenyl)-3-hydroxypropanoate (9n)**. <sup>1</sup>H NMR (500 MHz, CDCl<sub>3</sub>) δ 7.67 (dd, *J* = 6.0, 2.2 Hz, 1H), 7.65 – 7.59 (m, 1H), 7.21 (t, *J* = 8.6 Hz, 1H), 5.13



(dt,  $J = 8.0, 3.9$  Hz, 1H), 4.20 (q,  $J = 7.1$  Hz, 2H), 3.63 (d,  $J = 3.5$  Hz, 1H), 2.75 – 2.59 (m, 2H), 1.28 (t,  $J = 7.1$  Hz, 3H).  $^{13}\text{C}$  NMR (126 MHz,  $\text{CDCl}_3$ )  $\delta$  172.06, 163.62, 161.56, 139.82, 132.59, 132.52, 130.90, 116.75, 116.60, 113.94, 68.82, 61.40, 43.09, 14.24. HRMS (ESI+): calcd. for  $\text{C}_{12}\text{H}_{11}\text{FNO}_2$   $[\text{MH}-\text{H}_2\text{O}]^+$ : 220.0774, found: 220.0765.

**Ethyl 3-hydroxy-3-(pyridin-2-yl)propanoate (9s).**<sup>[10]</sup>  $^1\text{H}$  NMR (500 MHz,  $\text{CDCl}_3$ )  $\delta$  8.55 (d,  $J = 4.8$  Hz, 1H), 7.70 (td,  $J = 7.7, 1.7$  Hz, 1H), 7.42 (d,  $J = 7.9$  Hz, 1H), 7.21 (dd,  $J = 7.4, 4.9$  Hz, 1H), 5.18 (dt,  $J = 9.3, 5.0$  Hz, 1H), 4.28 (d,  $J = 5.6$  Hz, 1H), 4.18 (q,  $J = 7.1$  Hz, 2H), 2.90 (dd,  $J = 16.0, 4.1$  Hz, 1H), 2.76 (dd,  $J = 16.0, 8.4$  Hz, 1H), 1.25 (t,  $J = 7.1$  Hz, 3H). The  $^1\text{H}$  NMR data is in agreement with published data.<sup>[10]</sup>

### 6.2.2. Synthesis of compounds 5m-n and 5s

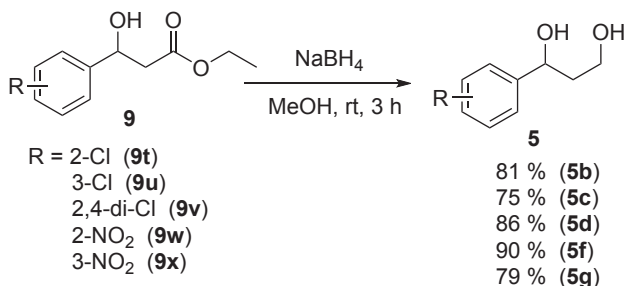
The compounds **5m-n** and **5s** were obtained by reducing **9m-n** and **9s** (1.0 equiv.) using  $\text{NaBH}_4$  (3.0 equiv.) following a procedure similar to that used for **5k**. The  $^1\text{H}$  NMR data of compounds **5m**<sup>[2b]</sup> and **5s**<sup>[5]</sup> match with earlier reported NMR data.

**1-(perfluorophenyl)propane-1,3-diol (5m).**  $^1\text{H}$  NMR (500 MHz,  $\text{CDCl}_3$ )  $\delta$  5.53 – 5.24 (m, 1H), 4.01 – 3.82 (m, 2H), 3.07 (s, 1H), 2.42 – 2.25 (m, 1H), 2.09 – 1.82 (m, 2H). The  $^1\text{H}$  NMR data is in agreement with published data.<sup>[2b]</sup>

**5-(1,3-dihydroxypropyl)-2-fluorobenzonitrile (5n).**  $^1\text{H}$  NMR (500 MHz,  $\text{CDCl}_3$ )  $\delta$  7.61 (dd,  $J = 6.0, 2.2$  Hz, 1H), 7.57 (ddd,  $J = 7.5, 5.1, 2.2$  Hz, 1H), 7.16 (t,  $J = 8.7$  Hz, 1H), 4.97 – 4.88 (m, 1H), 3.86 – 3.72 (m,  $J = 5.7$  Hz, 2H), 3.65 (s, 2H), 1.86 (q,  $J = 5.3$  Hz, 2H).  $^{13}\text{C}$  NMR (126 MHz,  $\text{CDCl}_3$ )  $\delta$  163.45, 161.39, 141.77, 141.74, 132.57, 132.50, 130.78, 116.63, 116.47, 114.15, 72.81, 61.37, 40.43. HRMS (ESI+): calcd. for  $\text{C}_{10}\text{H}_{11}\text{FNO}_2$   $[\text{M}+\text{H}]$ : 196.0729, found: 196.0769.

**1-(pyridin-2-yl)propane-1,3-diol (5s).**  $^1\text{H}$  NMR (500 MHz,  $\text{CDCl}_3$ )  $\delta$  8.43 (d,  $J = 4.9$  Hz, 1H), 7.67 (td,  $J = 7.7, 1.6$  Hz, 1H), 7.39 (d,  $J = 7.9$  Hz, 1H), 7.16 (dd,  $J = 7.2, 5.1$  Hz, 1H), 4.98 (dd,  $J = 8.9, 3.7$  Hz, 1H), 3.90 – 3.76 (m, 2H), 2.09 – 1.98 (m, 1H), 1.96 – 1.81 (m, 1H). The  $^1\text{H}$  NMR data is in agreement with published data.<sup>[5]</sup>

## 6.3. Synthesis of compounds 5b-d and 5f-g



The compounds **5b-d** and **5f-g** were obtained by reducing **9t-x** (1.0 equiv.) using NaBH<sub>4</sub> (3.0 equiv.) following a procedure similar to that used for **5k**. The <sup>1</sup>H NMR data of compounds **5b**<sup>[2b]</sup>, **5c**<sup>[11a]</sup>, **5d**<sup>[11b]</sup> and **5f-g**<sup>[2b]</sup> match with earlier reported NMR data.

**1-(2-chlorophenyl)propane-1,3-diol (5b)**. <sup>1</sup>H NMR (500 MHz, CDCl<sub>3</sub>) δ 7.57 (dd, *J* = 7.7, 1.6 Hz, 1H), 7.27 (qd, *J* = 7.7, 1.2 Hz, 2H), 7.17 (td, *J* = 7.6, 1.7 Hz, 1H), 5.29 (dd, *J* = 8.8, 2.9 Hz, 1H), 3.86 – 3.77 (m, 2H), 3.63 (s, 2H), 2.04 – 1.94 (m, 1H), 1.88 – 1.77 (m, 1H). The <sup>1</sup>H NMR data is in agreement with published data.<sup>[2b]</sup>

**1-(3-chlorophenyl)propane-1,3-diol (5c)**. <sup>1</sup>H NMR (500 MHz, CDCl<sub>3</sub>) δ 7.28 (s, 1H), 7.25 – 7.16 (m, 2H), 7.13 (dt, *J* = 6.7, 1.7 Hz, 1H), 4.78 (dd, *J* = 8.2, 4.4 Hz, 1H), 4.11 (s, 2H), 3.78 – 3.62 (m, 2H), 1.89 – 1.74 (m, 2H). The <sup>1</sup>H NMR data is in agreement with published data.<sup>[11a]</sup>

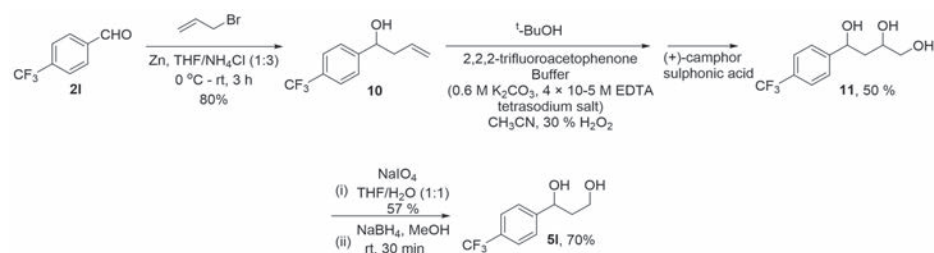
**1-(2,4-dichlorophenyl)propane-1,3-diol (5d)**. <sup>1</sup>H NMR (500 MHz, CDCl<sub>3</sub>) δ 7.45 (d, *J* = 8.4 Hz, 1H), 7.28 (d, *J* = 2.1 Hz, 1H), 7.21 (dd, *J* = 8.4, 2.0 Hz, 1H), 5.18 (dd, *J* = 8.9, 2.4 Hz, 1H), 4.55 (s, 1H), 3.80 – 3.76 (m, 3H), 1.96 – 1.88 (m, 1H), 1.79 – 1.68 (m, 1H). The <sup>1</sup>H NMR data is in agreement with published data.<sup>[11b]</sup>

**1-(2-nitrophenyl)propane-1,3-diol (5f)**. <sup>1</sup>H NMR (500 MHz, CDCl<sub>3</sub>) δ 7.86 – 7.80 (m, 2H), 7.62 – 7.55 (m, 1H), 7.38 – 7.31 (m, 1H), 5.42 (d, *J* = 7.7 Hz, 1H), 3.85 (m, 2H), 3.25 (s, 1H), 2.09 (s, 1H), 2.06 – 1.97 (m, 1H), 1.90 – 1.79 (m, 1H). The <sup>1</sup>H NMR data is in agreement with published data.<sup>[2b]</sup>

**1-(3-nitrophenyl)propane-1,3-diol (5g)**. <sup>1</sup>H NMR (500 MHz, CDCl<sub>3</sub>) δ 8.12 (s, 1H), 7.98 (dd, *J* = 8.1, 1.8 Hz, 1H), 7.61 (d, *J* = 7.7 Hz, 1H), 7.42 (t, *J* = 7.9 Hz, 1H), 4.98 (t, *J* = 6.1

Hz, 1H), 4.54 (s, 1H), 3.83 – 3.71 (m,  $J = 5.3$  Hz, 2H), 3.61 (s, 1H), 1.88 – 1.85 (m, 2H). The  $^1\text{H}$  NMR data is in agreement with published data.<sup>[2b]</sup>

#### 6.4. Synthesis of compound 51



Compound **51** was prepared according to a literature procedure and its  $^1\text{H}$  NMR data match with earlier reported data.<sup>[12]</sup>

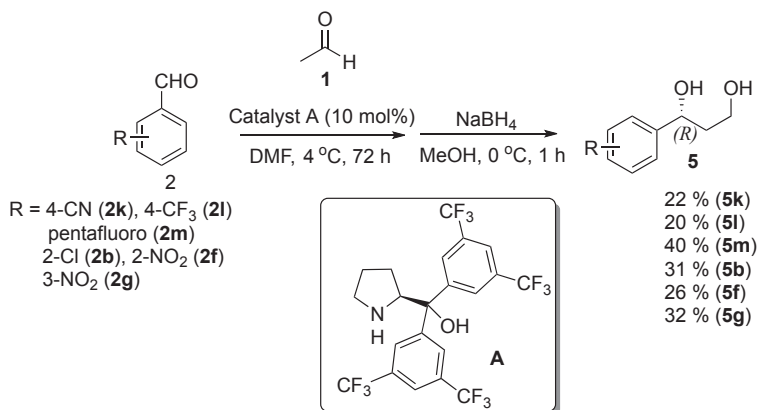
**1-(4-(trifluoromethyl)phenyl)but-3-en-1-ol (10)**.  $^1\text{H}$  NMR (500 MHz,  $\text{CDCl}_3$ )  $\delta$  7.61 (d,  $J = 8.2$  Hz, 2H), 7.48 (d,  $J = 8.1$  Hz, 2H), 5.87 – 5.71 (m, 1H), 5.23 – 5.14 (m, 2H), 4.81 (dd,  $J = 8.0, 4.7$  Hz, 1H), 2.59 – 2.51 (m, 1H), 2.51 – 2.41 (m, 1H), 2.13 (s, 1H). The  $^1\text{H}$  NMR data is in agreement with published data.<sup>[12a]</sup>

**4-(4-(trifluoromethyl)phenyl)butane-1,2,4-triol (11)**.  $^1\text{H}$  NMR (500 MHz,  $\text{CDCl}_3$ )  $\delta$  7.58 (d,  $J = 8.1$  Hz, 2H), 7.45 (d,  $J = 8.1$  Hz, 2H), 5.00 (dd,  $J = 10.0, 2.8$  Hz, 1H), 4.01 (ddd,  $J = 9.6, 6.5, 3.2$  Hz, 1H), 3.62 (dd,  $J = 11.2, 3.2$  Hz, 1H), 3.47 (dd,  $J = 11.2, 6.8$  Hz, 1H), 1.85 (dt,  $J = 14.6, 9.8$  Hz, 1H), 1.72 (dt,  $J = 14.5, 2.8$  Hz, 1H). The  $^1\text{H}$  NMR data is in agreement with published data.<sup>[12b]</sup>

**1-(4-(trifluoromethyl)phenyl)propane-1,3-diol (51)**.  $^1\text{H}$  NMR (500 MHz,  $\text{CDCl}_3$ )  $\delta$  7.61 (d,  $J = 8.2$  Hz, 2H), 7.50 (d,  $J = 8.1$  Hz, 2H), 5.05 (dd,  $J = 8.1, 4.1$  Hz, 1H), 3.90 (t,  $J = 5.6$  Hz, 2H), 2.03 – 1.91 (m, 2H). The  $^1\text{H}$  NMR data is in agreement with published data.<sup>[12d]</sup>

## 7. Chemical synthesis of enantiopure 1,3-diols using Jorgenson and Hayashi catalyst

### 7.1. Chemical synthesis of enantiopure 1,3-diols 5k-m, 5b and 5f-g



Compounds **5k-m**<sup>[2b]</sup>, **5b**<sup>[2b]</sup> and **5f-g**<sup>[2b]</sup> were prepared according to literature procedures and their <sup>1</sup>H NMR spectra match with earlier reported NMR data.<sup>[2b]</sup>

**(R)-4-(1,3-dihydroxypropyl)benzotrile (5k)**. <sup>1</sup>H NMR (500 MHz, CDCl<sub>3</sub>) δ 7.65 (d, *J* = 8.2 Hz, 2H), 7.50 (d, *J* = 8.4 Hz, 2H), 5.11 – 4.99 (m, 1H), 3.93 – 3.90 (m, 2H), 3.64 (dd, *J* = 11.0, 3.1 Hz, 1H), 3.40 (dd, *J* = 11.0, 7.8 Hz, 1H), 1.98 – 1.94 (m, 2H). The <sup>1</sup>H NMR data is in agreement with published data.<sup>[2b]</sup>

**(R)-1-(4-(trifluoromethyl)phenyl)propane-1,3-diol (5l)**. <sup>1</sup>H NMR (500 MHz, CDCl<sub>3</sub>) δ 7.62 (d, *J* = 8.1 Hz, 2H), 7.50 (d, *J* = 8.4 Hz, 2H), 5.06 (dd, *J* = 8.2, 4.1 Hz, 1H), 3.94 – 3.87 (m, 2H), 2.00 – 1.96 (m, 2H). The <sup>1</sup>H NMR data is in agreement with published data.<sup>[2b]</sup>

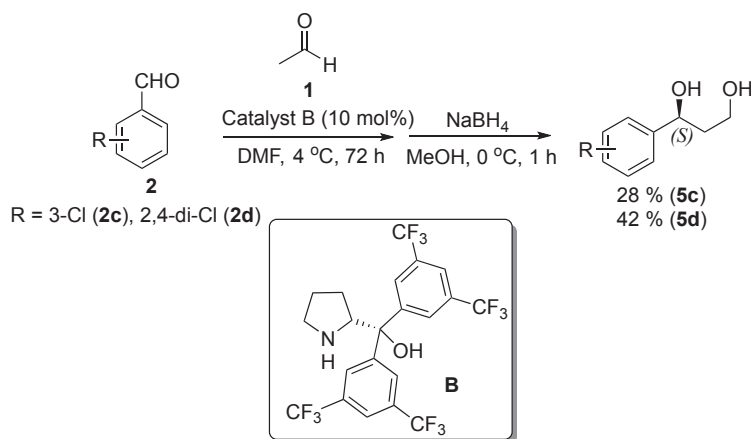
**(R)-1-(perfluorophenyl)propane-1,3-diol (5m)**. <sup>1</sup>H NMR (500 MHz, CDCl<sub>3</sub>) δ 5.40 – 5.32 (m, 1H), 4.00 – 3.87 (m, 2H), 3.02 (d, *J* = 4.4 Hz, 1H), 2.35 (ddt, *J* = 17.4, 9.2, 4.1 Hz, 1H), 1.97 (ddt, *J* = 14.4, 6.7, 3.9 Hz, 1H), 1.81 (s, 1H). The <sup>1</sup>H NMR data is in agreement with published data.<sup>[2b]</sup>

**(R)-1-(2-chlorophenyl)propane-1,3-diol (5b)**. <sup>1</sup>H NMR (500 MHz, CDCl<sub>3</sub>) δ 7.62 (dd, *J* = 7.7, 1.4 Hz, 1H), 7.34 – 7.26 (m, 2H), 7.20 (td, *J* = 7.6, 1.7 Hz, 1H), 5.34 (dd, *J* = 8.8, 2.7 Hz, 1H), 3.93 – 3.85 (m, 2H), 3.45 (s, 1H), 2.68 (s, 1H), 2.06 – 2.01 (m, 1H), 1.95 – 1.84 (m, 1H). The <sup>1</sup>H NMR data is in agreement with published data.<sup>[2b]</sup>

**(R)-1-(2-nitrophenyl)propane-1,3-diol (5f).**  $^1\text{H NMR}$  (500 MHz,  $\text{CDCl}_3$ )  $\delta$  7.93 (ddd,  $J = 11.6, 8.0, 1.4$  Hz, 2H), 7.67 (td,  $J = 7.7, 1.4$  Hz, 1H), 7.43 (ddd,  $J = 8.5, 7.4, 1.5$  Hz, 1H), 5.52 (dd,  $J = 9.0, 2.6$  Hz, 1H), 4.09 – 3.91 (m, 2H), 3.54 (s, 1H), 2.23 (s, 1H), 2.17 – 2.08 (m, 1H), 2.04 – 1.93 (m, 1H). The  $^1\text{H NMR}$  data is in agreement with published data.<sup>[2b]</sup>

**(R)-1-(3-nitrophenyl)propane-1,3-diol (5g).**  $^1\text{H NMR}$  (500 MHz,  $\text{CDCl}_3$ )  $\delta$  8.13 (s, 1H), 7.98 (ddd,  $J = 8.2, 2.3, 1.0$  Hz, 1H), 7.60 (d,  $J = 7.7$  Hz, 1H), 7.41 (t,  $J = 7.9$  Hz, 1H), 4.94 (t,  $J = 15.4, 9.7$  Hz, 1H), 4.91 (s, 1H), 4.03 (s, 1H), 3.80 – 3.67 (m, 2H), 1.89 – 1.83 (m, 2H). The  $^1\text{H NMR}$  data is in agreement with published data.<sup>[2b]</sup>

## 7.2. Chemical synthesis of enantiopure 1,3-diols 5c-d using Jorgenson and Hayashi catalyst



The compounds **5c**<sup>[2b]</sup> and **5d**<sup>[2b]</sup> were obtained by following a procedure similar to that used for **5k** using enantiocomplementary catalyst **B**, and their  $^1\text{H NMR}$  spectra match with earlier reported NMR data.<sup>[6,13]</sup>

**(S)-1-(3-chlorophenyl)propane-1,3-diol (5c).**  $^1\text{H NMR}$  (500 MHz,  $\text{CDCl}_3$ )  $\delta$  7.35 (s, 1H), 7.30 – 7.17 (m, 3H), 4.88 (dd,  $J = 8.2, 4.2$  Hz, 1H), 3.80 (m, 3H), 3.41 (s, 1H), 1.98 – 1.84 (m, 2H). The  $^1\text{H NMR}$  data is in agreement with published data.<sup>[6,13]</sup>

**(S)-1-(2,4-dichlorophenyl)propane-1,3-diol (5d).**  $^1\text{H NMR}$  (500 MHz,  $\text{CDCl}_3$ )  $\delta$  7.57 (d,  $J = 8.4$  Hz, 1H), 7.37 – 7.27 (m, 2H), 5.29 (d,  $J = 8.2$  Hz, 1H), 3.96 – 3.85 (m, 2H), 3.78 (s, 1H), 2.76 (s, 1H), 2.07 – 1.97 (m, 1H), 1.91 – 1.80 (m, 1H). The  $^1\text{H NMR}$  data is in agreement with published data.<sup>[13]</sup>

## 8. References

1. (a) Guo, C.; Saifuddin, M.; Saravanan, T.; Sharifi, M.; Poelarends, G. J. Biocatalytic Asymmetric Michael Additions of Nitromethane to  $\alpha,\beta$ -Unsaturated Aldehydes via Enzyme-bound Iminium Ion Intermediates. *ACS Catal.* **2019**, *9*, 4369–4373; (b) Wang, M. M.; Ning, X. S.; Qu, J. P.; Kang, Y. B. Dehydrogenative Synthesis of Linear  $\alpha,\beta$ -Unsaturated Aldehydes with Oxygen at Room Temperature Enabled by tBuONO. *ACS Catal.* **2017**, *7*, 4000–4003; (c) Hayashi, Y.; Sakamoto, D.; Okamura, Daichi. One-Pot Synthesis of (S)-Baclofen via Aldol Condensation of Acetaldehyde with Diphenylprolinol Silyl Ether Mediated Asymmetric Michael Reaction as a Key Step. *Org. Lett.* **2016**, *18*, 4–7; (d) Yamazaki, S.; Sugiura, H.; Ohashi, S.; Ishizuka, K.; Saimu, R.; Mikata, Y.; Ogawa, A. Intramolecular [2 + 2] and [4 + 2] Cycloaddition Reactions of Cinnamylamides of Ethenetricarboxylate in Sequential Processes. *J. Org. Chem.* **2016**, *81*, 10863–10886; (e) EL-Atawy, M. A.; Ferretti, F.; Ragaini, F. A Synthetic Methodology for Pyrroles from Nitrodienes. *Eur. J. Org. Chem.* **2018**, 4818–4825.
2. (a) Koszelewski, D.; Zysk, M.; Brodzka, A.; Żądło, A.; Paprocki, D.; Ostaszewski, R. Evaluation of a New Protocol for Enzymatic Dynamic Kinetic Resolution of 3-Hydroxy-3-(aryl)propanoic Acids. *Org. Biomol. Chem.* **2015**, *13*, 11014–11020; (b) Hayashi, Y.; Itoh, T.; Aratake, S.; Ishikawa, H. A Diarylprolinol in an Asymmetric, Catalytic, and Direct Crossed- Aldol Reaction of Acetaldehyde. *Angew. Chem. Int. Ed.* **2008**, *47*, 2082–2084.
3. Zandvoort, E.; Geertsema, E. M.; Baas, B. J.; Quax, W. J.; Poelarends, G. J. Bridging Between Organocatalysis and Biocatalysis: Asymmetric Addition of Acetaldehyde to  $\beta$ -Nitrostyrenes Catalyzed by a Promiscuous Proline-Based Tautomerase. *Angew. Chem. Int. Ed.* **2012**, *51*, 1240–1243.
4. Qiao, Y.; Chen, Q.; Lin, S.; Ni, B.; Headley, A. D. Organocatalytic Direct Asymmetric Crossed-Aldol Reactions of Acetaldehyde in Aqueous Media. *J. Org. Chem.* **2013**, *78*, 2693–2697.
5. Kumar, M.; Shah, B. A.; Tanejaa, S. C. Tandem Catalysis by Lipase in a Vinyl Acetate Mediated Cross-Aldol Reaction. *Adv. Synth. Catal.* **2011**, *353*, 1207–1212.
6. Erion, M. D.; Reddy, K. R.; Boyer, S. H.; Matelich, M. C.; Gomez-Galeno, J.; Lemus, R. H.; Ugarkar, B. G.; Colby, T. J.; Schanzer, J.; van Poelje, P. D. Design, Synthesis, and Characterization of a Series of Cytochrome P450 3A-Activated Prodrugs (HepDirect Prodrugs) Useful for Targeting Phosph(on)ate-Based Drugs to the Liver. *J. Am. Chem. Soc.* **2004**, *126*, 5154–5163.
7. Fan, X.; Rodriguez-Esrich, C.; Wang, S.; Sayalero, S.; Pericas, M. A. Highly Enantioselective Cross-Aldol Reactions of Acetaldehyde Mediated by a Dual Catalytic System Operating under Site Isolation. *Chem. Eur. J.* **2014**, *20*, 13089–13093.
8. (a) CARPINO, Albert, P. (PFIZER INC.) WO 2013/068875 A1, **2013**; (b) Korsager, S.; Nielsen, D. U.; Taaning, R. H.; Skrydstrup, T. Access to  $\beta$ -Keto Esters by Palladium-Catalyzed Carbonylative Coupling of Aryl Halides with Monoester Potassium Malonates. *Angew. Chem. Int. Ed.* **2013**, *52*, 9763–9766.

## Chapter 6

9. (a) Kim, H.; Lee, C. Nickel-Catalyzed Reductive Cyclization of Organohalides. *Org. Lett.* **2011**, *13*, 2050-2053; (b) Banik, S. M.; Mennie, K. M.; Jacobsen, E. N. Catalytic 1,3-Difunctionalization via Oxidative C–C Bond Activation. *J. Am. Chem. Soc.* **2017**, *139*, 9152-9155.
10. Baudoux, J.; Lefebvre, P.; Legay, R.; Lasne, M. C.; Rouden, J. Environmentally Benign Metal-free Decarboxylative Aldol and Mannich Reactions. *Green Chem.*, **2010**, *12*, 252-259.
11. (a) Borowiecki, P.; Wawro, A. M.; Winska, P.; Wielechowska, M.; Bretner, M. Synthesis of Novel Chiral TBBt Derivatives with Hydroxyl Moiety. Studies on Inhibition of Human Protein Kinase CK2a and Cytotoxicity Properties. *Eur. J. Med. Chem.* **2014**, *84*, 364-374; (b) Reddy, K. R.; Matelich, M. C.; Ugarkar, B. G.; Gómez-Galeno, J. E.; DaRe, J.; Ollis, K.; Sun, Z.; Craigo, W.; Colby, T. J.; Fujitaki, J. M.; Boyer, S. H.; van Poelje, P. D.; Erion, M. D. Pradefovir: A Prodrug That Targets Adefovir to the Liver for the Treatment of Hepatitis B. *J. Med. Chem.* **2008**, *51*, 666–676.
12. (a) Dam, J. H.; Fristrup, P.; Madsen, R. Combined Experimental and Theoretical Mechanistic Investigation of the Barbier Allylation in Aqueous Media. *J. Org. Chem.* **2008**, *73*, 3228-3235; (b) Theodorou, A.; Triandafillidi, I.; Kokotos, C. G. Green Organocatalytic Dihydroxylation of Alkenes. *Eur. J. Org. Chem.* **2017**, 1502–1509; (c) Bosset, C.; Angibaud, P.; Stanfield, I.; Meerpoel, L.; Berthelot, D.; Guérinot, A.; Cossy, J. Iron-Catalyzed Synthesis of C2 Aryl- and N-Heteroaryl-Substituted Tetrahydropyrans. *J. Org. Chem.* **2015**, *80*, 12509-12525; (d) Twilton, J.; Christensen, M.; DiRocco, D. A.; Ruck, R. T.; Davies, I. W.; MacMillan, D. W. C. Selective Hydrogen Atom Abstraction through Induced Bond Polarization: Direct  $\alpha$ -Arylation of Alcohols through Photoredox, HAT, and Nickel Catalysis. *Angew. Chem. Int. Ed.* **2018**, *57*, 5369–5373.
13. Kumar, M.; Kumar, A.; Rizvi, M.; Mane, M.; Vanka, K.; Taneja, S. C.; Shah, B. A. Synthesis of  $\alpha,\beta$ -Unsaturated  $\delta$ -Lactones by Vinyl Acetate Mediated Asymmetric Cross-Aldol Reaction of Acetaldehyde: Mechanistic Insights. *Eur. J. Org. Chem.* **2014**, 5247–5255.







# *Chapter 7*

*Summary and Future Perspectives*

---

## Summary

### Repurposing enzymes

Enzymes are highly complex molecules, carefully folded to position functional groups and form a structure that can accommodate substrate molecules and promote a chemical reaction. Despite their high complexity, enzymes can be precisely engineered using recombinant DNA technology. By mutating particular base pairs in the gene coding for an enzyme, a different enzyme variant can be produced with, sometimes, vastly different properties. However, predicting the effect of a mutation on the properties of an enzyme is still highly challenging. As a consequence, a common strategy to engineer enzymes is to generate many different mutants of an enzyme and test all mutants for a desired property. Individual mutants showing the desired property can then be selected and subjected to a new round of mutagenesis. This iterative process is referred to as directed evolution and has dramatically impacted the biocatalysis community, recently recognized by the awarding of one half of the 2018 Nobel Prize in chemistry to Frances H. Arnold for her work on the directed evolution of enzymes<sup>1</sup>.

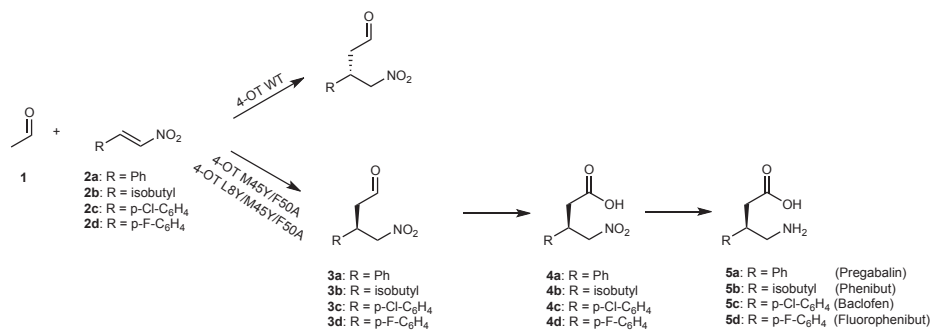
It is difficult to underestimate the power of directed evolution as scientists have demonstrated that computationally designed enzymes, with relative low activity, can be engineered to reach activities comparable to natural enzymes, simply by repeating the cycle of directed evolution many times<sup>2,3</sup>. In recent years there are some astounding examples of enzymes that were repurposed via directed evolution to catalyze abiological reactions<sup>4,5</sup>. The property of many enzymes to catalyze other reactions than their natural ones is also exemplified by the enzyme 4-oxalocrotonate tautomerase (4-OT), which has several promiscuous activities, including the ability to catalyze interesting C-C bond-forming reactions such as the Michael-type addition of aldehydes to nitroalkenes and the aldol condensation of acetaldehyde with benzaldehyde<sup>6,7</sup>.

### Engineering 4-OT

Although the activities are still modest, 4-OT might provide a good starting point for further engineering towards new C-C bond-forming 'Michaelases' or aldolases. The small size of 4-OT (only 62 amino acids) provides an excellent scaffold for an upcoming enzyme engineering strategy termed as mutability-landscape guided enzyme engineering (**chapter 1**). This strategy involves the systematic screening of (nearly) all single mutants of an enzyme for a desired property, the results of which can be graphically represented in a mutability landscape. Contrary to traditional engineering methodologies, this strategy does not only provide information on beneficial mutations, but also on neutral and detrimental mutations. This provides valuable information on the mutational robustness

of an enzyme and on functionally important residue positions. There are two approaches to obtain experimental data for the creation of a protein mutability landscape. The first approach involves the construction and characterization of a defined collection of (nearly) all single mutants and the second approach is called deep mutational scanning. The latter approach can only be used in combination with high-throughput screening assays because oversampling is required, whereas the first approach can also be used with lower throughput screening assays. Mutability-landscape guided enzyme engineering has been performed on several different enzymes, but the best characterized enzyme is 4-OT, with reported mutability landscapes on tautomerase activity, 'Michaelase' activity, enantioselectivity and enzyme expression.

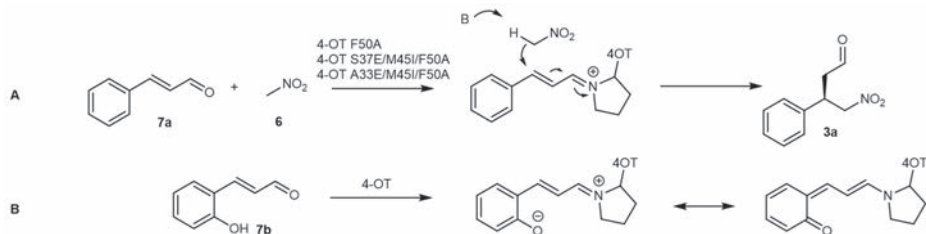
In a thorough study on the enantioselectivity of the 4-OT-catalyzed Michael-type addition, a double mutant, 4-OT M45Y/F50A was identified with inverted enantioselectivity towards the addition of acetaldehyde (**1**) to nitroalkene acceptors (**2a-d**)<sup>8</sup>. The resultant products **3a-d** are precursors for the pharmaceutically relevant enantiomers of a set of  $\gamma$ -aminobutyric acid (GABA) analogues (**5a-d**, Figure 1). However, the enantioselectivity of 4-OT M45Y/F50A is relatively poor, synthesizing **3a-d** with an enantiomeric ratio (e.r.) between 62:38 and 96:4. Therefore, in **chapter 2** we aimed to improve the enantioselectivity of 4-OT M45Y/F50A for the Michael-type addition of **1** to **2a-d**. We initially studied the effect of cosolvents on the enantioselectivity for the addition of **1** to **2b**, catalyzed by 4-OT M45Y/F50A. We found that DMSO and ethanol negatively affected the enantioselectivity of the enzyme, whereas small diols, such as 1,3-propanediol and ethylene glycol, positively affected the enantioselectivity improving the e.r. of **3b** from 90:10 to 95:5. However, enzyme engineering was necessary to further enhance the enantioselectivity. Guided by a reported crystal structure of 4-OT M45Y/F50A in complex with **2a**<sup>8</sup>, we randomized Ile-2 and Leu-8 and screened for mutants with improved enantioselectivity for the addition of **1** to **2b**. We found that substitution of Leu-8 to phenylalanine or tyrosine significantly improved the enantioselectivity, allowing the synthesis of **3b** with an e.r. of 99:1. 4-OT L8Y/M45Y/F50A proved to be more active and more stable in the presence of the cosolvent ethanol compared to 4-OT L8F/M45Y/F50A, and hence we continued with this mutant. Crystallization of 4-OT L8Y/M45Y/F50A revealed that the volume of the enzyme active site was reduced compared to 4-OT M45Y/F50A. We hypothesized that this reduces the rotational freedom of the nitroalkene substrate and hence promotes the enantioselective synthesis of products **3a-d**.



**Figure 1. Synthesis route for GABA analogues.** The asymmetric Michael-type addition of acetaldehyde (**1**) to nitroalkenes **2a-d** catalyzed by 4-OT yields  $\gamma$ -nitroaldehydes **3a-d**, which can be converted into  $\gamma$ -nitrocarboxylic acids **4a-d** and finally into GABA analogues **5a-d**.

Solubilization of nitroalkene substrates requires the use of cosolvents. High concentrations of cosolvents can significantly destabilize enzymes, as enzymes have evolved for catalysis in the aqueous environment of the cell. In **chapter 3** we employed mutability-landscape guided engineering to identify “hot-spot” positions that promote catalysis in the presence of high concentrations of ethanol. Two “hot-spot” positions were identified, Ser-30 and Ala-33. Targeting “hot-spot” position Ala-33 in the context of the highly enantioselective, but ethanol-sensitive, mutant 4-OT L8F/M45Y/F50A (**chapter 2**), resulted in a set of quadrupole mutants with significantly improved performance in high concentrations of ethanol. The best mutant, 4-OT L8F/A33I/M45Y/F50A, could be used to catalyze the Michael-type addition of **1** to **2a** in the presence of up to 40% v/v ethanol, whereas 4-OT L8F/M45Y/F50A lost all activity already in 10% v/v ethanol. This is an interesting showcase of the power of mutability-landscape guided enzyme engineering, where targeting of a single “hot-spot” position results in a dramatic increase in ethanol resistance of an enzyme.

We have shown that mutability-landscape guided enzyme engineering affords valuable information on positions in the enzyme that affect a desired characteristic. However, follow-up engineering can still be cumbersome, as even a library specifically targeting “hot-spot” positions often consists of a large fraction of inactive mutants. As a result, the screening of a library of mutants is typically the bottleneck in enzyme engineering<sup>9</sup>. Therefore, in **chapter 4**, we developed a novel pre-screening assay to exclude inactive mutants from the subsequent screening. We focused on the 4-OT-catalyzed Michael-type addition of nitromethane (**6**) to cinnamaldehyde (**7a**) forming **3a** (Figure 2). The mechanism of this reaction involves the formation of a covalent iminium species between **7a** and Pro-1 of 4-OT, which subsequently undergoes nucleophilic attack by **6**.



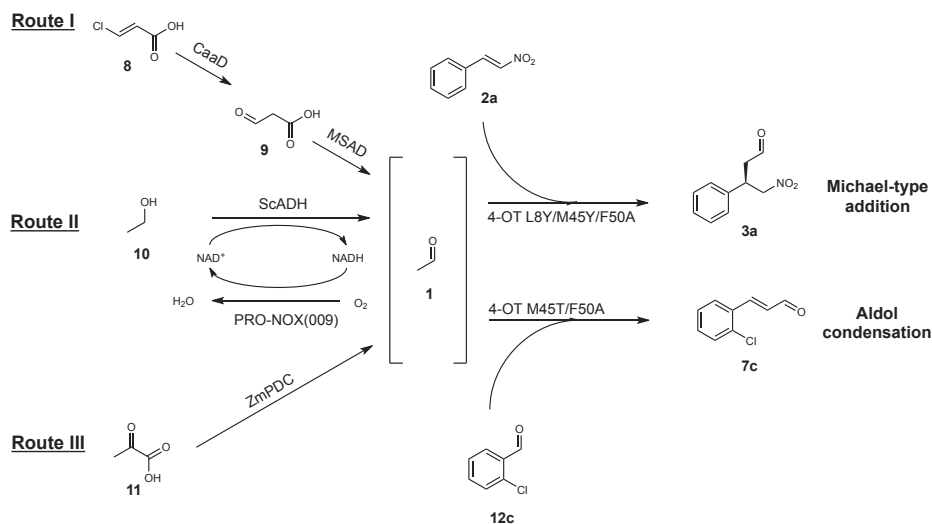
**Figure 2.** 4-OT catalyzed Michael-type addition of **6** to **7a** and iminium activation of **7b** by 4-OT. **A)** Enantioselective iminium-mediated Michael-type addition of **6** to **7a** catalyzed by 4-OT mutants. **B)** Iminium activation of **7b** by 4-OT forming a zwitterionic merocyanine-dye complex.

We showed that the reaction of Pro-1 with 2-hydroxycinnamaldehyde (**7b**) results in the formation of a zwitterionic, merocyanine-dye complex, with notable absorption at 516 nm that can clearly be observed by the naked eye. 4-OT mutants that did not form this brightly colored species proved to be poor catalysts for the Michael-type addition of **6** to **7a**. We exploited this in a facile solid-phase pre-screening assay, which reduced the screening burden up to 20-fold. After two rounds of directed evolution, two mutants were identified, 4-OT S37E/M45I/F50A and 4-OT A33E/M45I/F50A, which showed up to 39-fold improvement in activity.

### Applying 4-OT in multistep cascade reactions

The enzyme 4-OT provides an interesting platform to explore new (promiscuous) catalytic activities which could be applied for the synthesis of various pharmaceuticals. The successful engineering of 4-OT described above highlights the potential of 4-OT as an efficient biocatalyst. However, the synthesis of pharmaceuticals typically consists of many consecutive reactions, and product loss during intermediate workup can significantly reduce the overall yield of the process. The combination of (bio)catalysts in one-pot reaction cascades is an emerging concept in biocatalysis which prevents the necessity of intermediate purification. In **chapter 2** we use the newly engineered artificial ‘Michaelase’ 4-OT L8Y/M45Y/F50A in combination with a natural aldehyde dehydrogenase, PRO-ALDH(003), and a natural NADH oxidase, PRO-NOX(009) in a 2-step reaction cascade affording nitrocarboxylic acids **4a-d** in good isolated yield (up to 70%) and excellent enantiopurity (e.r. up to 99:1). Finally, using nickel boride as chemocatalyst, we also developed a three-step one-pot chemoenzymatic reaction cascade synthesizing the final GABA-analogues **5a-d** in good isolated yield (up to 70%) and with excellent e.r. values of up to 99:1.

4-OT is unique in that it can use acetaldehyde (**1**) as a nucleophile in Michael-type additions to nitroalkenes or the aldol condensation with benzaldehyde (**12a**). However, **1** is highly toxic and reactive and as a consequence requires intricate handling, which can impede its usage in practical synthesis. Therefore, in **chapter 5**, we report a study on three complementary routes for the *in situ* synthesis of **1** (Figure 3).

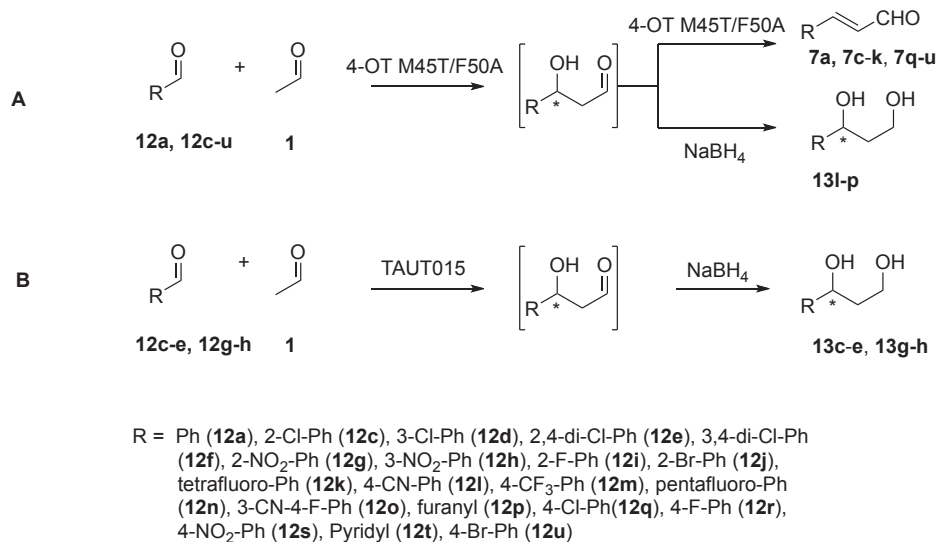


**Figure 3. Enzymatic complementary routes for *in situ* generation of acetaldehyde (**1**).** The *in situ* synthesized **1** is used as a substrate in a 4-OT L8Y/M45Y/F50A catalyzed Michael-type addition or a 4-OT M45T/F50A catalyzed aldol condensation reaction in one pot. Abbreviations: CaaD: chloroacrylic acid dehalogenase, MSAD: malonate semialdehyde decarboxylase, ScADH: alcohol dehydrogenase, PRO-NOX(009): NADH oxidase, ZmPDC: pyruvate decarboxylase.

Route I involves the dehalogenation of **8** into **9**, catalyzed by the enzyme chloroacrylic acid dehalogenase (CaaD), followed by the decarboxylation of **9** into **1** by the enzyme malonate semialdehyde decarboxylase (MSAD). Route II involves the oxidation of ethanol (**10**) into **1** catalyzed by the alcohol dehydrogenase from *Saccharomyces cerevisiae* (ScADH), in combination with the NADH oxidase, PRO-NOX(009). Route III involves the decarboxylation of pyruvate (**4**) into **1** using pyruvate decarboxylase from *Zymomonas mobilis* (ZmPDC). Route I and II proved to be effective routes for the *in situ* generation of **1** and could be used in combination with the 4-OT L8Y/M45Y/F50A catalyzed Michael-type addition of **1** to **2a** and the 4-OT M45T/F50A catalyzed aldol condensation of **1** with **12c**. These results demonstrate how modular reaction cascades can provide interesting artificial metabolic networks for the synthesis of important pharmaceuticals.

### Reaction engineering

Previously, the enzyme 4-OT M45T/F50A was identified as a promising catalyst for the aldol condensation of **1** and **12a**<sup>10</sup>. The aldol-condensation reaction comprises of two steps: the initial aldol coupling and the subsequent dehydration, which are both catalyzed by 4-OT<sup>11</sup>. Since, the aldol-coupling product contains an asymmetric C-atom, which is lost upon dehydration, in **chapter 6** we aimed to prevent dehydration and use 4-OT to synthesize chiral  $\beta$ -hydroxyaldehydes. We tested a range of aldol acceptors and found that primarily acceptors with electron withdrawing substituents were accepted as substrates by 4-OT M45T/F50A. Out of these acceptors, five yielded aldol products that did not undergo dehydration, and *in situ* reduction by NaBH<sub>4</sub> afforded 1,3-diols **13l-p** in good to excellent isolated yield (up to 92%) and in high enantiopurity (e.r. up to 99:1). To expand the scope of accessible chiral 1,3-diols, we screened through a panel of 4-OT homologues to find new promiscuous aldolase activities. The 4-OT homologue TAUT015 was identified as a promising biocatalyst and could be used for the enantioselective synthesis of 1,3-diols **13c-e** and **13g-h** (e.r. up to >99:1). As the aldol acceptors that were tested are not known to be accepted by natural aldolases, this study illustrates that members of the 4-OT superfamily, exemplified by 4-OT M45T/F50A and TAUT015, nicely complement the toolbox of biocatalysts available for asymmetric C-C bond-forming reactions.



**Figure 4.** Aldol condensations and coupling reactions catalyzed by 4-OT or TAUT015. **A)** Aldol condensation and coupling reactions catalyzed by 4-OT M45T/F50A. **B)** Aldol coupling reactions catalyzed by 4-OT homologue TAUT015.



In summary, we showcase the potential of 4-OT as a biocatalyst in new synthetic pathways towards various pharmaceuticals. Enzyme engineering afforded several mutants with improved ethanol resistance, improved enantioselectivity and improved ‘Michaelase’ activity. The engineered 4-OT mutants could be combined with other (bio) catalysts in one-pot cascade reactions providing artificial metabolic networks that can be used for the total synthesis of GABA analogues. Finally, via ‘substrate engineering’ and homologue screening, we obtained access to a range of chiral 1,3-diols via a new chemoenzymatic synthesis route.

### Future perspectives

The presented work in this thesis provides a stepping stone towards the development of novel 4-OT-based synthetic routes that could, for example, be applied for the synthesis of various GABA analogues. Significant progress has been made in terms of enantioselectivity, activity and solvent stability of 4-OT. However, industrial application of 4-OT still requires extensive engineering and reaction optimization. The activities of the engineered 4-OT variants (**chapter 2** and **chapter 4**) are still modest with a  $k_{\text{obs}}$  in the range of  $0.1 \text{ s}^{-1}$  for both the Michael-type addition of **1** to **2a-d** catalyzed by 4-OT L8Y/M45Y/F50A and the Michael-type addition of **6** to **7a** catalyzed by 4-OT S37E/M45I/F50A and 4-OT A33E/M45I/F50A. An increase in activity of at least 10-fold is desirable. The developed AICS pre-screening assay developed in **chapter 4** might be a helpful tool to achieve this. Furthermore, machine-learning algorithms are becoming a powerful tool to predict beneficiary mutations. The mutational data that has been obtained for 4-OT from mutability-landscape studies presented in this thesis, as well as in previous studies, are excellent training data for these algorithms. Several successful engineering studies have been reported using the ASRA and Innov’SAR algorithms<sup>12-14</sup> and, in collaboration with others, further engineering of 4-OT via machine-learning is ongoing.

The reaction cascades described in **chapter 2** and **chapter 5** nicely demonstrate the opportunities available for combining different catalysts in one pot, but further optimization is required. Ideally, these reaction cascades are not performed stepwise, but in continuous mode. A continuous system would allow for higher substrate loadings as the final GABA analogues are highly soluble under aqueous conditions. In addition, a continuous reaction cascade would overcome unfavorable reaction equilibria which would prevent the need of using one substrate in large excess compared to the other. However, the aldehyde dehydrogenase used in **chapter 2**, PRO-ALDH(003), could also oxidize **1** which prohibited an efficient continuous reaction cascade. Utilization of an aldehyde dehydrogenase that is selective towards  $\gamma$ -nitroaldehydes would solve this issue, but identification of an aldehyde dehydrogenase that is selective towards  $\gamma$ -nitroaldehydes

and can withstand high concentrations of **1** might be difficult. Alternatively, an aldehyde dehydrogenase could be used in combination with the 4-OT catalyzed Michael-type addition of **6** to **7a** (**chapter 4**). A stepwise reaction cascade using PRO-ALDH(003) has already been reported for this reaction<sup>15</sup>. A continuous system requires an aldehyde dehydrogenase that is selective towards **3a**, but does not oxidize **7a**. The identification of such an aldehyde dehydrogenase might be easier due to the marked electronic differences between the aldehyde functionality of **7a** compared to **3a**.

A fully continuous reaction cascade requires an alternative strategy for the reduction of the nitro group of **4a-d**, as the conditions use for chemical reduction (**chapter 2**) are not compatible with the previous enzymatic steps. Ideally, NiCl<sub>2</sub> would be replaced by a nitroreductase. However, no nitroreductases have been reported that can fully reduce nitro functionalities that are not in direct conjugation with an aromatic ring. Reported nitroreductases do not accept aliphatic nitro groups and often catalyze only partial reduction to reactive nitroso or hydroxylamine species, which results in the formation of unwanted side products<sup>16</sup>. Currently, work is ongoing in our lab to identify novel nitroreductases and explore reaction conditions that allow for the full reduction of the nitro functionalities present in **4a-d**. Our ultimate goal is to obtain a fully enzymatic system, either as whole cells or purified enzymes, which can be used to efficiently synthesize GABA analogues under environmentally friendly conditions.

## References

1. Bornscheuer, U. T., Hauer, B., Jaeger, K. E. & Schwaneberg, U. Directed Evolution Empowered Redesign of Natural Proteins for the Sustainable Production of Chemicals and Pharmaceuticals. *Angew. Chem. Int. Ed.* 58, 36–40 (2019).
2. Giger, L. et al. Evolution of a designed retro-aldolase leads to complete active site remodeling. *Nat. Chem. Biol.* 9, 494 (2013).
3. Obexer, R. et al. Emergence of a catalytic tetrad during evolution of a highly active artificial aldolase. *Nat. Chem.* 9, 50–56 (2017).
4. Kan, S. B. J., Lewis, R. D., Chen, K. & Arnold, F. H. Directed evolution of cytochrome c for carbon–silicon bond formation: Bringing silicon to life. *Science* 354, 1048– 1051 (2016).
5. Kan, S. B. J., Huang, X., Gumulya, Y., Chen, K. & Arnold, F. H. Genetically programmed chiral organoborane synthesis. *Nature* 552, 132–136 (2017).
6. Zandvoort, E., Geertsema, E. M., Baas, B.-J. J., Quax, W. J. & Poelarends, G. J. Bridging between organocatalysis and biocatalysis: asymmetric addition of acetaldehyde to  $\beta$ -nitrostyrenes catalyzed by a promiscuous proline-based tautomerase. *Angew. Chem. Int. Ed.* 51, 1240–1243 (2012).
7. Zandvoort, E., Baas, B.-J., Quax, W. J. & Poelarends, G. J. Systematic screening for catalytic promiscuity in 4-oxalocrotonate tautomerase: enamine formation and aldolase activity. *ChemBiochem* 12, 602–609 (2011).
8. van der Meer, J.-Y. et al. Using mutability landscapes of a promiscuous tautomerase to guide the engineering of enantioselective Michaelases. *Nat. Commun.* 7, 10911 (2016).
9. Acevedo-Rocha, C. G. et al. Directed Evolution of Proteins Based on Mutational Scanning. *Methods Mol. Biol.* 1685, 87–128 (2018).
10. Rahimi, M. et al. Mutations closer to the active site improve the promiscuous aldolase activity of 4-oxalocrotonate tautomerase more effectively than distant mutations. *ChemBioChem* 17, 1225–1228 (2016).
11. Zandvoort, E., Geertsema, E. M., Quax, W. J. & Poelarends, G. J. Enhancement of the Promiscuous Aldolase and Dehydration Activities of 4-Oxalocrotonate Tautomerase by Protein Engineering. *ChemBioChem* 13, 1274–1277 (2012).
12. Li, G., Dong, Y. & Reetz, M. T. Can Machine Learning Revolutionize Directed Evolution of Selective Enzymes? *Adv. Synth. Catal.* 361, 2377–2386 (2019).
13. Cadet, F. et al. A machine learning approach for reliable prediction of amino acid interactions and its application in the directed evolution of enantioselective enzymes. *Sci. Rep.* 8, 16757 (2018).
14. Feng, X., Sanchis, J., Reetz, M. T. & Rabitz, H. Enhancing the Efficiency of Directed Evolution in Focused Enzyme Libraries by the Adaptive Substituent Reordering Algorithm. *Chem. Eur. J.* 18, 5646–5654 (2012).

15. Guo, C., Saifuddin, M., Saravanan, T., Sharifi, M. & Poelarends, G. J. Biocatalytic Asymmetric Michael Additions of Nitromethane to  $\alpha,\beta$ -Unsaturated Aldehydes via Enzyme-bound Iminium Ion Intermediates. *ACS Catal.* 9, 4369–4373 (2019).
16. Roldán, M. D., Pérez-Reinado, E., Castillo, F. & Moreno-Vivián, C. Reduction of polynitroaromatic compounds: the bacterial nitroreductases. *FEMS Microbiol. Rev.* 32, 474–500 (2008).



# *Appendix*

*Nederlandse samenvatting voor  
de geïnteresseerde leek*

*Acknowledgement*

*List of publications*

---

## **Nederlandse samenvatting voor de geïnteresseerde leek**

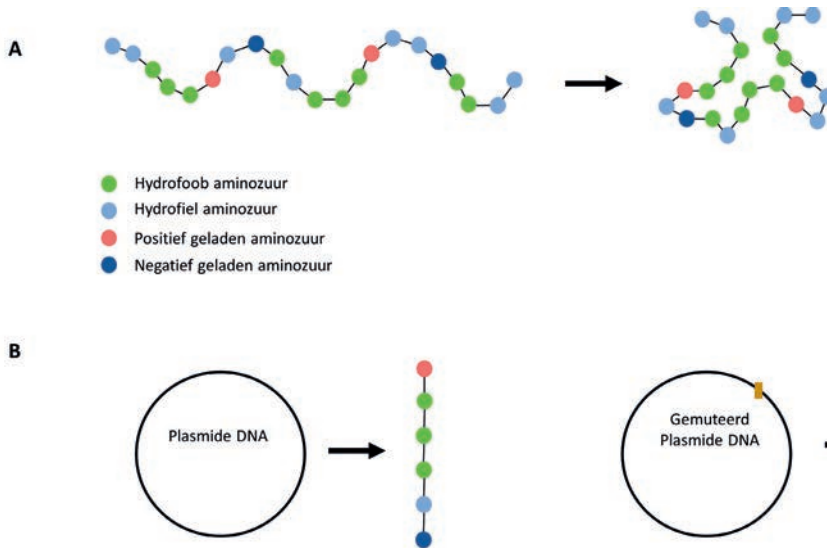
### **Enzym modificatie**

Enzymen zijn zeer complexe moleculen die bestaan uit lange ketens van aminozuren. Er zijn 20 verschillende aminozuren met allemaal een verschillend karakter. Er zijn bijvoorbeeld hydrofobe aminozuren (aminozuren die niet met water in contact willen komen) en hydrofiele aminozuren (aminozuren die juist graag met water in contact willen komen) en er zijn aminozuren met een positieve lading en aminozuren met een negatieve lading. Door deze verschillende eigenschappen is een enzym niet slechts een chaotische sliert van aminozuren, maar vouwt de sliert van aminozuren zich tot een complex driedimensionaal molecuul (Figuur 1A).

Alle levende organismen gebruiken enzymen als katalysator om chemische reacties mee te katalyseren. Een katalysator is een stof (bijvoorbeeld een enzym) die ervoor zorgt dat een chemische reactie sneller verloopt. In auto's zit bijvoorbeeld een katalysator die ervoor zorgt dat allerlei giftige gassen die uit de motor komen snel reageren tot koolstofdioxide, stikstof en water. Cellen van organismen hebben de beschikking over allerlei soorten enzymen die allemaal gespecialiseerd zijn in verschillende reacties. Hierdoor kunnen heel veel verschillende reacties plaatsvinden in een cel, die bijvoorbeeld nodig zijn voor het verteren van voedsel.

Veel wetenschappers zijn erin geïnteresseerd om enzymen te gebruiken als katalysator voor het maken van allerlei nuttige stoffen zoals medicijnen. Enzymen hebben namelijk een aantal grote voordelen ten opzichte van andere katalysatoren. Enzymen zijn bijvoorbeeld volledig hernieuwbaar en biologisch afbreekbaar (dit in tegenstelling tot de driewegkatalysator in auto's die bestaat uit zware metalen zoals platina, rodium of palladium).

Om enzymen te kunnen gebruiken in industriële processen, zoals de synthese van medicijnen, is het nodig om op grote schaal enzymen te produceren. Chemisch synthetiseren van enzymen is echter nagenoeg onmogelijk (en bovendien extreem duur) omdat enzymen zulke complexe moleculen zijn. Daarom wordt er gebruikt gemaakt van bacteriën om enzymen te produceren. Om ervoor te zorgen dat bacteriën het juiste enzym maken wordt aan het bacteriële DNA een klein stukje extra DNA toegevoegd. Dit stukje extra DNA, dat ook wel plasmide wordt genoemd, bevat het bouwplan voor het betreffende enzym, zodat de bacterie weet in welke volgorde de aminozuren aan elkaar moeten worden gekoppeld om het enzym te maken (Figuur 1B). De bacteriën kunnen eenvoudig op grote schaal worden gegroeid en geogst waarna het enzym uit de bacteriën kan worden gezuiverd.



**Figuur 1.** A) Schematische weergave van een enzym. Een enzym is een keten van aminozuren. De verschillende eigenschappen van de aminozuren zorgen ervoor dat het enzym zich vouwt tot een complexe structuur. B) Plasmide DNA bevat het bouwplan voor het enzym. Een mutatie in het DNA resulteert in een enzym met een kleine gewijzigde aminozuurvolgorde.

Enzymen hebben echter ook een aantal nadelen waarvoor oplossingen moeten worden gevonden voordat ze gebruikt kunnen worden als katalysator in een industrieel proces. Omdat enzymen vaak zeer gespecialiseerd zijn in het katalyseren van één specifieke reactie, is het niet altijd eenvoudig om een enzym te vinden dat geschikt is voor de reactie waarin bijvoorbeeld een medicijnfabrikant geïnteresseerd is. Het is daarom vaak nodig om natuurlijke enzymen te modificeren zodat ze geschikt zijn om te gebruiken in een industriële setting. Om enzymen te modificeren wordt het plasmide DNA een klein beetje gewijzigd (gemuteerd), zodat de bacterie vervolgens een enzym maakt met een klein beetje gewijzigde aminozuurvolgorde. Het is precies bekend hoe een bepaalde wijziging van het DNA zich vertaalt naar de aminozuurvolgorde van het enzym. Echter, het is veel moeilijker om te voorspellen hoe een gewijzigd enzym zich zal gedragen ten opzichte van het oorspronkelijke enzym. Daarom worden er meestal heel veel verschillende plasmiden gemaakt met allemaal een heel kleine wijziging in het DNA. Deze verschillende plasmiden worden allemaal aan aparte bacterieculturen toegevoegd waardoor er allemaal verschillende enzym varianten worden geproduceerd. Vervolgens wordt getest welke enzym variant het beste als katalysator werkt voor de reactie waarin men is geïnteresseerd. Het plasmide DNA van de beste enzym variant(en) wordt vervolgens uit de bacterie geïsoleerd en kan vervolgens opnieuw worden gemuteerd, waardoor de cyclus weer opnieuw kan beginnen. Dit proces van muteren van DNA en





selecteren van enzym varianten wordt vaak een aantal keren herhaald en kan resulteren in enzymen met zeer sterk verbeterde eigenschappen. Omdat dit proces veel lijkt op hoe in de natuur enzymen evolueren wordt dit proces vaak aangeduid als *directed evolution*.

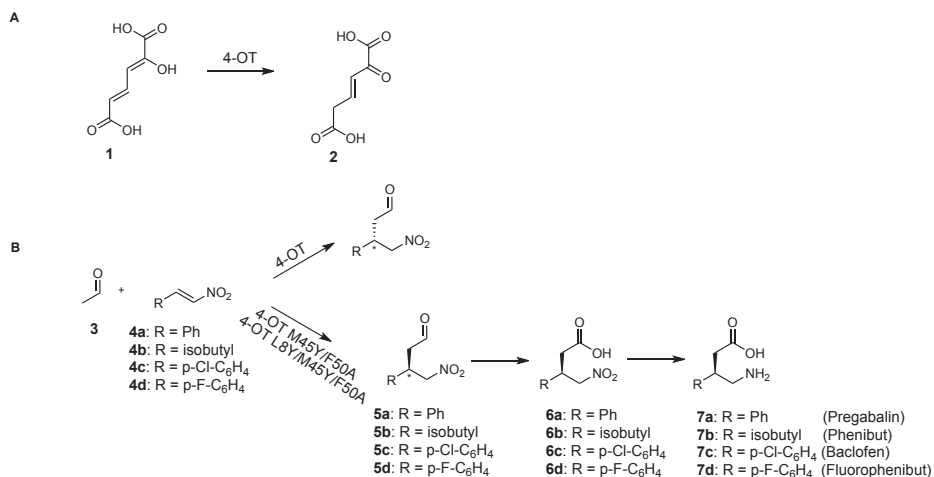
*Directed evolution* kan ook worden toegepast op enzymen die in de natuur totaal andere reacties katalyseren. Een goed voorbeeld hiervan is het enzym 4-oxalocrotonate tautomerase (4-OT) dat gevonden wordt in de bacterie *Pseudomonas putida*. De natuurlijke reactie die dit enzym katalyseert is de reactie van 2-hydroxy-2,4-hexadienedioate (**1**) naar 2-oxo-3-hexenedioate (**2**) (Figuur 2A). Wetenschappers hebben echter ontdekt dat dit enzym ook de reactie van acetaldehyde (**3**) met diverse nitroalkenen (**4a-d**) kan katalyseren. De producten van deze reactie (**5a-d**) kunnen via twee relatief eenvoudige reacties worden omgezet in een aantal belangrijke medicijnen zoals pregabalin (lyrica), phenibut en baclofen. Daarom is in deze thesis gekeken of, met behulp van *directed evolution*, 4-OT varianten konden worden gevonden die zeer goed waren in het katalyseren van de reactie van **3** met **4a-d**.

#### 4-OT modificatie

Om effectief enzym varianten te vinden met verbeterde eigenschappen is het belangrijk om via een slimme strategie het DNA te muteren. Omdat 4-OT een zeer klein enzym is van slechts 62 aminozuren, is het zeer geschikt om gebruik te maken van *mutability landscape* enzym modificatie. **Hoofdstuk 1** beschrijft een overzicht van studies die gebruik hebben gemaakt van deze techniek. Bij deze techniek wordt voor elke aminozuurpositie het oorspronkelijke aminozuur gemuteerd tot elk van de 19 andere aminozuren. In het geval van 4-OT komt dit theoretisch dus neer op 1178 ( $19 \times 62$ ) verschillende varianten. Het voordeel van deze techniek is dat het veel informatie oplevert over aminozuurposities waar mutaties een effect hebben op de eigenschappen van het enzym. Hierdoor wordt het eenvoudiger om in een volgende ronde van *directed evolution* effectief posities te kiezen om te muteren.

De stoffen **5a-d** hebben allen een asymmetrisch koolstofatoom, oftewel een koolstofatoom dat verbonden is met vier verschillende groepen (Aangeduid met [\*] in Figuur 2B). Daardoor bestaan er twee verschillende varianten van die stof die precies elkaars spiegelbeeld zijn. Deze varianten worden ook wel *enantiomeren* genoemd. Hoewel de beide *enantiomeren* van een stof dus zeer veel op elkaar lijken, kunnen ze een zeer verschillend effect hebben als medicijn. Een dramatisch voorbeeld hiervan is softenon, dat in de jaren vijftig werd voorgeschreven aan zwangere vrouwen tegen ochtendmisselijkheid. Ook de werkzame stof in softenon bevat een asymmetrisch koolstofatoom en later bleek dat het ene enantiomeer hielp tegen ochtendmisselijkheid, terwijl het andere enantiomeer

ernstige schade veroorzaakte bij het ongeboren kind. Daarom is het belangrijk om bij de synthese van stoffen met een asymmetrisch koolstofatoom ervoor te zorgen dat alleen het juiste enantiomeer wordt gemaakt. 4-OT bleek echter vrijwel uitsluitend het ‘verkeerde’ enantiomeer te maken, dat wil zeggen het enantiomeer dat uiteindelijk niet geschikt is als medicijn. In het verleden is met behulp van *mutability landscape* enzym modificatie een 4-OT variant gevonden die hoofdzakelijk het juiste enantiomeer maakt. Deze variant had een mutatie op positie 45 en positie 50 (4-OT M45Y/F50A) en kon de stoffen **5a-d** maken met een enantiomeer ratio variërend tussen de 62:38 en 96:4. Hoewel dit een grote verbetering is ten opzichte van 4-OT, is dit nog niet goed genoeg om te worden toegepast in de industrie. Daarom wordt in **hoofdstuk 2** een nieuwe ronde van *directed evolution* beschreven gericht op het verbeteren van de enantioselectiviteit. 4-OT M45Y/F50A wordt als start-variant gebruikt. Op grond van een kristalstructuur van 4-OT M45Y/F50A, waarin je in 3D kunt zien hoe het enzym eruitziet, worden twee posities uitgekozen (positie 2 en positie 8) om tegelijk gemuteerd te worden binnen 4-OT M45Y/F50A. Na het screenen van een serie van mutanten werden twee 4-OT varianten gevonden met sterk verbeterde enantioselectiviteit. Beide varianten hadden een extra mutatie op positie 8, waarbij het relatief kleine aminozuur leucine werd vervangen door het grotere aminozuur phenylalanine of tyrosine (4-OT L8F/M45Y/F50A en 4-OT L8Y/M45Y/F50A respectievelijk). 4-OT L8Y/M45Y/F50A bleek ook nog een significant snellere katalysator te zijn en bovendien stabiel in de aanwezigheid van ethanol vergeleken met start-variant 4-OT M45Y/F50A. De kristalstructuur van 4-OT L8Y/M45Y/F50A werd vervolgens opgehelderd en vergeleken met de kristalstructuur van start-variant 4-OT M45Y/F50A. Daaruit bleek dat de verbetering in enantioselectiviteit waarschijnlijk wordt veroorzaakt doordat het grote aminozuur tyrosine ervoor zorgt dat startmaterialen **4a-d** dieper in het enzym worden geduwd. Acetaldehyde (**3**) kan hierdoor slechts vanaf één kant reageren met **4a-d**, waardoor er slechts één enantiomeer wordt gemaakt.

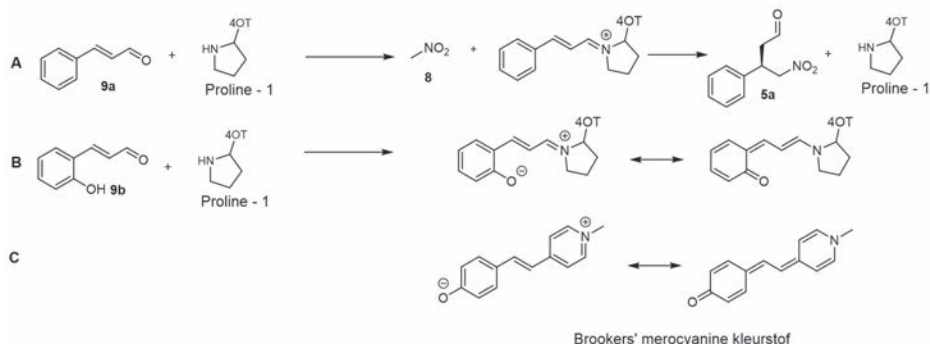


**Figuur 2.** A) Natuurlijke reactie van 4-OT. B) Onnatuurlijke reacties van **3** met stoffen **4a-d**. 4-OT maakt het ‘verkeerde’ enantiomeer, maar twee 4-OT mutanten maken het andere enantiomeer dat in 2 stappen kan worden omgezet in de stoffen **7a-d** hetgeen de werkzame stoffen in een aantal belangrijke medicijnen zijn.

De stoffen **4a-d** zijn slecht oplosbaar in water. Om toch voldoende op te kunnen lossen in water moeten er *co-solvents* aan worden toegevoegd. Dit zijn stoffen die goed oplossen in water en ervoor zorgen dat slecht oplosbare stoffen toch opgelost kunnen worden. Omdat enzymen gebouwd zijn om in water reacties te katalyseren, kan de aanwezigheid van hoge concentraties *co-solvents* enzymen sterk destabiliseren. Daarom is er in **hoofdstuk 3** gekeken naar mutaties die de stabiliteit van 4-OT in de aanwezigheid van *co-solvents* verbeteren. Er is specifiek gekeken naar de *co-solvent* ethanol, omdat ethanol een goedkope hernieuwbare *co-solvent* is en daarom erg aantrekkelijk om te gebruiken in de industrie. De katalytische activiteit van 4-OT voor de reactie van **3** met **4a** werd gemeten in de aanwezigheid van 5% en 25% ethanol. Het bleek dat 4-OT ongeveer 50% activiteit verloor bij 25% ethanol in vergelijking met 5%. Hetzelfde werd getest voor een collectie van 1040 4-OT varianten met elk één mutatie. De resultaten hiervan werden in een solvent-stabiliteit *mutability landscape* gezet. Dit is een soort kaart waaruit gemakkelijk kan worden afgelezen welke mutaties voordelig zijn en welke juist nadelig zijn. Het bleek dat mutaties op positie 30 en positie 33 veelal een positief effect hadden op de solvent stabiliteit van 4-OT. Echter, hoewel deze enkelmutanten van 4-OT stabielere waren in de aanwezigheid van ethanol, maken ze hoofdzakelijk het ‘verkeerde’ enantiomeer van **5a**. Daarom werd besloten om positie 33 te muteren in start-variant 4-OT L8F/M45Y/F50A (zie **hoofdstuk 2**). Deze mutant maakt hoofdzakelijk het juiste enantiomeer van **5a**, maar heeft een zeer lage stabiliteit in de aanwezigheid van ethanol. Het alanine-aminozuur op positie 33 werd gemuteerd tot alle 19 mogelijke andere aminozuren en

elke variant werd getest voor stabiliteit in de aanwezigheid van ethanol. De variant met een isoleucine-aminozuur op positie 33 (4-OT L8F/A33I/M45Y/F50A) bleek het meest stabiel in de aanwezigheid van ethanol. Deze variant kon gebruikt worden in de aanwezigheid van 40% ethanol, terwijl start-variant 4-OT L8F/M45Y/F50A al bij 10% ethanol geen enkele activiteit meer had.

Zoals aan het begin al werd gezegd is het heel moeilijk om vooraf te voorspellen of een bepaalde enzym-variant betere of juist minder goede eigenschappen heeft. Het is daarom vaak nodig om heel veel enzym-varianten te testen voordat er een variant wordt gevonden met de gewenste eigenschappen. Het is zelfs niet ongebruikelijk dat meer dan de helft van de geteste varianten geen enkele activiteit heeft omdat die specifieke mutatie het enzym kapot maakt. Om toch efficiënt enzymen te modificeren is het daarom noodzakelijk om technieken te ontwikkelen waarmee het mogelijk is om heel veel verschillende varianten snel te testen. In **hoofdstuk 4** wordt een nieuwe methode beschreven om op een eenvoudige wijze actieve varianten te scheiden van inactieve varianten. Het doel was om verbeterde 4-OT varianten te vinden voor de reactie van **8** met **9a** tot **5a** (Figuur 3A). Het product van deze reactie is dus hetzelfde als bij de reactie beschreven in **hoofdstuk 2** en **hoofdstuk 3**, maar er wordt gebruik gemaakt van verschillende startmaterialen. Om te begrijpen hoe de nieuwe methode werkt die in dit hoofdstuk wordt beschreven is het noodzakelijk om iets preciezer naar het *reactiemechanisme* van 4-OT te kijken. Het reactiemechanisme wil zeggen welke chemische stappen de startmaterialen ondergaan voordat ze reageren tot het uiteindelijke product. 4-OT heeft als eerste aminozuur een proline, die een heel belangrijke rol speelt in het reactiemechanisme. Deze proline kan namelijk reageren met **9a** om een *enzym-substraat* complex te vormen, een complex van het enzym met een van de startmaterialen. Dit is de eerste stap in het reactiemechanisme van 4-OT. Daarom geldt in principe dat als een 4-OT variant niet het *enzym-substraat* complex kan vormen, het geen goede katalysator kan zijn voor de reactie. In **hoofdstuk 4** wordt een assay beschreven die gebruik maakt van stof **9b**, een variant van de stof **9a**. Deze stof kan ook met 4-OT reageren tot een enzym-substraat complex en vormt dan een chemische structuur die veel lijkt op de structuur van bepaalde kleurstoffen en krijgt daarom een rode kleur (Figuur 3B). Wanneer een 4-OT variant dus geen rode kleur krijgt in de aanwezigheid van **9b**, dan betekent dit dat deze variant geen enzym-substraat complex kan vormen en daarom ook geen goede katalysator is. Met deze eenvoudige kleurreactie kan een collectie van 4-OT varianten eenvoudig worden voorgesorteerd en kunnen de inactieve varianten uit de collectie worden gefilterd. Met deze techniek, die *Activated Imminium Colony Staining* (AICS) wordt genoemd, kon tot wel 95% aan inactieve varianten uit een 4-OT mutanten collectie worden gefilterd. Na twee ronden van *directed evolution*, werden twee 4-OT varianten gevonden met een tot wel 39 keer hogere activiteit.



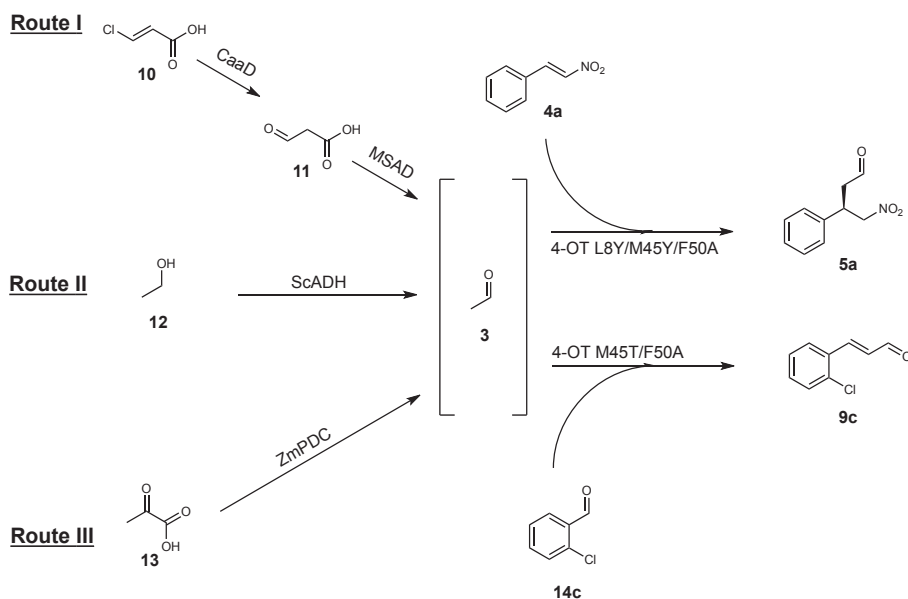
**Figuur 3.** A) Onnatuurlijke reactie van 4-OT waarbij **8** en **9a** reageren tot **5a**. De eerste stap in het reactiemechanisme is de reactie van de proline-1 van 4-OT met **9a**. Dit enzym-substraat complex reageert vervolgens met stof **8** tot stof **5a** en het vrije enzym. B) Reactie van proline-1 van 4-OT met **9b**. Het gevormde enzym-substraat complex lijkt op de structuur van bepaalde kleurstoffen en heeft daarom een rode kleur. C) De chemische structuur van de kleurstof Brookers' merocyanine.

### 4-OT in cascadereducties

We hebben hiervoor gezien dat 4-OT gebruikt kan worden als katalysator voor de synthese van **5a-d** via twee verschillende chemische routes. Echter, om de stoffen **5a-d** om te zetten in medicijnen zijn nog twee chemische stappen nodig: de reactie van **5a-d** tot **6a-d** gevolgd door de reactie tot **7a-d**, waar weer andere katalysatoren voor nodig zijn. Bij chemische syntheses wordt vaak na iedere reactiestap het product gezuiverd omdat de volgende reactiestap onder andere condities moet worden uitgevoerd. Een groot voordeel van enzymen is dat ze, hoewel ze heel verschillende reacties kunnen katalyseren, redelijk vergelijkbare reactiecondities nodig hebben. In cellen moeten alle enzymen immers ook samenwerken onder de reactiecondities van de cel. Hierdoor is het soms mogelijk om meerdere reacties achter elkaar uit te voeren zonder dat tussentijds producten moeten worden gezuiverd. Dit worden *cascadereducties* genoemd. In **hoofdstuk 2** wordt een 3-staps cascadereductie beschreven: de reactie van **3** met **4a-d** gekatalyseerd door 4-OT, gevolgd door de reactie van **5a-d** tot **6a-d** gekatalyseerd door een *aldehyde dehydrogenase* enzym en uiteindelijk de reactie van **6a-d** tot **7a-d** gekatalyseerd door het metaal *nickel chloride*. De uiteindelijke medicijnen (stoffen **7a-d**) kunnen via deze 3-staps cascadereductie worden verkregen in hoge opbrengst (tot wel 70%) en met zeer hoge enantio-zuiverheid (met een enantiomeer ratio van 99:1).

Stof **3** is een belangrijke bouwsteen die gebruikt kan worden door 4-OT in verschillende reacties. Deze stof is echter ook giftig en vrij reactief, waardoor het lastig is om reacties met deze stof op te schalen. In **hoofdstuk 5** worden drie enzymatische routes onderzocht om **3** te produceren (Figuur 4). Twee van deze routes werkten zeer goed en konden

worden gebruikt in *cascadereacties* met 4-OT. Hierdoor is het niet meer nodig om 3 apart aan reacties met 4-OT toe te voegen, maar kan het in hetzelfde reactiemengsel worden geproduceerd.

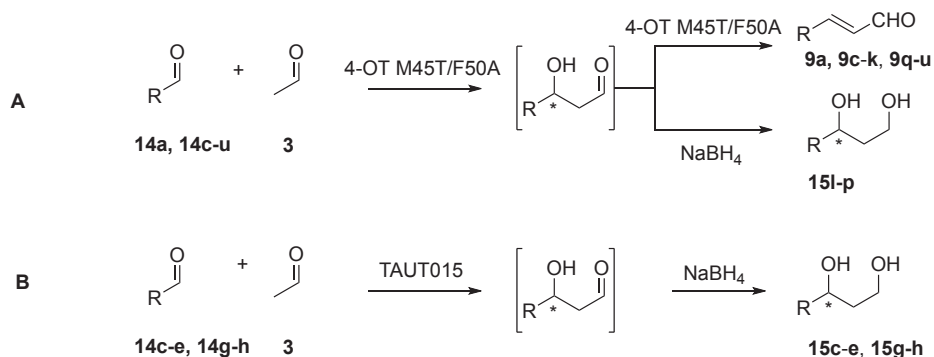


**Figuur 4.** Drie enzymatische routes om acetaldehyde (3) te synthetiseren. Het gesynthetiseerde 3 kan vervolgens door 4-OT worden gekoppeld aan 4a of 14c. Afkortingen: CaaD: chloroacrylic acid dehalogenase, MSAD: malonate semialdehyde decarboxylase, ScADH: alcohol dehydrogenase, ZmPDC: pyruvate decarboxylase.

### Reactiemodificatie

Een andere reactie die door 4-OT kan worden gekatalyseerd is de reactie van 3 met 14a tot 9a (Figuur 5). Deze reactie bestaat uit twee stappen. De eerste stap is de koppeling van stof 3 met stof 14a. Daarna volgt een dehydratie stap, waarbij water wordt geëlimineerd. Beide stappen worden door 4-OT gekatalyseerd. Hoewel stof 9a, en derivaten daarvan, van zichzelf ook interessant zijn om te maken, waren we geïnteresseerd om alleen de eerste koppelingsreactie te laten plaatsvinden en te voorkomen dat water geëlimineerd zou worden. In plaats van enzymmodificatie, zoals in **hoofdstuk 2, 3 en 4**, gebruikten we in **hoofdstuk 6** een andere strategie: *reactiemodificatie*. Bij *reactiemodificatie* is het doel om door kleine aanpassingen te maken aan één van de startmaterialen, de uitkomst van de reactie te veranderen. In dit geval werd de door 4-OT gekatalyseerde reactie van 3 met de stoffen 14c tot 14u getest om te zien of er reacties konden worden gevonden waarbij 3 wel gekoppeld zou worden, maar waarbij water niet geëlimineerd zou worden.

Bij een deel van deze derivaten vond zowel koppeling als dehydratie plaats (de derivaten **14c-k** en **14q-u**) maar de derivaten **14l-p** koppelden wel aan **3**, maar de resulterende producten werden niet gedehydrateerd (Figuur 5). Omdat deze ‘koppelingsproducten’ niet erg stabiel zijn, werd de stof  $\text{NaBH}_4$  aan de reactie toegevoegd die ervoor zorgt dat de koppelingsproducten reageren tot de stoffen **15l-p**. De stoffen **15l-p** (zogenaamde 1,3-diolen) hebben een asymmetrisch koolstofatoom. De stoffen **15l-p** die op deze manier gemaakt waren, hadden een zeer goede enantiomere verrijking (met enantiomeer ratio's tot wel 99:1). Vervolgens hebben we gekeken of we nog meer 1,3-diolen konden maken, door een panel van 4-OT homologen te testen als katalysator. Een 4-OT homologoog is een variant van 4-OT die in andere organismen wordt gevonden. Naast *Pseudomonas putida* zijn er namelijk veel meer organismen die een 4-OT enzym hebben, maar meestal met een enigszins gewijzigde aminozuurvolgorde. Uit een panel van vijftig 4-OT homologen vonden we dat de 4-OT homologoog TAUT015 goed gebruikt kon worden als katalysator voor de synthese van de stoffen **15c-e** en **15g-h**. Ook in dit geval konden de stoffen **15c-e** en **15g-h** met een zeer goede enantiomere verrijking worden verkregen (met enantiomeer ratio's tot wel meer dan 99:1).



R = Ph (**14a**), 2-Cl-Ph (**14c**), 3-Cl-Ph (**14d**), 2,4-di-Cl-Ph (**14e**), 3,4-di-Cl-Ph (**14f**), 2-NO<sub>2</sub>-Ph (**14g**), 3-NO<sub>2</sub>-Ph (**14h**), 2-F-Ph (**14i**), 2-Br-Ph (**14j**), tetrafluoro-Ph (**14k**), 4-CN-Ph (**14l**), 4-CF<sub>3</sub>-Ph (**14m**), pentafluoro-Ph (**14n**), 3-CN-4-F-Ph (**14o**), furanyl (**14p**), 4-Cl-Ph (**14q**), 4-F-Ph (**14r**), 4-NO<sub>2</sub>-Ph (**14s**), Pyridyl (**14t**), 4-Br-Ph (**14u**)

**Figuur 5.** A) 4-OT gekatalyseerde reactie van **3** met **14a**, **14c-u**. Een deel reageert tot het gedehydrateerde product (**9a**, **9c-k**, **9q-u**) en een deel reageert tot 1,3-diolen na toevoeging van  $\text{NaBH}_4$  (**15l-p**). B) De reactie van **3** met **14c-e**, **14g-h** gekatalyseerd door de 4-OT homologoog TAUT015. In dit geval reageert alles tot 1,3-diolen na toevoeging van  $\text{NaBH}_4$ .

Samengevat laat deze thesis het potentieel van 4-OT zien als katalysator voor verschillende koolstof-koolstofbindingvormende reacties. Nieuwe 4-OT varianten zijn ontdekt met een zeer goede enantioselectiviteit en solvent-stabiliteit. Gecombineerd met andere (bio)katalysatoren konden deze varianten worden toegepast in cascadereducties waardoor een serie van farmaceutisch relevante GABA-derivaten zonder tussentijdse zuivering konden worden gesynthetiseerd vanuit zeer simpele bouwstoffen. Verder laten we zien dat 4-OT kan worden toegepast in een nieuwe syntheseroute voor 1,3-dienen, belangrijke bouwstenen voor de productie van medicijnen.



## **Acknowledgement**

As I am writing this, we are in the middle of the Corona-lockdown. Everyone is urged to work from home, and refrain from social contacts. Public gatherings are not allowed and hence all public PhD defences have been postponed until further notice. I am also not certain if my initially planned defence can take place, or if it will be postponed. PhD theses can only be defended via a (public) livestream, which is, in my opinion, a relatively poor substitute. A PhD is often just as much a team effort as it is an individual achievement and therefore it is much preferred to defend a thesis in the (physical) presence of the many people that helped reaching this goal. Looking back on my PhD journey, I feel very blessed for all the help I received from so many people. Therefore, I would like to take this opportunity to express my gratitude to everyone who contributed to this thesis.

Firstly, I would like to thank my primary supervisor Prof. Dr. **Gerrit J. Poelarends**. Gerrit, thank you so much for providing me with the opportunity to perform my PhD research in your group. Already from the beginning, I appreciated your effective communication style, promptly answering E-mails and always willing to help. I also really appreciated your feedback style. Rather than simply correcting mistakes, you took the time to explain why particular parts should be corrected, which greatly helped me to improve my skills. The fair way you try to treat all members of the group helped to provide for a nice working environment that I really enjoyed.

I would like to thank my second supervisor Prof. Dr. **Wim J. Quax**. Although you were less directly involved in my research, I really appreciated the supporting questions during our work discussions and the way you managed the whole department.

I would like to thank the members of the reading committee, Prof. dr. ir. **M. W. Fraaije**, Prof. dr. **A. K. H. Hirsch** and Prof. dr. **U. Hanefeld** for taking the time to read and evaluate my thesis.

I would like to express my gratitude to all members of the CarbaZymes Consortium. Thanks to Prof. Dr. **W. D. Fessner** and **A. Theodoridou** for leading the consortium and taking care of many administrative tasks. Prof. **S. J. Charnock** and Dr **Y. Qi**, thank you for your efforts in the fruitful collaboration we had on the cascade synthesis of  $\gamma$ -aminobutyric acids described in chapter 2. In addition, I would like to thank Dr. **P. Domingues de Maria** for his valuable advice concerning the patent application based on this work.

My sincere gratitude to Dr. **Andy-Mark W. H. Thunnissen** and Dr. **Tjaard Pijning** who were involved in the crystallisation of the 4-OT mutant described in chapter 2.

Thanks to the technicians in our group, **Ronald, Rita** and **Pieter**. Thank you for all the things you took care of to keep the lab organized and functioning. You made certain that I could fully focus on my own PhD project and that I did not have to worry about so many other things. **Ronald**, thank you for helping me with the pipetting robot, saving me a lot of manual pipetting labour. **Rita**, thank you for taking care of all my sequencing samples (which amounted to almost 600 samples for my entire PhD). **Pieter**, thank you for teaching me many organic synthesis techniques and helping me, as a biologist, to feel comfortable in the chemistry lab. I often think back on the *wie is de mol* discussions we had with a broad smile.

Many thanks to the master students I supervised **Max Lubberink** and **Fabiola Zuchi**. **Max**, thank you for all your hard work, which eventually nicely fitted in chapter 3. I am happy that you found a PhD position in Manchester. **Fabiola**, thank you for your efforts as a master student. I was highly impressed by your meticulous way of performing research. I am happy you could continue your research as a PhD student and I am sure you will defend your PhD in Groningen one day as well.

My sincere thanks to all (former) members of the CFB group. **Jan-Ytzen**, thank you for introducing me with 4-OT. Much of what has been researched in this thesis is based on knowledge from former 4-OT researchers like yourself. **Ykelien**, thank you for all the nice lunches we had, discussing important, but primarily not-so important topics. I seldomly met someone with more passion for *wie is de mol* which surely invigorated me. I suppose I will never become such a good mole-hunter as you are. **Chao**, you were with me on the same CarbaZymes train. Thank you so much for the nice collaborations we had and the fun time. I still cherish many atypical and curious anecdotes. Thank you **Saravanan** for our pleasant collaboration, which eventually resulted in a nice publication. I have learned a lot from you, with respect to science and Indian culture. Also, a big thanks to **Saif**. We had a very fruitful collaboration and I have learned a lot from all your chemistry knowledge. **Jielin**, you were the only girl in our office for such a long time, keeping us, boys, in check. Thank you for the very nice time we had. I really enjoyed being your paranymp. **Haigen**, thanks for the good time we had. I have a very high opinion of your research qualities and I have learned a lot from you. I am happy that I could witness your defence as paranymp. **Jan**, thank you for the nice lunches we had, discussing things in Dutch, in English and discussing Dutch in English. When you feel eloquent enough, we might consider discussing English in Dutch. Also, thank you

## *Acknowledgement*

for being my paranymph. Thanks to my former Indonesian colleagues **Tjie** and **Zainal**. **Tjie**, I am very happy you managed to defend your PhD and that I could witness this as your paranymph. **Zainal**, I am sure you will be able to defend your thesis soon as well.

Many thanks to the new '4-OT boys' **Andreas**, **Michele** and **Guangcai**. I am sure you will teach this small, awesome enzyme some new amazing tricks. **Michele**, thank you for all your efforts as my paranymph. Also, my thanks to the 'Oxytrain girls' **Marie** and **Eleonora**. I think you both have contributed a lot to the very nice atmosphere in the group. My thanks to **Alex** and **Faizan**. I wish you all the best for your PhD research.

Thanks to Prof. Dr. **Frank J. Dekker**, Prof. **Hidde J. Haisma** and Dr. **Robbert H. Cool** for insightful discussions that helped me to become a better scientist. Thank you, **Janita** and **Yvonne** for your support with all of the paperwork. A big thanks to all other former and current members of CFB.

Met het afronden van dit proefschrift sluit ik ook mijn periode in Groningen af. Naast mijn directe collega's heb ik hier ook veel andere goede vrienden ontmoet. Juist omdat veel van het onderzoek dat ik heb gedaan best lastig te begrijpen is voor veel niet-onderzoekers, bestaat het gevaar dat je jezelf opsluit in een kleine onderzoekswereld. Ik ben ontzettend dankbaar voor veel vrienden die, juist als het onderzoek misschien weleens even tegen zat, mij konden opvangen.

Alle leden van de eetgroep, **Anneke**, **Annet**, **Gerd**, **Hans**, **Klaas** en **Theo**, dank jullie wel voor de fijne eetavonden op woensdag. Ik heb me er altijd erg thuis gevoeld, zeker in het begin toen ik nog niet zoveel mensen kende in Groningen. Ik heb altijd zeer genoten van jullie baksels en culinaire hoogstandjes. Ik hoop dat jullie ook hebben genoten van de maaltijden bij mij thuis, opgepropt in mijn kleine huisje aan de Timpweg waarvan de deurbel het soms deed maar meestal niet. Bedankt **Klaas** en **Anneke** dat jullie mij hebben geïntroduceerd binnen de eetgroep.

Aan alle mensen die betrokken waren bij het +FM-jeugdwerk dat ik vier jaar heb mogen doen, heel erg bedankt! **Dorothea**, dank je wel dat je twee jaar mijn medeleider wilde zijn. **Klaas** en **Yvonne**, dank jullie wel dat jullie samen met mij de andere twee jaar hebben willen leiden. Ik ben redelijk onervaren aan het jeugdwerk begonnen, maar dankzij jullie heb ik hier nieuwe vaardigheden mogen leren, hetgeen mij ook de ruimte bood om even van het wetenschapswereldje weg te stappen. **Amelia**, **Anne**, **Birgit**, **David**, **Jesse**, **Julia**, **Loïs**, **Maarten**, **Maartje**, **Rian**, **Roel**, **Talitha**, **Thomas** en **Wilco**, bedankt voor jullie bijdragen aan de gesprekken die we mochten hebben. Ik bewaar nog steeds goede en

mooie herinneringen aan de +FM-events die we samen hebben meegemaakt. Ga met God en Hij zal met je meegaan, waar je ook naartoe gaat en welke beslissingen je ook neemt.

Dan wil ik ook nog alle mensen van de Oosterkerk heel hartelijk bedanken. Toen ik in Groningen kwam wonen, kende ik daar niemand. Maar ik heb bij jullie een fijne plek gevonden. Daarbij wil ik in het bijzonder de leden van **kring 2d** heel erg bedanken. Ik heb erg genoten van onze kringbijeenkomsten en heb daar veel mooie mensen mogen ontmoeten.

Ook wil ik mijn vader, broer en zusje heel erg bedanken. Kort nadat ik begonnen was aan mijn PhD, in September 2015, overleed mijn moeder. We hebben haar mogen begraven in de wetenschap dat we elkaar na dit leven terug zullen zien. In deze tijd zijn jullie een enorme steun voor mij geweest. Ik ben er altijd zeer dankbaar voor geweest dat we een zeer hechte familieband mogen hebben en dat er geen conflicten of spanningen zijn. Ik heb dan ook regelmatig de lange treinreis van Groningen naar Barneveld (en weer terug) gemaakt. **Papa**, bedankt voor de interesse in mijn werk, voor zoveel praktische hulp en dat ik altijd welkom ben thuis thuis. **Maaike** en **Pieter**, bedankt voor de fijne sfeer die er altijd mocht zijn. Ik ben me bewust van de uitzonderlijke situatie dat ik een broer en een zusje heb die allebei affiniteit met wetenschap hebben. Hierdoor kon ik de mooie en lastige aspecten met jullie delen. Heel hartelijk dank daarvoor.

Ten slotte, terugkijkend naar de afgelopen vijf jaar is er één person die nooit van mijn zijde is geweest. Is er één persoon die echt een doel en richting heeft gegeven. Die mij bewaard heeft en die mij zal bewaren, dwars door het onbekende van de toekomst. *Aan hem die door de kracht die in ons werkt bij machte is oneindig veel meer te doen dan wij vragen of denken, aan hem komt de eer toe, [...] tot in alle eeuwigheid.* (Ef. 3:21)

Lieuwe Biewenga



April 2020  
Veldhoven

## List of publications

- 2016 van der Meer J.Y., **Biewenga L.**, and Poelarends G.J., The Generation and Exploitation of Protein Mutability Landscapes for Enzyme Engineering. *ChemBioChem* (2016), 17,19: 1792-1799, 10.1002/cbic.201600382
- 2019 **Biewenga L**†, Saravanan T†, Kunzendorf A., van der Meer J.Y., Pijning T., Tepper P.G., van Merkerk R., Charnock S.J., Thunnissen A.M.W.H. and Poelarends G.J., Enantioselective Synthesis of Pharmaceutically Active  $\gamma$ -Aminobutyric Acids Using a Tailor-made Artificial Michaelase in One-pot Cascade Reactions. *ACS catalysis* (2019), 9,2. 1503-1513, 10.1021/acscatal.8b04299
- 2019 Guo C†, **Biewenga L**†, Lubberink M., van der Meer J.Y., van Merkerk R., Poelarends G.J., Tuning Enzyme Activity for Nonaqueous Solvents: Engineering of an Enantioselective ‘Michaelase’ for Catalysis in High Concentrations of Ethanol *ChemBioChem* (2019), 10.1002/cbic.201900721
- 2019 **Biewenga L**†, Kunzendorf A†, Poelarends G.J., In Situ Acetaldehyde Synthesis for Carbologation Reactions. *ChemBioChem* (2019), 10.1002/cbic.201900666
- 2020 **Biewenga L.**, Crotti M., Mohammad S., Poelarends G.J., A Selective Colorimetric “Turn-on” Probe for Efficient Engineering of Iminium Biocatalysis. *ACS Omega* (2020), 5, 5, 2397-2405, 10.1021/acsomega.9b03849
- 2020 Mohammad S., Guo C., **Biewenga L.**, Saravanan T., Charnock S.J., Poelarends G.J., Enantioselective Aldol Addition of Acetaldehyde to Aromatic Aldehydes Catalyzed by Proline-based Carbologases. *ACS catalysis* (2020), 10,4. 2522-2527, 10.1021/acscatal.0c00039

### Before Groningen:

- 2020 Hartman R., **Biewenga L.**, Munson-McGee J., Refai M., Boyd E.S., Bothner B., Lawrence M., Young M., Discovery and Characterization of Thermoproteus Spherical Piliferous Virus 1: a Spherical Archaeal Virus Decorated with Unusual Filaments. *Journal of Virology* (2020), 94 (11) e00036-20, 10.1128/JVI.00036-20

† These authors contributed equally to this work