TECHNISCHE UNIVERSITÄT DARMSTADT

„Exploration of Metagenomic Racemases for Biosynthetic Cascades“

Vom Fachbereich Chemie
der Technischen Universität Darmstadt

zur Erlangung des angestrebten akademischen Grades
Doctor rerum naturalium
(*Dr. rer. nat.*)
genehmigte

Dissertation

M. Sc. Silvan Wolf Poschenrieder

geboren in München

Erstgutachter: Prof. Dr. Wolf-Dieter Fessner

Zweitgutachter: Prof. Dr. Harald Kolmar

Darmstadt, 2024

Poschenrieder, Silvan Wolf: Exploration of Metagenomic Racemases for Biosynthetic Cascades
Darmstadt, Technische Universität Darmstadt,
Jahr der Veröffentlichung der Dissertation auf TUprints: 2025
Tag der mündlichen Prüfung: 28.10.2024
Tag der Einreichung: 16.07.2024
URN: urn:nbn:de:tuda-tuprints-290464
DOI: https://www.doi.org/10.26083/tuprints-00029046

Die experimentellen Arbeiten der vorliegenden Dissertation wurden unter der Leitung von Prof. Dr. Wolf-Dieter Fessner am Clemens-Schöpf-Institut für Organische Chemie und Biochemie der TU-Darmstadt im Zeitraum vom Juli 2019 bis Dezember 2022 durchgeführt.

Erklärungen

§8 Abs. 1 lit. c der Promotionsordnung der TU Darmstadt

Ich versichere hiermit, dass die elektronische Version meiner Dissertation mit der schriftlichen Version übereinstimmt und für die Durchführung des Promotionsverfahrens vorliegt.

§8 Abs. 1 lit. d der Promotionsordnung der TU Darmstadt

Ich versichere hiermit, dass zu einem vorherigen Zeitpunkt noch keine Promotion versucht wurde und zu keinem früheren Zeitpunkt an einer in- oder ausländischen Hochschule eingereicht wurde. In diesem Fall sind nähere Angaben über Zeitpunkt, Hochschule, Dissertationsthema und Ergebnis dieses Versuchs mitzuteilen.

§9 Abs. 1 der Promotionsordnung der TU Darmstadt

Ich versichere hiermit, dass die vorliegende Dissertation selbstständig und nur unter Verwendung der angegebenen Quellen verfasst wurde.

§9 Abs. 2 der Promotionsordnung der TU Darmstadt

Die Arbeit hat bisher noch nicht zu Prüfungszwecken gedient.

Darmstadt, den …….

(Silvan Poschenrieder)

Danksagung

Frei nach Arnold Schwarzenegger's: „Never call me a self-made man.“, möchte ich diese Gelegenheit nutzen um allen zu danken, die mir geholfen haben der Chemiker zu werden, der ich heute bin.

Bezüglich der Doktorarbeit, möchte ich an erster Stelle Prof. Wolf-Dieter Fessner danken, der mir die Möglichkeit gegeben hat, in seiner Arbeitsgruppe einer forschenden Tätigkeit nachzugehen und so die Grundlage für diese Arbeit zu schaffen.

Ich danke Prozomix, insbesondere Dr. Samantha Gittings und Prof. Simon Charnock, die mir eine Racemase Bibliothek zur Verfügung gestellt haben. Danke auch dafür, dass ihr vielen meiner Wünsche, bezüglich der Expression von Enzymen nachgekommen seid und mich mit fachlichem Rat unterstützt habt.

Ohne die tatkräftige Unterstützung all meiner Kollegen: Bettina, Michael, Yuriy, Anna, Alex, Fabian, Sjoerd, Inès, Tobi, Michel und Alex, wäre ich jedoch nicht weit gekommen. Vielen Dank, dass ihr mit mir die Infrastruktur am Laufen gehalten und die vielseitigen Aufgaben mit mir bewältigt habt.

Außerdem möchte ich mich bei der NMR-Abteilung, für die Messung meiner Proben und bei Dr. Schiesser, für die chirale Analytik mittels GC/MS, bedanken.

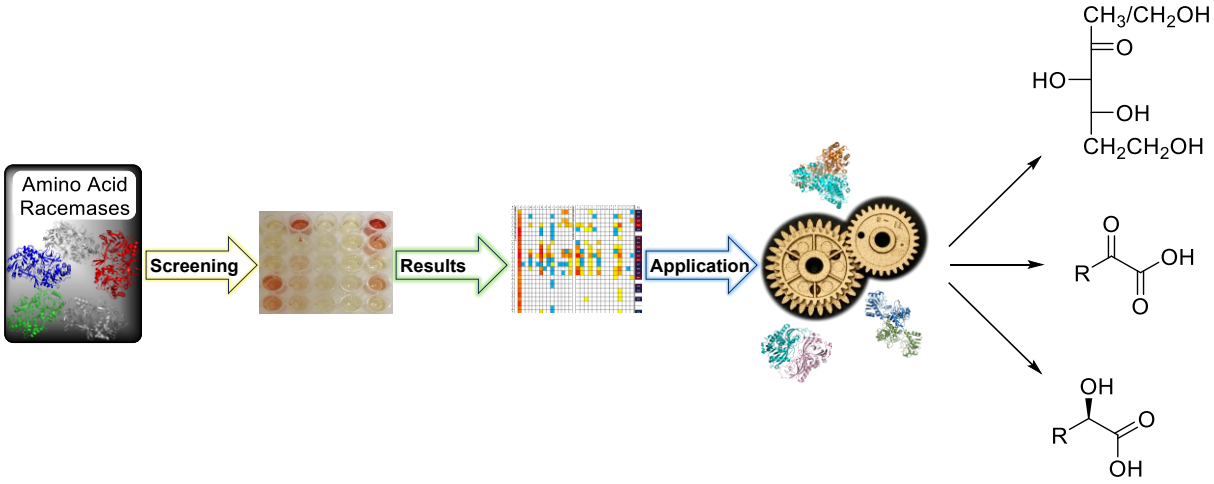
Ich möchte auch nochmal all jenen Danke sagen, die mich fachlich auf meinem langen Weg unterstützt und begleitet haben, insbesondere Prof. Kroutil, Dr. Christoph Winkler, Dr. Michael Fuchs und der gesamten Elk-Crew, Prof. Breinbaur und Dr. Patrik Dobrounig von der TU-Graz, Dr. Christoph Kring von der Uni Heidelberg und all jenen, die mir das Wissen und die Techniken anvertraut haben, die mich heute als Chemiker ausmachen. Ich werden versuchen davon etwas weiterzugeben.

Ganz besonderer Dank gilt denen, die mir persönlich und moralisch den Rückhalt gegeben haben, den man braucht um dicke Bretter zu bohren, meiner Frau Hanna, meinen Eltern Bianka und Wolfgang, meinem Bruder Pascal und meiner Oma Käthe. Ohne euch hätte ich diesen Berg nie erklimmen können.

Abschließen und in meine Arbeit überleiten, möchte ich mit einem Zitat von Louis Pasteur (1822 - 1895): „Ich will Ihnen das Geheimnis verraten, das mich zum Ziel geführt hat. Meine Stärke liegt einzig und allein in meiner Beharrlichkeit.“

Table of Contents

Zusammenfassung in Deutscher Sprache



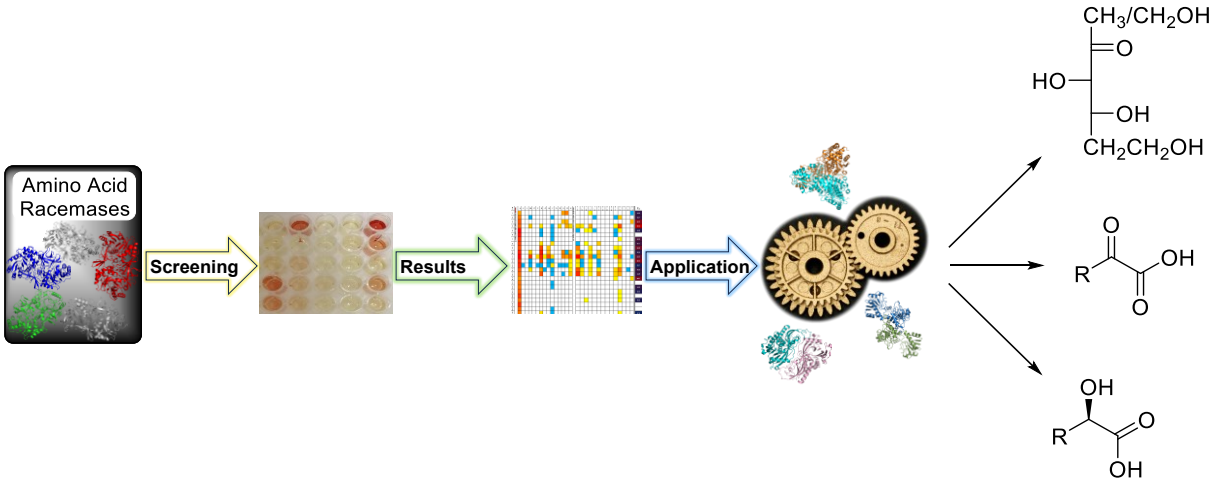
Die Verwendung von umweltverträglicheren Reagenzien aus erneuerbaren Rohstoffen, wie zum Beispiel Aminosäuren, gewinnt in der derzeitigen Transformation der chemischen Industrie immer mehr an Bedeutung. Denselben Bedeutungszuwachs erfährt auch die Biokatalyse, da Enzyme in der Lage sind solche Materialien stereoselektiv umzusetzen, oft sogar unter dynamischer kinetischer Racematspaltung. Diese Form der Racematspaltung wird als ein Schlüsselfaktor für ökonomisch sinnvolle stereoselektive Produktionsprozesse angesehen. So wird ein endloser Bedarf an neuen Enzymkatalysatoren geschaffen. Eine Enzymklasse, mit der man dynamische kinetische Racematspaltungen etablieren kann, ist die der Aminosäureracemasen. Diese wirken jedoch meistens selektiv auf nur wenige verschiedene Aminosäuren, was ihre Anwendbarkeit begrenzt. Um dieses Anwendungsgebiet zu erweitern, ist diese Arbeit dem Auffinden neuer interessanter Aminosäureracemasen innerhalb einer Enzymbibliothek gewidmet, welche von unserem Industriepartner Prozomix Ltd. zur Verfügung gestellt wurde. Sie entstand durch Similaritätssuche in der metagenomischen Prozomix-Datenbank. Das Ziel dieser Doktorarbeit war es also, Racemasen mit breitem oder ungewöhnlichem Substratspektrum zu finden und deren Potential in biokatalytischen Eintopf-Kaskadenreaktionen unter Beweis zu stellen.

Dazu wurde im ersten Schritt ein spektroskopischer Assay adaptiert. Dieser war die Grundlage um die 56 Bibliotheksenzyme auf ihre Fähigkeit zu prüfen, 16 kanonische und 14 ausgewählte nicht-kanonische Aminosäuren umzusetzen. Für die Bewertung der Enzymstabilität wurden die Enzymschmelzpunkte mittels nanoDSF bestimmt. Nachfolgend wurden die besten Racemasen in mehreren biokatalytischen

Kaskadenreaktionen eingesetzt: (i) Um die seltenen Zucker 5-Desoxy- und 1,5-Didesoxyfruktose herzustellen, wurden Racemasen mit D-Aminosäureoxidase und Transketolase kombiniert und diese Kaskadenreaktion mit einer alternativen Racemase/Transaminase/Transketolase-Kombination verglichen. (ii) Durch Kombination von Racemasen mit Oxidase, bzw. mit Oxidase und Dehydrogenase wurden wertvolle 2-Keto- und D-2-Hydroxysäuren hergestellt. Zusätzlich wurden neue Techniken entwickelt, um einen entscheidenden Teilprozess solcher Kaskadenreaktionen im Labor zu verbessern: die Versorgung der Oxidase mit Sauerstoff. Benötigte Substrate für das Screening und Cosubstrate für die Transketolase wurden unter Optimierung existierender Routen und unter Entwicklung neuer, verbesserter Synthesewege und Methoden hergestellt.

Mehrere Racemasen mit bemerkenswertem Substratspektrum wurden entdeckt, zum Beispiel Enzyme, die Tyrosin racemisieren, oder das Dipeptid β-Aspartam akzeptieren. Einige der Bibliotheksenzyme haben ein breites Substratspektrum und eine hohe Thermostabilität von T_m > 69 °C. Das Potential der Enzyme wurde demonstriert, indem erfolgreich seltene Zucker- und Aminosäurederivate hergestellt wurden. Dabei wurden Produktivitäten von bis zu 15.3 g D-Phenylmilchsäure pro Gramm zell-freiem Racemase Extrakt erreicht. Die Zielsetzung, neue taugliche Racemasen für dynamische kinetische Racematspaltungen zu finden, wurde erreicht und einige sind sogar vielversprechende Startpunkte für Enzyme Engineering.

Abstract



Biocatalysis becomes more and more appealing to industry due to its inherent ability to address contemporary demands for ecological, stereoselective synthesis from benign and biologically sourced materials, like amino acids. This creates an endless desire for new, diverse enzyme catalysts. One of the advantages of biocatalysis is to inherently provide catalysts for dynamic kinetic resolution like amino acid racemases, which is key to economic stereoselective synthesis. Drawback of most known amino acid racemases is their specificity for only few amino acids. Therefore, this work is dedicated to finding new and unusual amino acid racemases within a library that was created by our industry partner Prozomix Ltd. from their metagenomic database. Besides discovering enzymes with broader, or rare reactivities, the goal was also to demonstrate their synthetic potential in biocatalytic one-pot sequential cascade reactions.

First step was to adapt a spectrophotometric assay to enable efficient probing of the 56 library enzymes for activity on 16 canonical and 14 selected non-canonical amino acids. In addition, the enzymes' melting points were measured with nanoDSF to assess their thermal stability. Subsequently, biocatalytic reaction cascades were implemented. For the production of the rare sugars 5-deoxy- and 1,5-dideoxyfructose, the detected racemase-hits were combined with D-amino acid oxidase and transketolase, and comparing this oxidase pathway with an alternative transaminase pathway. Valuable 2-keto acids, and D-2-hydroxy acids were produced, employing selected racemases together with oxidase and stereoselective dehydrogenase enzymes. Moreover, new techniques were developed to improve one of these cascades' critical sub-process, the provision of oxygen to the oxidase at laboratory-scales. Required substrates for

screening and sugar-producing cascades were synthesized, optimizing existing protocols and developing new, improved routes and methodologies.

Several racemases with remarkable substrate ranges were discovered, for example enzymes racemising tyrosine, or acting on the dipeptide β-aspartame. Some of the panel enzymes show both, broad substrate range and high thermostability of T_m > 69 °C. Their synthetic potential was successfully demonstrated by producing rare sugars and amino acid derivatives with productivities of up to 15.3 g D-phenyllactic acid per gram of cell-free racemase extract. Thus, the research goal to discover useful, innovative racemases, fit for synthetic application in dynamic kinetic resolutions was achieved and these novel enzymes may even serve as starting points for enzyme engineering.

1 Introduction

1.1 Biocataysis, Chirality, and Industry Perspectives

Following the thalidomide scandal, in 1962, US congress passed the Kefauver-Harris amendment to the federal food, drug, and cosmetic act, defining extended requirements for the admission of drugs. According to the amendment, new drugs not only had to be proven to be safe, but now also had to be proven effective for the product's intended use. The scandal around thalidomide strengthened the FDA and put the importance of stereocentres in the safety and efficacy of substances in the limelight. Only 30 years later, in 1992, the FDA finally published a new guideline for the development of new stereoisomeric drugs, acknowledging that two biologically distinguishable stereoisomers are basically two different drugs. [1]

Before that, most drugs were chemically produced as racemates. These guidelines today are regarded as the dawn of a new paradigm and a revived interest in stereospecific synthesis. Not only did it lead to the development of new synthetic methods, but it also defined a new convincing argument for the development of biocatalytic synthesis, a relatively new and often underestimated field of chemistry. One could claim that biosynthesis is stereospecific by nature.

With the increasing acuteness of the climate crisis, rising costs of mineral oil and CO_2-taxation on the horizon, biocatalytic synthesis gains another auxiliary: the urge for natural resources. Not only is the production of biocatalysts based on natural resources, but these are also the predetermined substrates for biocatalysts.

From an industrial point of view, biocatalysis can bring another economical advantage: a substance produced biocatalytically may be termed biologic, organic, or bio-based. Even if it cannot be differentiated from a product produced with traditional organic synthesis, it now may be marketed as "bio" and has the potential to achieve much higher prices (*e.g.* raspberry ketone). Furthermore, a new way to produce a substance can be secured by the producer as new intellectual property (IP), which forms the fundamental basis for financial success of chemical companies.

So, there are a lot of advantages to promote the development of biocatalysis and in fact, the field is developing rapidly. More and more enzyme classes become commercially available and also the biocatalytic toolbox is rapidly expanding and

improving, for example through the methods of directed enzyme evolution in vitro, as demonstrated by Arnold and others leading the way to improved biocatalysts for industrial processes.

One big asset of biocatalysis is the possibility to create artificial biocatalytic cascade reactions, exploiting the compatibility of enzyme catalysts. This way, multistep reactions can be designed, converting a cheap biological feedstock into a valuable fine chemical without the need of isolation and purification of reaction intermediated. Such cascade reactions will be explained in greater detail and applying these techniques is a core piece of this work.

In the past, there were a lot of drawbacks when it came to employing enzymes as catalysts. For example, difficulties in finding specific enzymes for certain transformations or that found enzymes had limitations in applicable substrate concentrations, stability, or turnover. However today, many tools are available to speed up enzyme discovery and evaluation, like genome mining from natural DNA samples, high throughput screening platforms, or even AI-based *a-priori* design of enzyme catalysts. Also, the growing general knowledge about coenzyme handling and generation facilitates the application of enzymes.

Focus of this work is the evaluation of novel amino acid racemases.

While racemases on their own produce racemates from enantiopure substances, thus greatly decrease the value of single enantiomer substrates, in combination with other enzymes or processes, they are a valuable tool for dynamic kinetic resolution. Dynamic kinetic resolution combines two reaction steps, one a racemisation and the other one driving the equilibrium of this racemisation step completely to one side, ideally forming 100% of an enantiopure product from a racemic educt.

This has been demonstrated in racemases for example by Yu *et al*. in the production of *S*-propyl-L-cysteine from D-serine (Figure 1). [2]

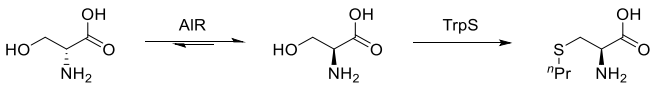


Figure 1: Dynamic kinetic resolution of D-serine in the course of *S*-propyl-L-cysteine, employing alanine racemase (AIR) and tryptophan synthase (TrpS). [2]

Hence, the enzyme class has the potential to make the cheaper enantiomer, or even racemates applicable in cases, where beforehand, only the more expensive

enantiomer could be used. This makes racemases broadly applicable in biocatalytic processes.

Enzymes, interacting with amino acids have in general great potential for application, because L-and D-amino acids are important building blocks in synthetic chemistry. Both enantiomers of amino acids occur as basic structural elements in the quickly extending class of peptide pharmaceuticals, for example in the antidiabetic drug Semaglutid, or the chemotherapy medication Actinomycin D. Additionally, natural amino acids and their derivatives are common building blocks in synthetic chemistry due to their availability from cheap biological feedstocks.

Non-canonical and non- natural amino acids often exhibit surprising compatibility with the existing enzymatic toolbox, even though sometimes containing rare functionalities, often foreign to nature. As such, they have been established for example as biological probes.

With the ever-increasing importance of stereoselective synthesis and individually tailored drug therapies, these building blocks do not only hold great value due to their contained stereocentre (α-amine), but also due to the simplicity in controlling the stereocentre with kinetic resolution, exploiting the existing biocatalytic toolbox.

In this context, amino acid racemases incorporate nicely into the existing toolbox, enabling the kinetic resolution of racemic amino acids to be performed in a dynamic fashion, increasing yields and reducing side products and waste. It has been found that contrary to wide belief, there are amino acid racemases available in nature, acting on diverse amino acid substrates, not just single substrate enzymes, like alanine, or serine racemases. Additionally thermostable amino acid racemases can be found, as potent catalytic tools and good starting points for enzyme engineering.

2 Theoretical Background

The following chapter provides an accessible entry into the theoretical background, initially recapitulating the basis for understanding amino acids and enzymes. Then the current knowledge about the employed enzyme classes is summarised. Finally, the importance of biocatalytic cascades for industry application of biocatalysts is pointed out.

2.1 Amino acids

2.1.1 From Amino acids to Proteins

To ease our way towards more complicated topics and to give the non-biochemist reader the opportunity to enjoy a snippet of this work, this section deals with the topic of amino acids. The following paragraphs will clarify important terms, provide a historical background and underline the biological importance of this class of molecules.

From the perspective of an organic chemist, amino acids are simply organic molecules, possessing at least one amino (-NRR') and one carboxylic acid (-COOH) functional group. However, the general population associates the term majorly with (*S*)-configured L-α-amino acids, bearing a primary amine (-NH_2) in α-position to an acid functionality and the C-atom, adjacent to the amine, carrying specifically one hydrogen atom and an additional variable organic residue. The variable organic group is then further narrowed down to little over 20 different residues that are present in canonical amino acids. This very narrow association of the term amino acid is imprinted into humans, because of the uttermost importance of this particular group of molecules for the existence of life on earth as we know it. In this work, the term amino acid is used a little more broadly, but will always refer to α-amino acids with primary amines.

As described above, amino acids always contain two compatible functional groups (amine-, carboxylic acid-group). Since, relatively stable bonds can be formed between the amine group of one amino acid and the carboxylic acid group of another, they are building blocks for larger oligo- or polymers, also called peptides. The bond between an amine and a carboxylic acid functional group is called an amide- or peptide-bond. Such bonds possess a partial double bond character due to electron delocalization rendering it kinetically stable against hydrolysis and the amide group planar. Also, the peptide-bond does not freely rotate, providing a structurally defining rigidity to peptides (Figure 2).

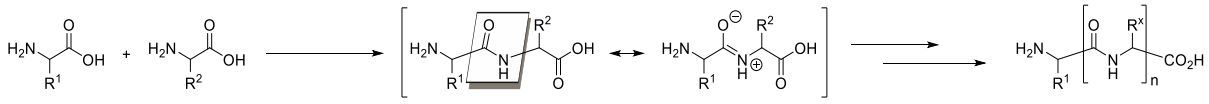


Figure 2: Amino acids, the planar amide bond and oligomeric peptides.

Contrary to the chosen depiction in Figure 2, amino acids are zwitterionic in aqueous solution at pH-values around neutral, meaning the amine is protonated (ammonium) and positively charged and the acid-group is deprotonated and negatively charged (carboxylate). In other words, the ammonium-group is less acidic than the acid group. For this reason, small amino acids are very polar and well soluble in water.

The described inherent modularity, diversity, and structural determination of peptides form the foundation for the importance of these oligo- and polymers not only in the life of a chemist.

A defined set of L-α-amino acids represents the building blocks (monomers) of proteins which are linear peptides and the structural basis of living organisms. The linear sequence of these building blocks in a protein (written from the N- to C-terminal), also called the primary structure of the protein, is coded within the DNA of an organism. The complete set of proteins in the DNA of an organism, called its genome, determines its genotype and every species has a unique and defining set of proteins coded within its DNA. Leaving aside mitochondria, an organism's genome holds all information needed to build the organism from scratch and change it during its life cycle, according to its age and environmental influences leading to the observed phenotype.

However, not all possible L-α-amino acids are built-into proteins. Instead, organisms use a small predefined subset, the proteinogenic amino acids, in humans 21. This is valid for all organisms, with minor variations. 20 of these are called canonical amino acids (selenocysteine in humans is excluded), presenting the standard set of amino acids in living organisms. Sometimes, proteins contain slight variations of these proteinogenic amino acids, like methylated or phosphorylated versions, which occur due to post-translational (after translation of DNA into protein) modifications. All of this does not mean that organisms do not employ other amino acids, just not for building proteins. One example of such other metabolically important amino acids is L-homoserine, an intermediate in the biosynthesis of methionine, threonine and isoleucine.

Eight amino acids are the so called essential amino acids, among them are for example histidine, lysine, and phenylalanine, meaning that humans cannot produce these in sufficient quantities, rendering them an essential component of their diet.

The linear sequence, or primary structure directly defines how the built-up macromolecule will fold into a well-defined three-dimensional structure, called the proteins tertiary structure, composed of smaller reoccurring secondary structures. These secondary structural elements are repeating, recognizable structural elements, like helices (*e.g.* α-helix) and planes (*e.g.* β-fold), building up rigid zones inside the peptide molecule, resembling bars, plates with flexible hinges, the so-called loops, in between.

All of these secondary elements then build up a larger structure, which is called the tertiary structure. It is fascinating and its own field of expertise, how the linear sequence of the amino acid chain, coded in the DNA already defines the later overall tertiary structure. The folded amino acid strands then potentially interact with other tertiary structures and smaller, permanently incorporated small organic molecules forming even larger, fully functional quaternary structures.

One group within the proteins is called enzymes. These enzymes can specifically catalyse chemical reactions, by binding substrate molecules into the protein's quaternary structure, lowering the transition state energy of a specific reaction and thus, decreasing the activation energy, needed to turn the substrates into products.

Coming back to the machine-inspired picture of bars, plates and hinges, the thought that the protein's overall 3D structure is static, represented in the famous "lock-and-key analogy", introduced in 1894 by Emil Fischer, is long overhauled. [3] Even though Fischer mainly wanted to refer to the strict stereospecificity of enzymes, a picture of enzymes as rigid, static moulds was projected, where certain molecules can fit in and then react. Instead, it is now well known that the overall structure of enzymes is highly flexible, interacts and reacts on changes in its surrounding in a purposeful manner. Therefore the "lock-and-key analogy" has been mostly replaced by Koshland's "induced-fit" model, describing that an enzyme changes its shape, responding on a substrate that only then perfectly fits inside, or onto a surface. [4] Such behaviour is also observed for example in the prominent SARS-CoV-2 spike protein, extending its stinger, upon interaction with the human ACE2 receptor (*vide infra*).

It might well be that interactions with small molecules on one side of a protein are cascaded through the whole structure, through lever- and gear-like movements, leading to specific changes on the very opposite side of these, sometimes huge protein molecules. But also, this picture of a machine, reacting on a provided material is spoofed through the lively, breathing like motions, going on in these molecules, when they are put into their natural surroundings.

This dynamic nature of proteins also makes the elucidation of their structure very difficult. State-of-the-art methods are based on crystallization of proteins, often with a bound substrate. These crystals are then subjected to radiation, for example X-rays. The three-dimensional structure is deduced from the refraction patterns, produced by the static atoms in the crystal upon radiation. However, there is no guarantee that the dynamic protein has the same structure in its natural habitat as in the static crystal and it is not easy to deduce the dynamics of an enzyme from the picture of a static crystal structure.

Last but not least, many proteins also interact with copies of themselves, forming dimers or higher oligomers, or with other proteins, forming bigger units, called quaternary structures, sometimes incorporating also smaller organic molecules and inorganic ions (cofactors, coenzymes), crucial to render them fully functional.

One of the most prominent examples of a protein in recent times, the aforementioned SARS-CoV-2 spike protein is displayed in Figure 3, to illustrate the concept of quaternary, tertiary and secondary structure.

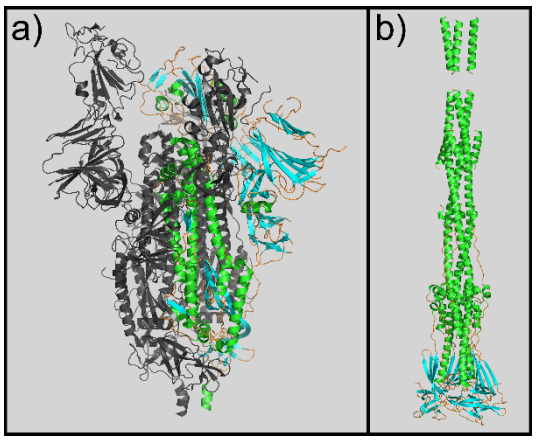


Figure 3: a) SARS-CoV-2 spike protein trimer, prefusion, stabilized (pdbid: 7AD1). [5] In one monomer, secondary structural elements are color-coded, helices in green, sheets in cyan, and loops in orange. Retracted length: ~160 Å. b) SARS-CoV-2 spike protein trimer, post-fusion (pdbid: 7E9T). [6] All monomers coloured. Extended length: ~240 Å.

The spike protein, extending from the virus surface is constituted of three identical subunits, bound to each other by intermolecular forces (quaternary structure). Two of these are displayed in grey. The third subunit is coloured according to the respective secondary structural elements. The complex function of the spike can be described as follows in a simplified way. At the bottom of the picture the spike protein extends into the virus membrane and at the top, there are areas for interaction with the human ACE2 receptor, located on the surface of human cells, for example a lung-cell. This interaction leads to major structural rearrangements of the spike protein, initially cutting off the top sections. Then, the core helices extend, puncture the human cell membrane, anchoring inside. Ultimately, the process culminates into the fusion of both membranes, allowing the virus to spill its molecular machinery into the host cell and begin replication. [6,7,8]

Proteins are subdivided into several groups, most frequently according to their shape (globular, fibrous, and membrane proteins), or function (*e.g.* structural proteins, signalling proteins, enzymes).

This work is dedicated to enzymes, proteins that catalyse specific chemical reactions, by lowering the internal energy of these reactions' transition states. Their size usually lies in the range between 30 and 100 kDa, with the smallest subunits starting from 62 amino acids. [9] (smallest known enzyme: 4-oxalocrotonate tautomerase, with a subunit mass of less than 7 kDa) Even though, theoretically, by fusion infinitely large, the largest structure protein, found in humans is titin, forming filaments in our muscles with about 34000 amino acid in the protein chain and a weight of about 3.6 MDa ($6*10^{-12}$ µg).

2.1.2 Amino acid Enantiomers

Amino acids in natural proteins are present only in the L-form (exception: peptidoglycan, *vide infra*), one of the two possible enantiomers of this type of small molecules, a fact that is described with the term "biological homochirality". The word chiral (Greek: chéri - hand) describes the fact that the two corresponding molecules cannot be superposed on their mirror image, like two hands. The two different versions (hands) of such molecules are called enantiomers. In the case of amino acids, they are defined in the spatial arrangement of the four substituents of the α-C atom. The historic D/L-nomenclature stems from the Fischer projection and an amino acid or sugar is called either L (Latin: laevus – left), or D (Latin: dexter – right), according to the position of

either the α-carbon (in amino acids), or the lowest substituent in the carbon chain (in sugars), either located on the left, or right in the molecule's Fisher projection. Nowadays the D/L-nomenclature is only used for sugars and amino acids and in the more recent and general CIP-nomenclature, L-amino acids usually correspond to the (*S*)-enantiomer. This is only reversed, when the variable substituent at the α-carbon has a higher priority, according to CIP, than the carboxylic acid group, for example in cysteine.

Even though there are some convincing explanations for the origin of biological homochirality, based on biological or physical chiral amplification, [10] we are still fishing in the dark for answers on why gene-coded proteins in nature are exclusively build up from L- and not D-amino acids, with hypotheses ranging from chirality inducing radiation to the hand of god, or nowadays preferred by mostly atheist scientists: pure chance. However, if the reason really is pure chance, then homochirality is a pretty strong argument for the common origin of all life on planet earth.

However, the generalization that nature exclusively utilizes L-amino acids has long been overcome. There are many examples on the purposeful employment of D-amino acids and this is entirely logical. A prominent example is the incorporation of D-amino acids into structure-proteins of cell walls (peptidoglycan). Think of it this way: since every organism is tuned on utilizing L-amino acids and devouring organic matter (*e.g.* proteins) to make a living, it should also possess a sophisticated machinery to break down and oxidize these proteins, splitting them into small building blocks, harvesting energy on the way. Enzymes catalysing this crackdown of peptides are called peptidases or proteases which are present everywhere around us, not only in our intestinal tract. Many microorganisms secrete such enzymes into their surroundings, for example as a defence mechanism or to break down larger proteins before the smaller fragments can be taken up as nutrients. [11] As efficient as these proteases work, they are also very specific for proteins made from L-amino acids. Therefore, microorganisms protect themselves, by equipping their cell walls with these structural peptides, build from D-amino acids.

In fact, bacteria construct a macromolecule from linear polysaccharides, interconnected with short peptide strands, the so called murein (Latin: murus – wall) or peptidoglycan (glycans: polysaccharides). Incorporated into these peptide strands are D-amino acids, like D-alanine, D-glutamate, or D-serine.

However, the peptidoglycan is not the only example for D-amino acid utilization in living beings. We now know many examples for purposefully employed D-amino acids in nature and a lot of research in this area is currently ongoing. One exemplary research question is: why do some bacteria specifically produce and secrete monomeric D-amino acids and what signalling cascades are set into motion through this course of action? [12]

Another example for D-amino acids in bacterial excretion is poly-γ-glutamate, a peptide biopolymer, employed by bacteria as incapsulating slimy substance, with adhesive properties, used for example by *Bacillus subtilis* for defence and nutrient storage. [13] This macromolecule has attracted some attention from industry for pharmaceutical formulations or bioplastics, because it does not trigger an immune response and is biodegradable. [14]

In humans and other vertebrates, D-amino acids also play an important role. For example, D-serine is an important regulatory factor in the human brain as a co-agonist of the *N*-methyl-D-aspartate (NMDA) receptor. This receptor plays a key role in excitatory synaptic transmission and is associated with learning, memory and emotions. [15,16,17]

D-Aspartate is also an important regulator in mammals. In a rat model, it has been shown to be present in higher concentrations in the brain shortly after birth, rapidly declining and after three weeks finally no longer detectable. Thus, it is believed to play an important role in the early development of the mammalian brain. [15] Also, it is believed to regulate endogenous hormone synthesis, as it mainly occurs in parts of the endocrine system, like the testis and adrenal gland. [18]

But if α-D-amino acids, like their L-counterparts are present and necessary for the proper functioning of living organisms, how are they produced, when considering the abundance of L-amino acids with seemingly the entire enzymatic machinery being directed at the utilization of these L-enantiomers? Even though the loss of stereoconfiguration can occur in amino acids upon heating, radiation or unspecific chemocatalysis, for appropriate function, an organism must have access to a method for the production of a specific enantiomer on demand.

2.2 Racemases and Epimerases

Racemases are enzymes that catalyse the fully reversible interconversion of one enantiomer of a chiral molecule into the other enantiomer. They belong to the enzyme family of isomerases (EC: 5.1: racemases and epimerases). As catalysts, they do not influence the position of the equilibrium between the two enantiomers and naturally produce the racemate. However, in many racemases, reaction rates for the interconversion of individual enantiomers can differ, thus in a dynamic equilibrium situation, the concentration of both enantiomers may not be equal.

Epimerases are a subtype of racemases, acting on one stereocentre of molecules that possess at least two stereocentres.

Racemases and epimerases are usually named after their best substrate and not exclusive to the molecular group of amino acids (*e.g.* α-methylacyl-CoA racemase). Examples for amino acid racemases are: alanine racemase, serine racemase, glutamate racemase, lysine racemase, arginine racemase, proline racemase and hydantoin racemase.

2.2.1 Amino acid racemases

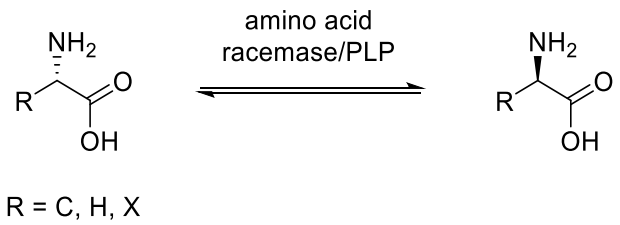


Figure 4: Net-reaction, catalysed by amino acid racemases.

There is a wide variety of racemases in nature with different reaction mechanisms. Since the racemase panel examined in this work was selected based on sequence homology with the broad-spectrum racemase from *Vibrio cholerae* (BsrV), which is a fold-type III (*vide infra*), PLP-dependent racemase, these PLP-dependent racemases need to be discussed in greater detail. Broad-spectrum or broad-specificity racemases (EC: 5.1.1.10) are structurally and sequence-wise close relatives to fold-type III alanine racemases (EC: 5.1.1.1). Molecular structure and catalytic machinery of fold-type III racemases are described in the following section. Further information about process parameters and substrate spectrum are given in the discussion part of this work.

The coenzyme present in this enzyme class is PLP (Figure 5).

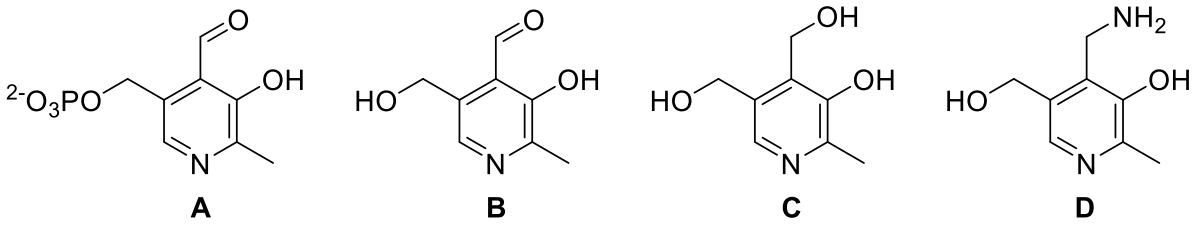


Figure 5: Molecular structures of vitamin B_6 vitamers, including the racemase coenzyme pyridoxal-5'-phosphate (**A,** PLP), pyridoxal (**B**), pyridoxine (**C**) and pyridoxamine (**D**).

PLP is the phosphorylated derivative of pyridoxal, one of the three vitamin B_6 vitamers, besides pyridoxine and pyridoxamine (Figure 5B - D). The organic molecule is composed basically three parts: an electrophilic aldehyde functionality, conjugated to the second part, an electron poor, phenolic six-membered nitrogen heterocycle and the third part, a non-conjugated phosphate group.

The phosphate group is connected *via* a conformationally flexible methylene group to the meta position of the pyridine ring. Due to the methylene bridge, it has little influence on the electronic properties of the mesomeric system. Nonetheless, this group, as in many other biological molecules, increases the solubility in water and represents a handle for polar binding inside an enzyme's active site (*vide infra*).

The pyridine ring, as all aromatic 6-ring heterocycles, is electron deficient, since the lone pair at the ring nitrogen cannot donate its electrons into the aromatic system, due to geometric limitations. The ring carries two -M substituents (ring nitrogen, aldehyde group) and one +M substituent (phenol group). The two alkyl substituents have minor effects on the π-system. Due to its property of readily accepting electrons, donated para to the ring nitrogen, the conjugated π-system of this molecule is often described as an electron sink. (Figure 5)

The aldehyde is able to form Schiff bases with primary, secondary and tertiary amines through nucleophilic attack of an amine lone pair. The formed imines (except those with tertiary amines) are readily protonated and can undergo imine-enamine tautomery. Imine formation also drastically lowers the pK_A of any protons, available at the amine carbon. These properties are the basis for all in-enzyme functionalities of PLP.

PLP and analogues and even simple aldehydes can catalyse a variety of reactions *via* formation of Schiff bases, also without an enzyme scaffold (Figure 6).

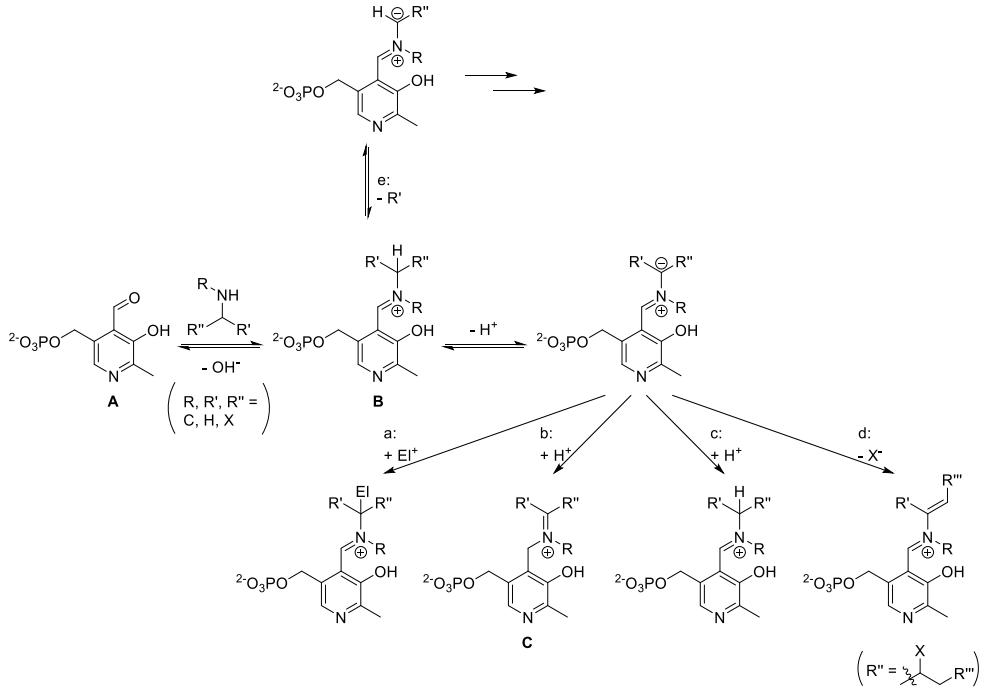


Figure 6: Reactions of PLP: after formation of an imine and deprotonation of the latter, the formed carbanion can undergo several subsequent reactions. a: addition of an electrophile (El^+), b: isomerization, c: racemization, d: β-elimination. The imine can also lose another substituent, *e.g.* decarboxylation, retro-aldol reaction. [19,20,21] **A**: PLP, **B**: aldehyde imine (aldimine), **C**: ketone imine (ketimine).

Many of these reactions are based on the increased α-CH-acidity of an imine bound amine. In general, anions are stabilized that are generated through heterolytic cleavage of covalent bonds at the α-C of the former amine. Among the catalysed reactions are additions of a generated carbanion to electrophiles, racemisation reactions, eliminations and imine isomerisations. However, in many cases these reactions require elevated temperatures (*e.g.* PLP racemisation, decarboxylation reported for 80 – 100 °C) [19,20], the presence of another base, or the presence of additional metal ions in aqueous medium. [22]

Another factor that makes PLP a versatile coenzyme is the tunability of the electronic properties of the conjugated π-system *via* ionization. The ring nitrogen, the phenolate and the catalytically resulting imine group can optionally be protonated. Additionally, an imine-bound amino acid can be ionized at the carboxylate functionality. All these possible ionizations have large effects on the pK_A of protons available at the amine carrying carbon. In the context of an active site, ionizable groups can be hydrogen-bound (H-bound) to distinct residues and thus the ionization state can be selectively controlled.

There has been a detailed study on the effect of imine formation and the ionization of PLP and the carboxylate of amino acid substrates. [23] For this purpose, a model Michael-type reaction with the pyridoxal-5'-phosphate analog 5-deoxypyridoxal was performed at different pH-values in protonated and deuterated water. It could be shown that the negative charge at an amino acid's carboxylate raises the α-C-H acidity by 5 units, and deprotonation of the PLP's phenol raises it by 6 units. Comparing model glycine derivatives led to the conclusion that the α-CH-pK_A is 34 in the negatively charged free amino acid, 29 in the zwitterionic, 21 in the positively charged methyl ester and 14 in the positively charged, protonated imine (with acetone) of the methyl ester. Concluding, there is tremendous (20 units) potential for alteration of the C-H acidity in a PLP-dependent enzyme's active site pocket.

The great versatility of PLP-catalysis can also be found in enzymes, as this coenzyme is present in many different enzyme-classes. Even all of the reaction types depicted in Figure 6 are also taking place in various enzymes. For example, a: 5-aminolevulinic synthase, b: transaminases, c: racemases, d: tryptophanase, e: decarboxylases. These categories are sometimes not exclusive, and some PLP-dependent enzymes cannot be grouped like this at all (*e.g.* glycogen phosphorylase). A feature common to all PLP-dependent enzymes is that the negatively charged phosphate group acts as a handle, establishing polar contacts to the enzyme's peptide and holding the coenzyme in place. Also, all of these enzymes form a covalent bond with the PLP aldehyde group and an active site lysine, the so-called internal aldimine (within the enzyme). This lysine has been shown to be essential for enzymatic catalysis. For example, in alanine racemase from *Geobacillus stearothermophilus*, a K39A mutation renders the enzyme inactive. [24]

Molecular mechanisms of PLP-dependent enzymes are initiated with an attack of a substrate amine group on the internal aldimine, liberating the PLP-binding lysine side chain. Glycogen phosphorylase is an exception, where substrates exclusively interact with the PLP's phosphate group.

One question, subject to extensive research is, how these enzymes create selectivity for one of the reaction pathways (a – e) even though PLP is able to catalyse so many different reactions. Many substrates can possibly undergo different reactions in the active site, nonetheless these enzymes are usually very selective. This is even more surprising in active sites, bearing all features necessary for more than one type of

reaction. For example, bacterial alanine racemases are closely related to eukaryotic aspartate aminotransferases. The amino acid substrate is bound in a very similar way, and after deprotonation of the aldimine, the reactant can be either protonated from the other side, resulting in racemisation, or at the imine carbon, producing the ketimine, thus carrying out the first half reaction of transamination. Indeed, this transamination side reactivity has been examined in alanine racemase from *Geobacillus stearothermophilus* (AlR *G.st.*), however, only occurring about every 10^7 turnovers, depending on the reaction conditions. [25,26] This side reactivity has been investigated also in mutational studies. [27]

Another side reactivity often occurring in racemases is β-elimination of heteroatoms, for example in β-halogenated alanine derivatives in this case leading to enzyme deactivation. [28,29,30] Also, elimination of the serine hydroxy group, termed dehydratase activity, has been reported. The resulting enamine is liberated and tautomerizes to the corresponding imine, which is converted to pyruvate in water. This phenomenon is described for example for serine racemase from *Orizae sativa* (rice) [31] and has also been found in the racemase panel under study (see discussion part).

Aside from reactivity, PLP-dependent enzymes are mainly classified by their fold-type. There are seven different fold-types, based on amino acid sequence and three-dimensional arrangement of secondary structural elements. PLP-dependent amino acid racemases, subject to this thesis, belong either to fold-type II or III. Since the selection of the examined racemase panel was based on sequence-homology to the broad-spectrum racemase from *Vibrio cholerae*, which is a fold-type III enzyme, it is assumed that all examined enzymes belong to fold-type III, an assumption that has never been contradicted during the conducted experiments. Also, when conducting homology modelling of two panel sequences, these have been found to belong to type III. Thus, only this enzyme's type is discussed in greater detail. In general, it is assumed that bacterial alanine racemases, most bacterial serine racemases, eukaryotic aspartate racemases and the known broad-specificity/broad-spectrum racemases belong to fold-type III, while fungal alanine racemases, eukaryotic and some bacterial serine racemases belong to fold-type II. [32]

The first crystal structure of an alanine racemase (AlR) was published in 1997 by Shaw, Petsko and Ringe with good resolution (Figure 7, 1SFT, 1.9 Å). [33]

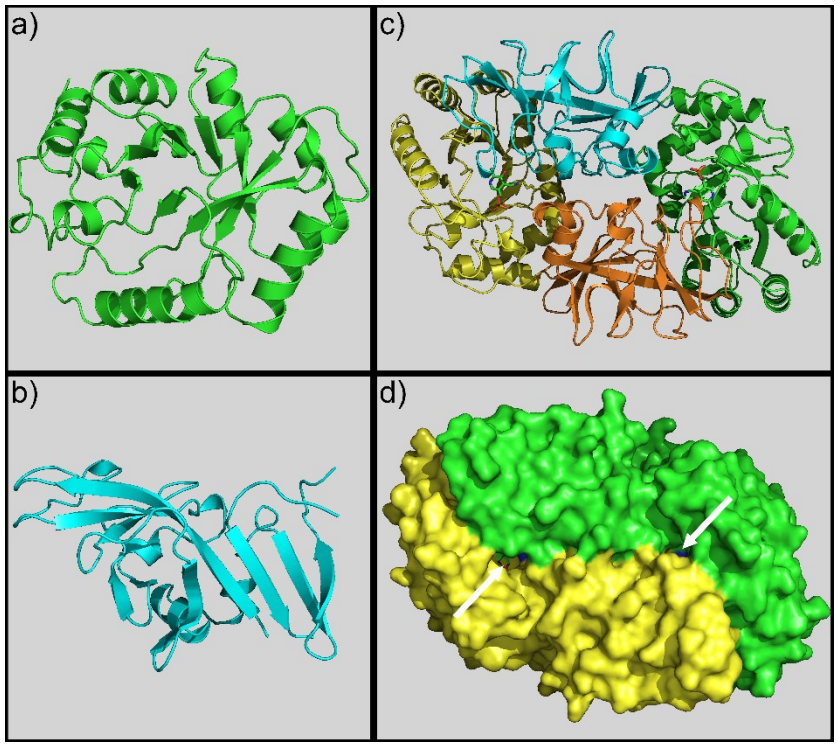


Figure 7: Crystal structure of alanine racemase from *Geobacillus stearothermophilus*, PDBID: 1SFT. [33] a: Domain 1, eight stranded α/β-barrel, b: domain 2, c: domain assembly in the dimer with one monomer in green/cyan and the other in yellow/red, coenzyme bound as internal aldimine, d: enzyme surface of dimeric assembly with active site entry marked with arrows.

Since then, the group of professor Ringe has published many structures of the enzyme with different ligands [26,27,33,34,35] and the protein database rcsb.org contains more than 50 crystal structures of fold-type III alanine racemases from 16 different organisms (February 2023). The enzyme crystallized in 1997 originates from *Geobacillus stearothermophilus*, can be viewed as a prototype fold-type III AlR and displays a dimeric enzyme, each monomer consisting of two domains: an 8-stranded α/β-barrel, found in many enzymes [36.] and a more unique domain, consisting mostly of β-strands. The PLP cofactor is bound covalently as the so-called internal aldimine to the Lys39-N^{ε} (numbered as in PDBID 1SFT) and centred on top of the opening of the barrel. The second domain of the other monomer is pushed onto the top of the barrel, sandwiching the PLP cofactor in between the barrel of one monomer and the other domain of the other monomer, forming a C2-symmetric dimer resulting in a two active site holoenzyme. Both active site entrances are located on the same side of the dimeric enzyme.

The active site and molecular mechanism of BsrV and alanine racemase from *Geobacillus stearothermophilus* (AlR_{Gs}) are almost identical. Also, there are many publications about AlR_{Gs} crystal structures (*vide supra*) and its function. Therefore, the following section will focus on describing the active site and reaction mechanism of AlR_{Gs} and then point out the most important differences between BsrV and AlR_{Gs}. For more information about differences between the two types of amino acid racemases, visit the discussion part of this work and the studies of Espaillat et al. 2014. [37]

When looking at enzyme mechanisms, there are three key features to be considered.

I: What is the general procedure and what active site residues enable this procedure to happen. In other words, how is the reaction happening in the enzyme?

II: The second question is, which step in the mechanism is rate determining, thus between which two intermediates lies the highest activation barrier and how is this barrier lowered compared to a non-catalysed reaction in solution. In other words, why is the reaction happening.

III: Especially interesting in the case of racemases: What is the reason for the high selectivity of the enzyme's main reaction, over the many other possible pathways, supported by PLP?

I: The how-question has been looked into a lot in the case of racemases, but there are still aspects that are not entirely clarified. The general procedure can be looked upon from the perspective of the substrate molecule. The first step is the diffusion of the amino acid substrate from the bulk solvent into the active site. During this step, the substrate still interacts with the solvent, but also already comes into contact with the electrostatic surface of the enzyme and residues at the active site entrance. This can lead for example to a change in the ionization state of the substrate, or to the repulsion of molecules that have a certain charge or polarity. This effect becomes even stronger in the channel leading to the active site. Certain aspects of this will be discussed in the following paragraph and in the discussion part of this work.

Upon arrival of the substrate in the active site, many enzymes are reported to respond to substrate binding with closing of the active site, to exclude water and to move certain residues in positions, beneficial to catalysis. This usually includes loop, inter-domain, and/or inter-mer movements.

However, to the best of my knowledge, there are no large structural rearrangements upon substrate binding in broad-spectrum racemases. Crystal structures of internal and external aldimines (expected to correspond to open and closed forms of the enzymes) do not exhibit significant differences in domain positions. It might be concluded that active site closure would lead to a negative effect on the efficiency of catalysis, considering than many molecules in biological systems are able to enter the active site. For example, acetate, often used as a buffer component, can be found in the crystal structure that was reported first for alanine racemase (PDBID 1SFT). [33] Also, a continuous connection to the bulk water might be beneficial (Grotthuss-mechanism, [38] *vide infra*).

Before substrate binding and molecular reaction mechanism, the binding and positioning of the coenzyme should be described.

A crystal structure of AlR_{Gs} with bound (*R*)-1-aminoethylphosphonic acid (L-AlaP, substrate analogue as external aldimine) and the corresponding journal article describe the mode of binding of both coenzyme and substrate. [34] Enzyme amino acid numbers are given for AlR_{Gs} (PDBID 1BD0), with the respective numbers for BsrV (PDBID 4BEU) [37] in brackets. The active site geometry is described relative to the two faces of the PLP coenzyme, given for the prochiral carbonyl function of free PLP. The C-C bond between the phenyl moiety and the carbonyl does not rotate, due to a hydrogen bond between the phenol oxygen and the imine nitrogen of the PLP, bound in the active site as internal or external aldimine. Thus, the two faces of the imine also define the two sides of the pyridine moiety.

Figure 8 displays the described binding amino acid side chains as in the aforementioned crystal structure of AlR_{Gs} with bound (*R*)-1-aminoethylphosphonic acid (PDBID: 1BD0). [34]

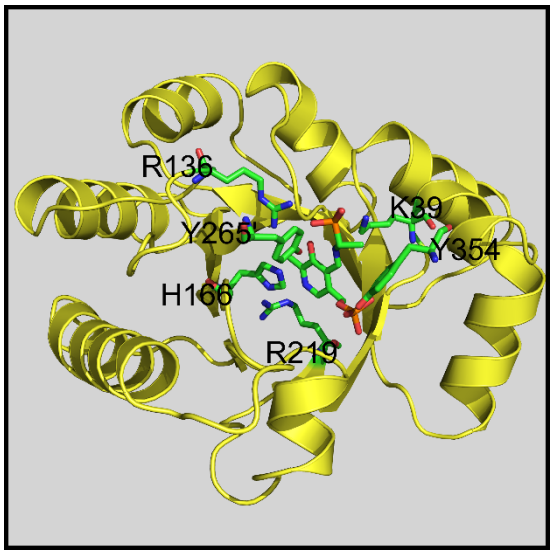


Figure 8: Excerpt from the crystal structure of AIR_{Gs} with bound (*R*)-1-aminoethylphosphonic acid (PDBID: 1BD0), [34] containing the described amino acid residues and the α/β-barrel structure.

The PLP's *re*-side is stacked against nonpolar amino acid residues from the α/β-barrel-opening (*vide supra*). Its pyridine nitrogen is hydrogen bound to an Arg219 (R260) guanidyl nitrogen, thus is mostly deprotonated. Many polar and basic amino acid residues are located on the PLP *si*-side, with the His166 (H204) imidazol ring stacked to the aromatic pyridine ring being one of them. This histidine is a feature in many PLP-dependent enzymes. The phosphate group is directed towards the PLP carbonyl, on the *si*-side, pointing towards the active site entrance. Two of its free oxygen atoms form hydrogen bonds with three amide groups of the enzyme's main chain, one serine hydroxy and one tyrosine phenol. The third oxygen (O5) is in bonding distance with a phenol of Tyr354 and two water molecules. Interestingly, this latter tyrosine is not present in BsrV and instead, there are additional water molecules. This increases the active site volume for larger substrates. O5 is connected with Tyr265' (Y299'), one of the catalytic bases *via* a water network. When a substrate is bound, this water H-bond-network is also connecting the other catalytic base Lys39 (K74), mediated *via* the PLP phenol and the substrate's carboxyl (*vide infra*). Due to that, both catalytic bases are potentially connected *via* a H-bond network, theoretically allowing fast proton transfer between them *via* a Grotthuss mechanism. The PLP phenol is also bound to Arg136. In BsrV, this arginine (R173) is retracted a little bit further away from the coenzyme, mediating its contact to the PLP phenol *via* a water molecule and again creating more space in the active site (*vide infra*).

Upon entry of the amino acid substrate into the active site, it passes the PLP phosphate. This probably helps in orientation of the substrate, since it repels the negatively

charged carboxyl group which is thus pushed towards the PLP phenyl group. It might also assist in deprotonation of the substrate's amino group, if not already deprotonated.

The substrate's carboxyl group is bound tightly in many amino acid-transforming enzymes and this carboxylate often acts as a handle for substrate recognition and orientation (*e.g.* D-amino acid oxidase, D-amino acid dehydrogenase). [39,40] Racemases also showcase this feature by the fact that short chain organic acids, if small enough to fit into the active site, are inhibitors. [41] A crystal structure of propionate bound to the active site of AlR_{Gs}, partly reflecting the substrate's carboxylate binding, reveals one of its carboxylate oxygens bound to the Met312' (M347') main chain nitrogen and the other one bound to the PLP phenol, Tyr265' (Y299') and the aforementioned conserved arginine Arg136 (R173). This arginine seems to be essential for carboxylate positioning in type III racemases, since it is held in position in alanine racemases and broad-spectrum racemases through secondary interactions, even though different ones. In alanine racemase, there is a *N*-carboxylated Lys139, generating hydrogen bonds with the two guanidyl-nitrogens of Arg136 that do not participate in substrate binding. In broad spectrum racemases, there is a negatively charged chloride ion, which coordinates to these two guanidyl nitrogens, N174, and two water molecules, which in turn are positioned through secondary interactions. [37]

In 2002, a study containing the crystal structures of bound D- and L-alanine revealed that the mode of carboxylate binding is changed between the two enantiomers. [42] D-alanine is bound very similar to propionate, except the oxygen that binds to the Met312' amide nitrogen also binds to Y284' and one water molecule. The other oxygen atom binds to Tyr265' and Arg136, but not to the PLP phenol. Instead, there is another water molecule in H-bond distance, which in turn is bound to the Asp313' side and main chain. Interestingly, in BsrV, this aspartate is replaced by asparagine (N348').

The L-alanine carboxylate binding is altered so that the oxygen that binds to Y265', now binds to K39. This is a very important finding, because it leads to the conclusion that the abstracted proton is directly transferred *via* the two active site bases, mediated by the substrate carboxylate group.

The molecular reaction mechanism is depicted in Figure 9 and described in the following paragraph.

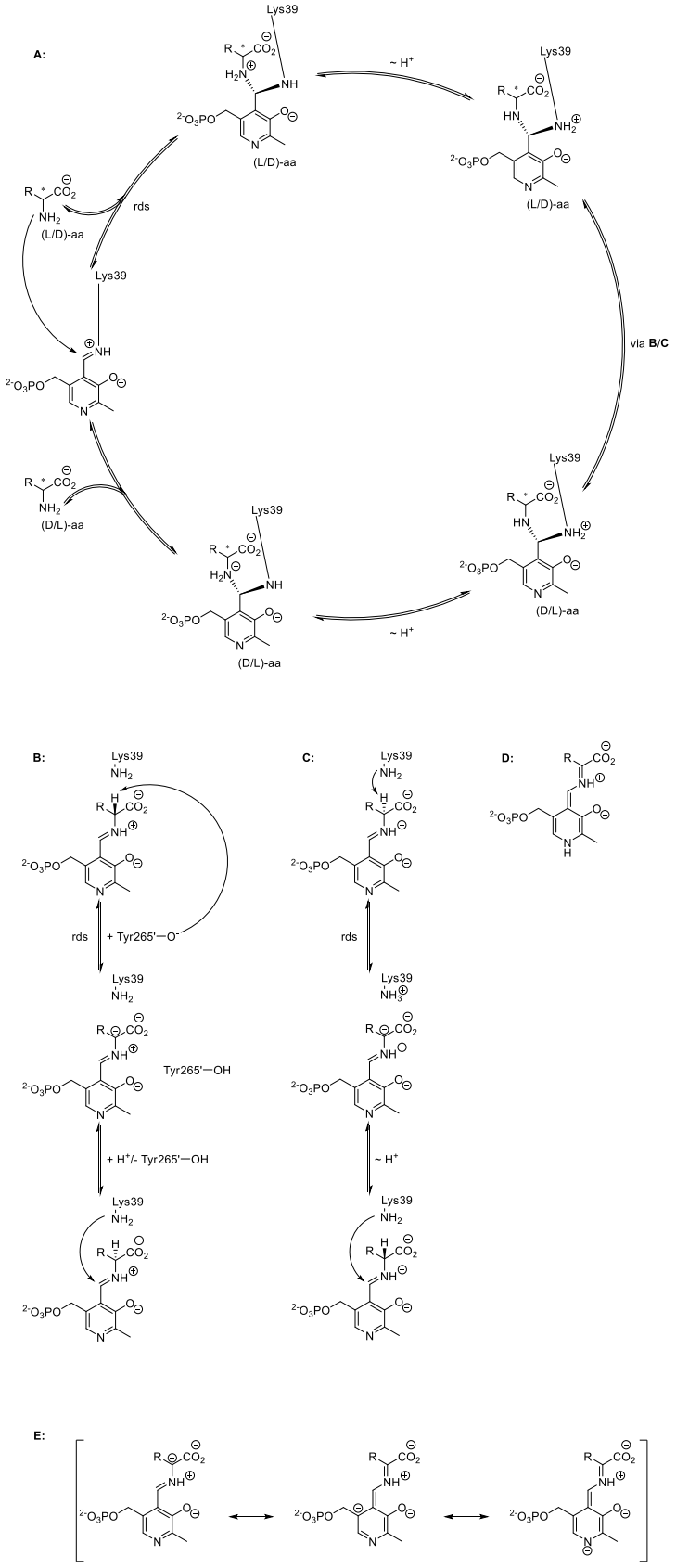


Figure 9: A: Catalytic cycle of racemases. B: L- to D-amino acid racemization. C: D- to L-racemization. D: quinoid intermediate in aspartate aminotransferases, observed spectra prove little relevance of quinoid structures in racemases. E: relevant anion stabilization *via* the cofactor. aa: amino acid, rds: candidates for rate determining steps.

First step of the reaction mechanism is the attack of the substrate's amine on the *si*-face of the internal aldimine, resulting in the formation of a tetraedric geminal diamine. Then, the Lys39 (K74) amine is liberated on the *re*-face of the formed external aldimine. Depending on the chirality of the amino acid substrate it is now deprotonated by one of the two active site bases. The previously bound Lys39 (K74) deprotonates the α-carbon of D-amino acids and the phenolate oxygen of Tyr265' acts as a base on the corresponding L-amino acid. The Tyr265's phenol group has been reported to have an unusually low pK_A of 7.9 at pH 7.0. [43]

The formed negative charge is usually thought to be stabilized through delocalisation into the PLP and carboxylate π-orbitals. However, in both crystal structures of bound L- and D-alanine, the imine nitrogen does not lie within the pyridine plane, but on its *re*-face, thus not in a position to donate electrons into the aromatic system. [42] The deprotonated external aldimine is subsequently protonated on the opposite side. The opposite active site base is believed to perform this protonation and that the proton is transferred from the protonated active site base to the other one, *via* the substrate carboxylate group. Possible hydrogen bonds that would provide the necessary network for this proton transfer are depicted in Figure 10.

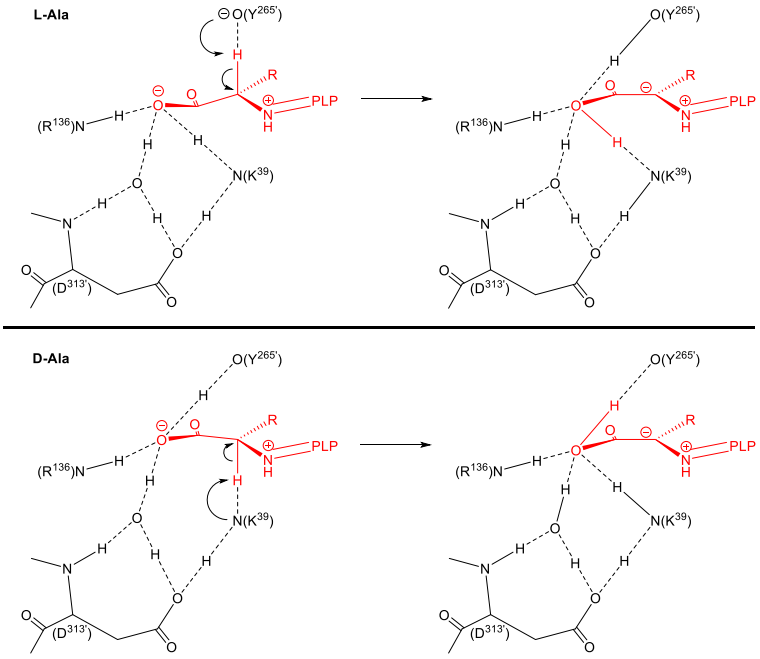


Figure 10: Current view on the proton transfer between the two active site bases Lys39 and Tyr265', based on the H-bond situation in the crystal structures of AlR$_{Gs}$ with L- and D-alanine. [42] Charges, valency and H-bonding only exhaustively displayed for the participating carboxylate oxygen and the two active site bases.

If there is no additional proton present, as displayed in Figure 10, between the respective deprotonated base and the substrate carboxylate group, the Grotthus

mechanism discussed below is ruled out and movement of the protonated base towards the carboxylate group through C-C or C-heteroatom-bond rotation is necessary to pass on the proton.

However, in theory, also another active site acid, for example a water molecule could act as proton donor in this step, even restoration of the original ionization state after product release is possible.

Subsequent to protonation of the external aldimine from the other side, a second transaldimination by the active site lysine liberates the amino acid with switched chirality and reforms the internal aldimine.

It should be clear that individual steps in the mechanism are reversible, due to microscopic reversibility, but that it is far from sure that the overall catalytic cycle is reversed for racemisation of the respective enantiomer. In fact, many aspects speak against this theory. It might even be that racemisation of L- and D-amino acids follows different pathways and protonation of the anion intermediate does not include the corresponding protonated base. [32]

The reader shall not be deprived of the fact that to racemise the amino acid, it would also be possible to flip the positions of the carboxylate group and the side chain of the deprotonated, bound amino acid, rotating the anion around the C-N single bond. However, it is argued that the carboxylate is, as in many enzymes acting on amino acids, tightly bound and held in position by active side residues (handle, *vide supra*) and secondly, the found second active site base would not be necessary in this case. This second tyrosine-base, however, has been identified and proven to be crucial for racemisation, so the two base mechanism is not debated in literature today. A lot of research has been performed on this behalf. For example, the effect of exposing racemases to cycloserine, an irreversible inhibitor for alanine racemase and an antibiotic agent, has been studied. Cycloserine cannot be racemised and instead performs the first half reaction of transamination, leading to the formation of a stable ketimine. While D-cycloserine still irreversibly inhibits an Y265F mutant of AIR_{Gs}, like both enantiomers do for the wild-type enzyme, L-cycloserine inhibition becomes reversible. [27] This among many other experiments strongly supports the two-base mechanism. The second base Tyr265 has been reported in 1999 to deprotonate L-alanine. [44]

Its noteworthy that the search for racemase inhibitors is a promising field, since a specific racemase inhibitor would probably exhibit strong antibiotic properties against many gram-negative bacteria, among them various human pathogens. (*e.g. Vibrio cholerae*, multiresistant staphylococci) But even though this was investigated for the past century, still no breakthrough has been made. The research has been focused on the active site and most inhibitors lack specificity. Blocking the active site entry might be more promising.

II: Describing, what happens with a substrate during its transition to a product does not really clarify, why the enzyme catalyses the reaction at all. Deprotonation and racemisation of an amino acid can occur also in water, even more so, if a catalytic Schiff-base forming aldehyde is present in solution. However, amino acid racemisation with PLP alone is only observed at high temperatures (80 – 100 °C) and is very slow. Protonation of an anion with a very high internal energy is fast and occurs spontaneously in protic solvents. Hence, the fact that racemases catalyse racemisation must be due to a decrease in the activation energy of the deprotonation step. In theory, this can either be achieved by stabilizing the anion, decreasing the C-H acidity, or by increasing the energy of the ground state, for example through the provision of a strong base, which might be precisely localized in an enzyme's active site.

A homogenic reaction in bulk solvent energetically proceeds from educt to product *via* a transition state. This transition state has a higher internal energy than both product and educt and is not stable. The energy difference between the educt and the transition state is called activation energy and is linked *via* the Arrhenius equation to the reaction rate k, where A is the frequency factor (collision factor), within limits independent of the temperature, E_A is the activation energy, R the gas constant and T is the temperature.

$$k = A \cdot e^{\frac{-E_A}{R \cdot T}}$$

Most enzyme catalysed reactions are multistep reactions. Therefore, it is important to clarify that the speed of the complete reaction is determined by the activation energy of the step that has the highest energy barrier, the so-called rate determining step (rds).

Enzymes speed up reactions by a factor of 10^{10} – 10^{20}. This translates into a decrease of the rds-activation energy of about 50 to 100 kJ/mol at 30 °C. In enzymes, this can be achieved by either lowering the energy of the transition state, or by increasing the energy of the Michaelis complex.

In the theory of Michaelis-Menten, describing enzyme kinetics, it is assumed that enzymes first form a complex with their substrate, the Michaelis complex. Then the Michaelis complex undergoes a reaction to form the product complex and in the final step, the product dissociates from the enzyme (Figure 11).

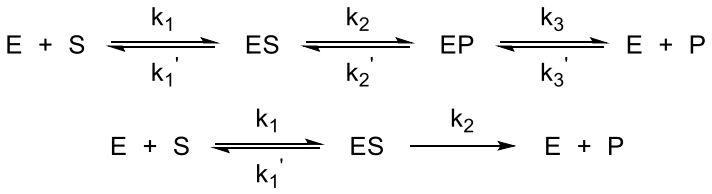


Figure 11: Michaelis Menten reaction equations. Top: full equation, bottom: simplified version. E: enzyme, S: substrate, ES: Michaelis complex, EP: enzyme-product complex, P: product.

To analyse enzyme kinetics, the reaction equations are further simplified. For this purpose, the reaction is described at the very beginning, where the substrate concentration is at a maximum and the product concentration is close to zero. Also, it is assumed that the substrate concentration is orders of magnitude higher than the enzyme concentration (pseudo first order). Under these conditions, k_3' is close to zero and product dissociation becomes quasi-irreversible. Additionally, for most enzymes k_3 has been reported to be much lower than k_2, which also seems to be the case for racemases. This theory is used to describe the tendency of molecules to bind to an enzyme and the reaction rate of a substrate at the beginning of the reaction (v_0), depending on the substrate's concentration.

Another important theory in this context is Hammond's postulate. It states that the transition state geometry is always nearest to the ground state with the closest free energy in the energy diagram of a reaction step (Figure 12).

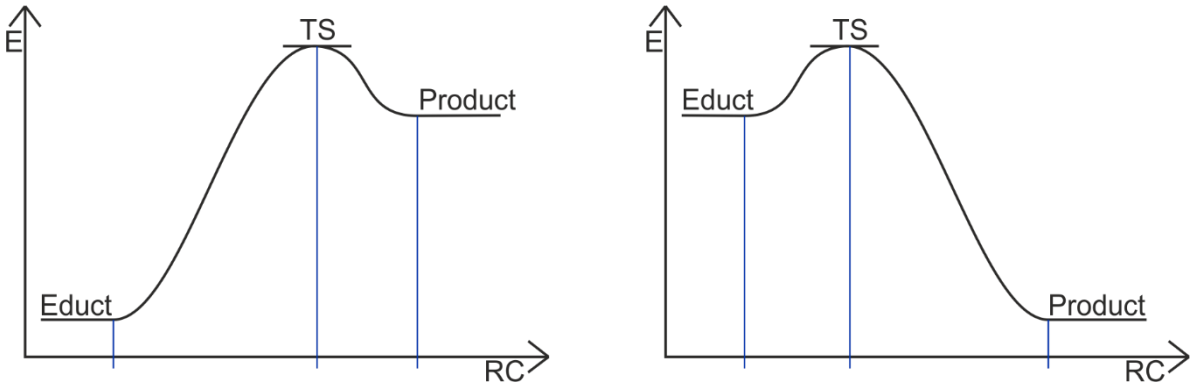


Figure 12: Hammond's postulate: E: free energy, RC: geometric reaction coordinate, TS: transition state.

Looking at the racemisation reaction in water, it should be clear that both, L- and D-amino acid have the same free energy and that the intermediate anion has much more.

Considering that the anion has so much higher internal energy, the transition state should be close to that in terms of free energy and geometry. Thus, deprotonation of the external aldimine as rds, proposed for example by Major *et al.* seems logical. [45] However, most publications report a transaldimination step to be rate limiting. [33,42,46] Since the activation barrier of the enzyme's rds should be at least 50 kJ/mol lower, it seems counterintuitive that transaldimination from the internal to the external aldimine ought be the rate determining step. But, as reported by Warshel *et al.*, [47] not the practically occurring reaction in water should be compared with the enzyme mechanism, but the same steps that occur in the enzyme, should be compared to the same as if occurring in water. Even though complex to dive into, this review is recommended, because it clarifies many misunderstandings about why enzymes catalyse reactions, strongly arguing for an electrostatic basis of lowering the transition state's free energy. So, in general it seems to be fair to conclude that the rate determining step is the formation of the geminal diamine, even though this has yet to be proven.

III: To understand the following explanation, the possibility of different reaction pathways needs to be kept in mind (*vide supra*). In fact, the first half reaction of transamination has been investigated as a side reaction in AlR_{Gs}, but, occurs apprx. only every 4 million turnovers, [26] depending on the reaction conditions. If the basic Arg219 (R260), which coordinates to the pyridine nitrogen, is exchanged through a glutamate in the R219E mutant, the ratio of transamination to racemisation is increased apprx. 600-fold and also the rate of transamination is slightly increased. [48]

In aspartate transaminases, this pyridine nitrogen is always coordinated to an acidic aspartate side chain, but why does its protonation disfavour racemisation, but not transamination? Aspartate transaminases do not carry a second catalytic base, as found in racemases. Instead, lysine deprotonates the external aldimine and then transfers the proton to the imine carbon, generating a ketimine from the aldimine. This happens *via* conformational changes in the lysine side chain, including several C-C bond rotations, expected to take more than 100 ps. For comparison, the rotation around the internal butane C-C single bond was calculated to take about 40 ps in a CCl_4 solution at room temperature (r.t.). [49] Proton transfer from a water molecule or *via* a Grotthus mechanism on the other hand has been shown to occur on a 0.5 – 5 ps timescale, [50] so roughly 20 to 200-fold faster.

It seems reasonable to assume that this could be the main reason for the preference of racemisation over transamination. Destabilization of the anion through deprotonation of the pyridine nitrogen, disfavouring a quinoid intermediate, favours faster proton transfer from a weaker base along already existing hydrogen bonds. In fact, this quinoid intermediate, previously believed to be part of the reaction mechanism of racemases, is found to be present in transaminases, but has been shown not to occur in the wild-type AlR. [51,52,53]

If there is a direct H-bond connection between both active site bases with the protonating base protonated prior to α-C deprotonation by the deprotonating base, even a concerted proton transfer without the occurrence of an anion intermediate would be possible. However, this has been shown not to be the case in AlR_{Gs} through kinetic isotope effect experiments. [54]

Another possible reason for the preference of racemisation over the first step of transamination might be that the basicity of the α-carbon is kept a lot higher through destabilization in racemases, compared to transaminases. This way, proton transfer to this carbon centre is favoured over proton transfer to the less basic aldimine carbon. However, for the overall reaction to be catalysed, the energy difference between the educt and the transition state, which is highest in energy, must be lowered. So, to compensate a high energy anion, which is, due to Hammond's postulate, similar to the transition state, the enzyme has to compensate a lot of this energy through internal electrostatic tensions.

To conclude, even though research about fold-type III racemases has been going on for a long time and many publications about mechanistic details have been published, several aspects require further clarification. Nonetheless, the picture about molecular interactions involved in substrate binding to active site residues is already quite clear and represents a solid base for mutational studies and computational docking experiments, potentially even leading to new broad-spectrum antibiotic agents.

As already mentioned, more information about differences between alanine racemases in terms of molecular structure, substrate spectrum and process parameters is collected in the discussion part and the work of Espaillat et al. 2014. [37] This publication describes a molecular fingerprint of broad-spectrum racemases and an additional N-terminal extension, not present in other racemase types, which binds to the same

extension of the other monomer in the homodimer. This extension is not present in alanine racemases and no information about its purpose has been found in literature. However, removal of the N-terminal drastically lowers the enzyme's activity. The same paper also states that the entry channel to the active site is approximately two times wider due to shorter loops in the peptide chain of the enzymes and that the entry site carries a negative charge in BsrV. Also, broad-spectrum racemases have been recently found to bind to and be inhibited by peptides occurring in the biosynthesis of mureins, which might be important for this enzyme's regulation. [55]

Apart from amino acid racemases, other enzymes used in reactions described in this thesis were D-amino acid oxidases (DAAOs), D-hydroxy acid dehydrogenase, an enzyme derived from D-aminopimelic acid dehydrogenase (D-HicDH), and transketolase (TK). They all act on amino acids, or compounds that are derived from amino acids. The following chapters contain brief descriptions of their structure and mechanism.

2.3 D-Amino acid oxidase

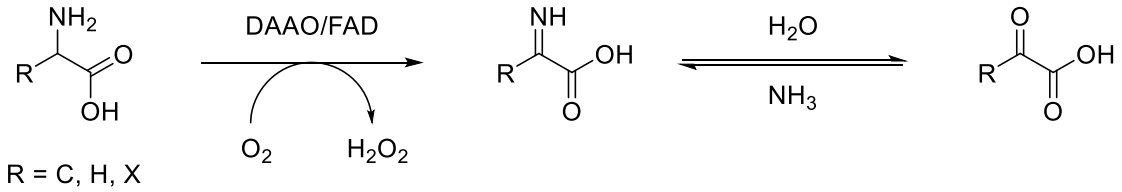


Figure 13: Net-reaction, catalysed by D-amino acid oxidase (DAAO).

D-Amino acid oxidases (DAAO, EC 1.4.3.3) are oxidoreductases, which oxidize D-amino acids to the corresponding imino acids. Redox equivalents are transferred to molecular oxygen *via* the coenzyme flavin adenine dinucleotide, yielding hydrogen peroxide. The product imino acid spontaneously hydrolyses to the corresponding 2-keto acid (also called α-keto, 2-oxo and pyruvic acid). The reverse reaction is not detected.

The oxidation of D-glutamate was discovered in the 1930s in liver and kidney tissues of different vertebrates. [56,57] Shortly thereafter, oxidation of D-alanine and D-valine was also discovered in a resting cell preparation of *Bacillus proteus*. [58] Today, D-amino acid oxidase (DAAO) from pig kidney is commercially available, mainly because it is an important medicinal tool in amino acid analysis.

Besides its importance as a bioanalytical tool, the oxidase is used for the industrial production of 7-aminocephalosporanic acid (7-ACA), the β-lactam core and basis of semi-synthetic, broad-spectrum cephalosporin antibiotics. The first large-scale production process was published in 1995. [59] The process is based on the use of immobilized DAAO in a reactor, pressurized with oxygen. These parameters already point out certain limitations of the oxidase, among them its limited stability and low K_M for molecular oxygen. More recent research includes the enzyme's use in electrochemical biosensors, [60] and its medicinal importance. [61]

In nature, the enzyme also has various purposes. In mammals, the enzyme serves to eliminate detrimental indigested D-amino acids and to regulate endogenous D-amino acid concentrations (D-aspartate, D-serine, *vide supra*). [39,62,63,64] Since mammals produce catalase in specific organelles, their DAAO is only expressed inside these peroxisome organelles to protect other tissues. This way, the toxic DAAO-byproduct H_2O_2 is directly quenched by peroxisomal catalase. Microorganisms are enabled to grow on D-amino acid substrates by the enzyme. [65]

2.3.1 Structure and Mechanism

D-Amino acid oxidases are structurally diverse enzymes, presumedly due to development of the enzyme during an early stage of evolution. The following discussion of DAAO structure will be based on the commercial enzyme from pig kidney (*Sus scrofa*) and the yeast DAAO from *Rhodotorula gracilis*, because these were the enzymes employed in this work.

To understand an enzyme, one should first look at the coenzyme. The coenzyme of DAAO, flavin adenine dinucleotide (Figure 14) contains riboflavin (vitamin B_2) and is common to many oxidoreductases.

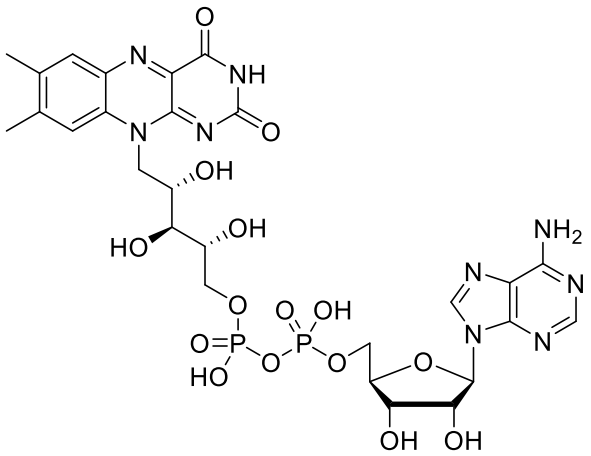


Figure 14: Molecular structure of the DAAO coenzyme flavin adenine dinucleotide (FAD).

The molecule has two functional units, the riboflavin, taking part in the molecular catalysis, and the adenine nucleoside diphosphate, being bound to the enzyme *via* polar contacts.

The overall structural monomer assembly of DAAO matches the PHBH-fold, named after *p*-hydroxybenzoate hydrolase, a flavin dependent monooxygenase. It occurs in many, but not exclusively flavoenzymes, even though PHBH-fold enzymes sometimes have very low sequence similarity (< 20%). The topology is characterized by the existence of two distinct domains, an FAD-binding domain and an interface domain. The FAD-binding domain consists of a β-sheet with several α-helices on both sides and the interface domain has a β-sheet flanked on one side by α-helices. Both domains are connected *via* several flexible loops. Even though the overall fold is similar, PHBH-fold enzymes are often very different, with certain secondary structural elements largely modified, eliminated, or added and also often possess different catalytic

residues. However, beside their overall fold, they also have almost identical FAD-binding domains with the nucleoside moiety tightly bound inside the FAD binding domain and the flavin isoalloxazine ring-system located at the interface of the two domains. [66]

Distinct amino acid residues from the peptide chain have not been identified to largely contribute to the stabilization of anionic intermediates, instead there is an α-helix's N-terminal pushed towards the N5 of the isoalloxazine from the opposite side to where the substrate is bound. [67] This helix's dipole is believed to be the structural element stabilizing reaction intermediates. Another common feature is that they close the active site upon substrate entry, plausibly to exclude water to protect reactive reaction intermediates. [68]

Depending on the source, the enzyme can have different oligomer assemblies with one active site per monomer. For example, the pig kidney DAAO is monomeric and head-to-head dimerization upon higher concentration leads to inhibition, whereas the yeast enzyme (*e.g. Rhodotorula gracilis*) is a head to tail dimer (Figure 15). [69,70,71]

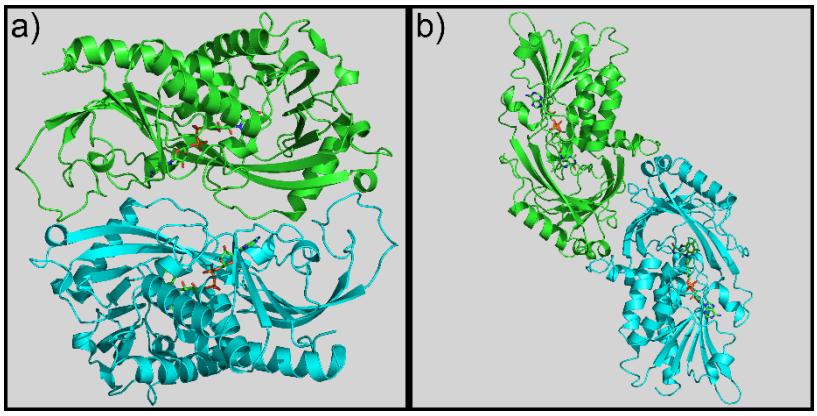


Figure 15: DAAO dimers with bound FAD. a: DAAO from *Rhodotorula gracilis* (PDBID: 1C0L), [70] b: DAAO from porcine kidney (PDBID: 1KIF). [69]

The fact that the yeast DAAO is a dimer leads to a higher stability of the enzyme, due to a different deactivation mechanism with dimer dissociation as the first reversible step.

The molecular reaction mechanism is described as binding of the amino acid's carboxylate *via* polar contacts and H-bonds to three conserved residues (Y224, Y228 and R283, labelled as in $DAAO_{pk}$), leading to active site closure. [72,73,74,75] The anionic amino acid is positioned with its α-C hydrogen atom close (3.4 Å) [69] to the isoalloxazine N5 on the *re*-side of the aromatic system. Practically irreversible hydride

shift leads to iminium formation and reduction of the isoalloxazine. [70,76,77] For flavin reoxidation with molecular oxygen, there are two possible pathways: reoxidation occurs (i) after imine release, representing a classical ping-pong mechanism, or (ii) while the product is still bound (ternary complex). The latter is true for the yeast enzyme and for the oxidation of non-basic amino acids by the pig kidney DAAO. In general, O_2 reactivity is defined by active site electrostatics and oxygen accessibility. [78] For example, DAAO$_{Rg}$ has an O_2-channel leading to the isoalloxazine N5 on its *si*-side. The developing positive charge at the iminium nitrogen is expected to be transported as proton *via* a water-network to the *si*-sided HO_2^- species. [79]

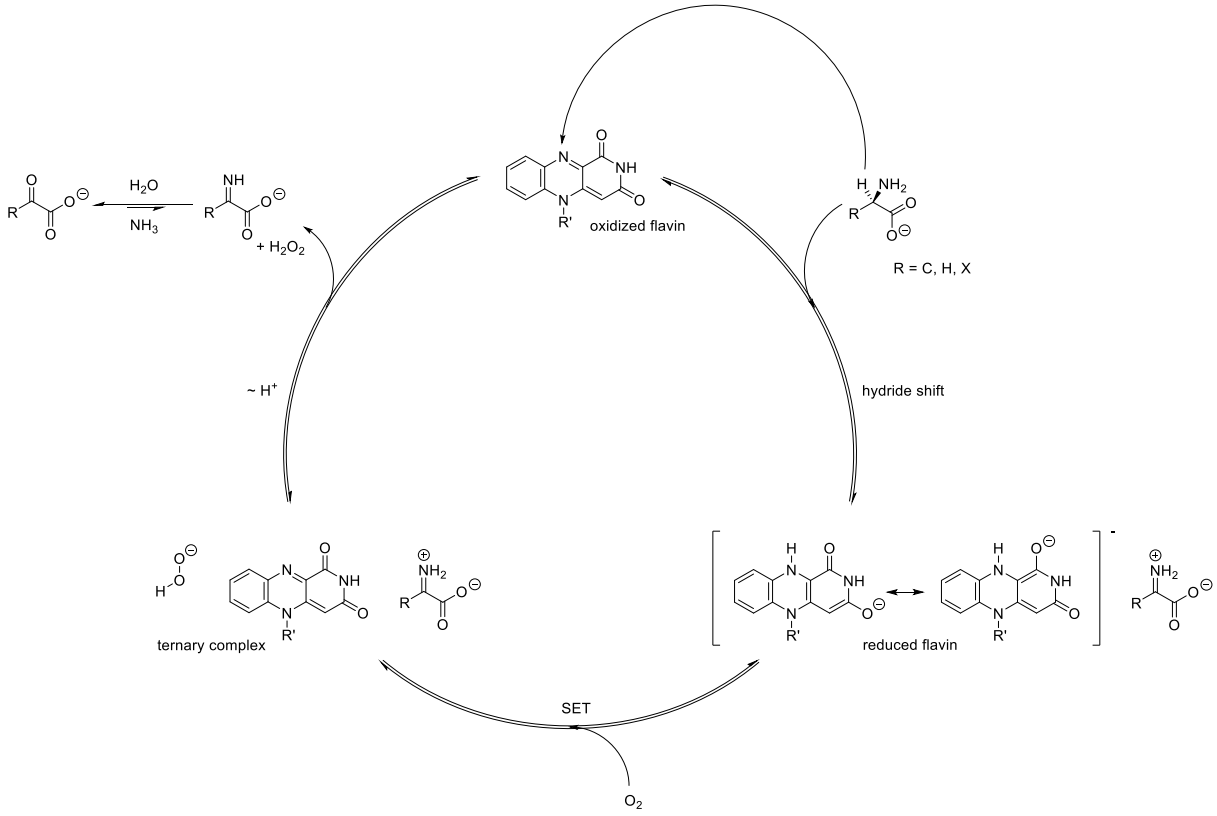


Figure 16: Molecular mechanism of DAAO$_{Rg}$, SET: single electron transfer, R': see *Figure 14*.

Determining steps have been shown to be product release in mammalian DAAO, supposedly due to an additional active site loop, restricting active site access and space. FAD reduction is the rds in the yeast enzyme. [39,80] Flavin oxidation, expected to proceed *via* single electron transfer (SET), [81] is always faster than reduction. [72,82]

Substrate scope and process parameters will be discussed in Chapter 4.3. It should be kept in mind that the enzyme has limited stability and in addition produces highly reactive H_2O_2. Both, the low stability and hydrogen peroxide production pose serious

limitations on biocatalytic reactions and require extensive optimization and method-adaptation.

2.4 D-Hydroxyisocaproic acid dehydrogenase

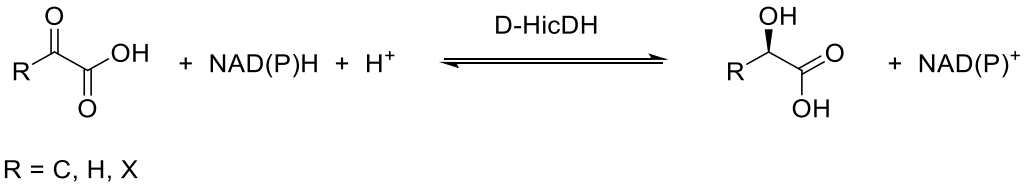


Figure 17: Net-reaction, catalysed by D-hydroxy acid dehydrogenase (DHADH).

D-Hydroxyisocaproic acid dehydrogenase D-HicDH from *Lactobacillus paracasei* (D-HicDH$_{Lp}$, EC: 1.1.1.345) is an enzyme from the large family of D-2-hydroxy acid dehydrogenases. [83] Well-known relatives are for example D-lactate dehydrogenase from *Treponema pallidum* and phosphite dehydrogenase from *Stutzerimonas stutzeri* with sequence homologies of 36% and 29%, respectively. Native substrate is (*R*)-2-hydroxyisocaproic acid, or D-leucic acid, the hydroxy acid analogue of leucine. D-2-Hydroxyacid dehydrogenases catalyse the reversible reduction of α-keto acids to the corresponding D-2-hydroxy acids. Like all dehydrogenases, they utilize the reduced and oxidized form of β-nicotinamide adenine dinucleotide (NADH/NAD$^+$), or its phosphorylated derivative nicotinamide adenine dinucleotide phosphate (NADPH/NADP$^+$) as their cosubstrate. Biotechnological alternatives for D-HicDH$_{Lp}$ are for example the D-HicDH from *Haloferax mediterranei*, [84,85] the hydroxypyruvate dehydrogenase from *E. coli* (YiaE), [86,87,88,89] and the D-lactate dehydrogenase from *Lactobacillus delbrueckii ssp. bulgaricus*. [90,91,92,93]

D-2-Hydroxyacid dehydrogenases have diverse roles in organisms. D-HicDHs take part in fatty acid metabolism. Biotechnologically, they are applied for the stereoselective production of enantiopure α-hydroxy acids, some of which are valuable building blocks for biodegradable polymers or complex molecules. The application of final compounds ranges from active pharmaceutical ingredients (APIs), [94,95,96] over scaffolds for orthopaedic implants, or drug delivery systems, [97,98,99] to fillers in cosmetic formulations. [100]

Due to high structural similarities in dehydrogenases, many aspects of D-HicDH are discussed, based on related enzymes. The dehydrogenase cosubstrate NAD(P)$^+$/NAD(P)H is expensive. Therefore, different dehydrogenases are often combined in biocatalytic reactions for the purpose of cosubstrate recycling. For example, formate dehydrogenase in combination with formate salt is a desirable NADH

recycling system, due to its intrinsic driving force and the potential of formate salts to constitute the volatile NH_4HCOO buffer system, facilitating workup (see discussion part). Relevant in the light of climate change, the formate dehydrogenase from *Thiobacillus sp.* KNK65MA has been found to possess a comparably high activity for CO_2 reduction, rendering formate dehydrogenases a potential research target for highly desired CO_2 fixation. [101]

2.4.1 Structure and Mechanism

The dehydrogenase cosubstrates $NAD(P)^+$/NAD(P)H are dinucleotides derived from niacin (vitamin B_3). Its adenoside part serves for recognition and binding to the enzyme, while the ribose bound nicotinamide functions as an electron acceptor, forming dihydronicotinamide upon reduction. In cells, the direction of dehydrogenase redox reaction is governed by the proportion between reduced and oxidized forms of the two cosubstrates. While the non-phosphorylated version in organisms is mainly used for the catabolic oxidative direction (NAD^+>NADH) and most generated NADH is used to produce ATP, the phosphorylated version facilitates many anabolic reduction reactions. [102]

D-HicDH is selective, but does not exclusively work with NAD^+/NADH (Figure 18), with about only 10% of its activity retained with $NADP^+$/NADPH. [103]

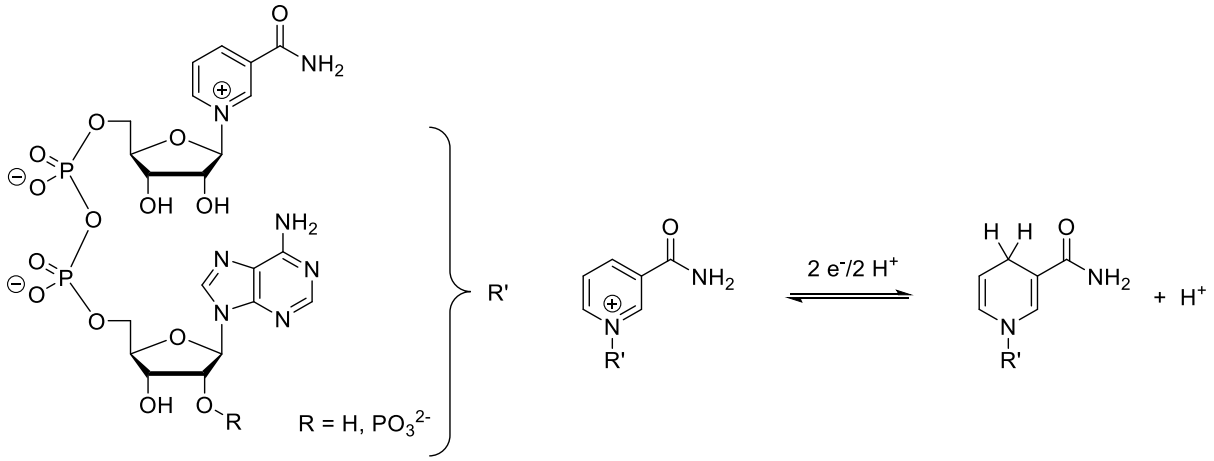


Figure 18: Molecular structure of the D-HicDH cosubstrate nicotinamide adenine dinucleotide (phosphate) ($NAD(P)^+$(NAD(P)H + H^+).

D-Hydroxy acid dehydrogenases have recently been partitioned into 22 sub-classes, based on phylogenetic sequence analysis, with 9 of them biochemically characterized. D-HicDH has been placed into the sub-group of bacterial enzymes within the subfamily of D-lactate dehydrogenases. Inside its subgroup D-HicDH is unique, because contrary

to other D-lactate dehydrogenases, it has a lower k_{cat}/K_M for pyruvate than for other, larger substrates. [104]

Most D-hydroxy acid dehydrogenases are dimeric enzymes, with the monomers consisting of two domains. The nicotinamide binding domain is located in the middle of the peptide sequence and has a classical nucleotide binding Rossmann-fold with alternating β-strands and α-helices in its sequence. The β-strands are three-dimensionally arranged to form a seven-stranded parallel β-sheet, with the α-helices lining the sheet on both sides. The so-called catalytic, or substrate-binding domain has a similar Rossman-like fold, with a five stranded parallel β-sheet, constructed from C- and N-terminal secondary structural elements. Designations as substrate binding or catalytic domain are misleading, since substrate binding is accomplished by residues from both domains and since side chains representing a catalytic acid/base stem from the nicotinamide binding domain. Both domains are connected *via* two flexible loops, forming a hinge between them. The active site lies at the bottom of a deep cleft, formed by residues of both domains and the flexible hinge. This overall fold with the deep cleft is typical for dehydrogenases in general and catalysis is believed to be governed by large inter-domain movements. These movements are described as a rolling motion upon transition from an open, to a super-closed conformation of the enzyme. [92,105,106] Since L-, and D-amino acid dehydrogenases and L-, and D-hydroxy acid dehydrogenases are closely related structurally and catalytically, and since L-dehydrogenases are very well described in literature, some of the following explanations will be based on research reports about L-specific enzymes.

Figure 19 displays the D-HicDH enzyme dimer, as in the pdb-structure 1DXY. [83]

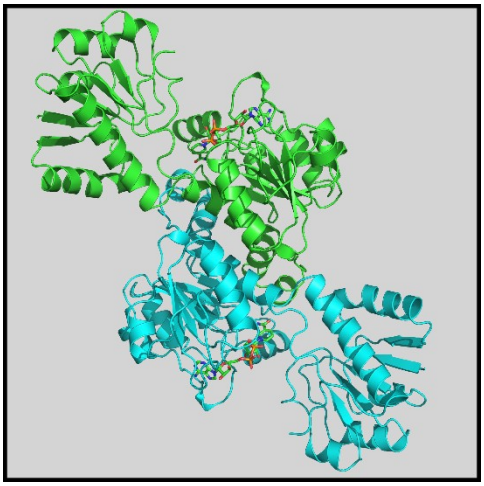


Figure 19: D-HicDH dimer with bound NAD^+ (pdbid: 1DXY). [83]

The cofactor is bound *via* many noncovalent contacts. The adenine moiety is located inside a hydrophobic pocket, also able to accommodate other small aromatic molecules. The pyrophosphate linker is bent into a gauche conformation, with the negatively charged oxygen atoms bound *via* an α-helix dipole, salt bridges and H-bonds to polar side chains (His155) of the peptide chain. Both ribose units are bound *via* hydrophobic interactions and H-bridges to their hydroxy groups. One important feature, shown to make dehydrogenases differentiate between NAD^+ and $NADP^+$, is a conserved aspartate (D175), repulsing the phosphate group of $NADP^+$. Mutational studies have been able to convert NAD^+ specific D-lactate dehydrogenase (D-LDH) from *Lactobacillus bulgaricus* to prefer $NADP^+$, by exchanging this amino acid through arginine (pdbid: 2DBQ, D-leucine dehydrogenase triple mutant D176S/I177R/F178T). [93] The surrounding of the nicotinamide is overall hydrophobic, with the amide group bound *via* polar interactions, leading to stronger binding of the non-charged, reduced cosubstrate.

Developed acceptance and preference of D-hydroxy acid dehydrogenases for NADPH is sought to be exploited for the production of valuable D-hydroxy acids by heterologous expression into cyanobacteria. This facilitates NADPH regeneration through photosynthesis, considered an environmentally friendly and energy efficient route towards the production of these chiral alcohols. [103]

While the cofactor is only bound by the nicotinamide binding domain, the hydroxy, or keto-acid substrate is bound by both domains in the closed and super-closed

conformation of the enzyme. In D-HicDH, the three oxygen atoms of substrates are H-bound to His295 and Arg234 of the nicotinamide binding domain, Asn76 and Tyr100 of the catalytic domain and main chain atoms of Val77 and Gly78. [83]

Whether catalysis is initiated by independent binding of substrate and cosubstrate, or if cosubstrate binding causes the enzyme to transition to a semi-closed conformation, leading to substrate binding is still under debate. [83] However, successive cosubstrate and substrate binding is likely, because substrate binding without bound cosubstrate is unproductive, the substrate is bound by residues of both domains in the closed conformation and because the cosubstrate cannot access the active site in the closed conformation. Successive binding of cosubstrate and substrate is also described for other dehydrogenases of the same family, leading to further domain closure. [107,108,109]

After complete domain closure, the R234 approaches H295, rendering the imidazolium moiety more acidic, facilitating 2-keto acid protonation, thus activating it for hydride transfer. [110,111] The two basic amino acids histidine and arginine are accompanied by glutamate (E263), stabilizing the imidazolium moiety of H295. This catalytic triad is highly conserved in D-hydroxy acid dehydrogenases. [108] In the oxidative direction, the deprotonated imidazole receives the proton upon hydride transfer from the hydroxy acid to the nicotinamide. The substrate's α-carbon is brought into close proximity (~2.8 Å) to the nicotinamide C4, with the hydrogen atom perpendicularly to the aromatic ring, to facilitate hydride transfer.

Charge repulsion of the developing positive charge of the oxidized nicotinamide from the predominantly hydrophobic interdomain contact zone [83] leads to active site opening. In the reductive direction, H195 imidazolium repulsion from R234 causes the active site to open and products to be released (Figure 20).

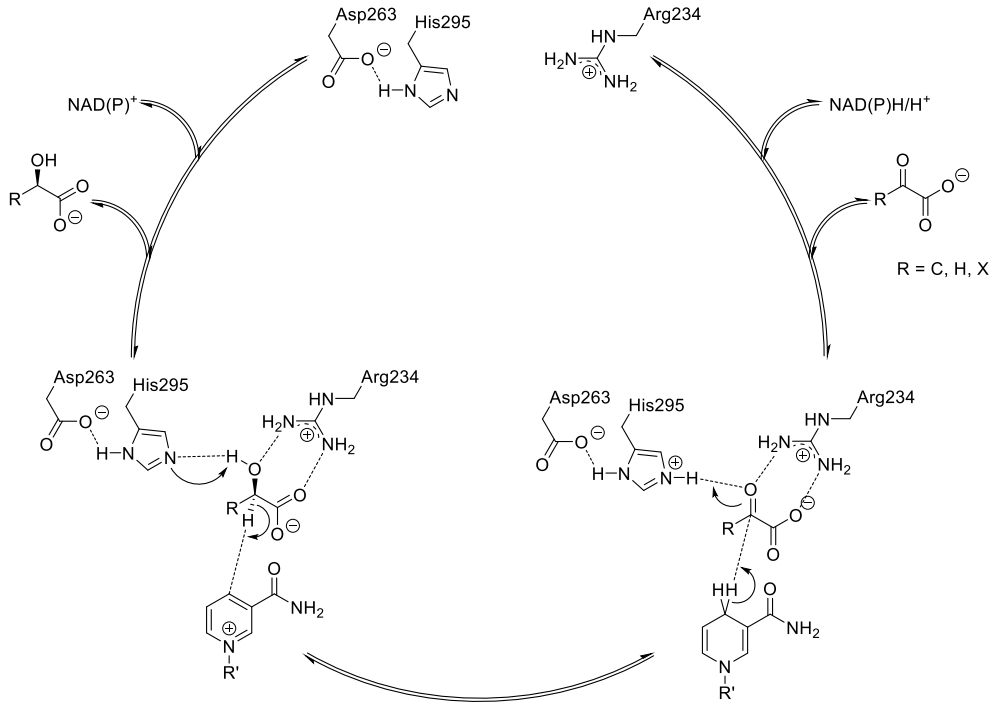


Figure 20: Molecular mechanism of D-HicDH. R': see *Figure 19*.

Definition of a rds has not been found in the literature. For L-specific hydroxy acid dehydrogenases, either inter domain movement prior to hydride transfer, or cofactor release are considered to be possible rate-determining steps. [107,112]

2.5 Transketolase

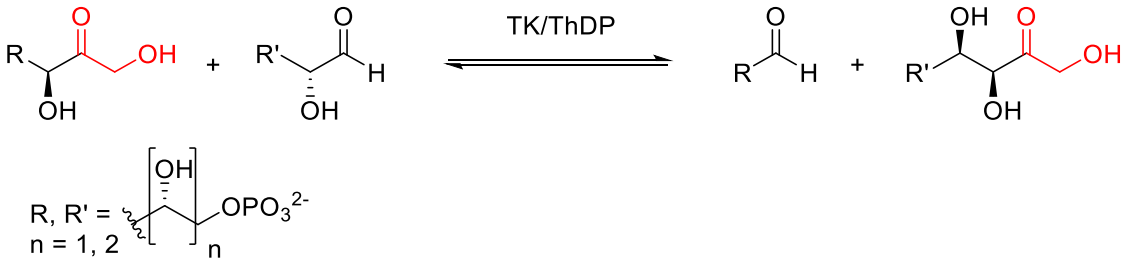


Figure 21: Natural net-reactions, catalysed by transketolase (TK).

Transketolase (EC: 2.2.1.1) is a central metabolic enzyme from the nonoxidative part of the pentose phosphate pathway and a key enzyme in all organisms. [113,114] The enzyme catalyses the reversible transfer of a two carbon hydroxyketone unit ("ketol") from a ketose to an aldose, transforming the aldose into an elongated ketose and the original ketose into a shortened aldose. During one turnover, the ketose, donating the ketol unit is called the donor and the aldehyde is called the acceptor. The natural substrates are terminally phosphorylated D-forms of sedoheptulose, fructose, ribose, xylulose, erythrose and glyceraldehyde, so C3 to C5 aldoses and C5 to C7 ketoses. In mammals, the enzyme is expressed in all tissues, underlining its central role in sugar metabolism, linking glycolysis with the pentose phosphate pathway and supplying important intermediates for nucleotide, aromatic amino acid, and vitamin biosynthesis. [115] In plants, it also connects the photocatalytic Calvin-cycle to the aforementioned metabolic functions by supplying ribulose-5-phosphate. [116] Compared to other enzyme families, variations in structure and function are small between transketolases from different sources.

The enzyme used in this thesis, is the transketolase from the thermophile *Geobacillus stearothermophilus*. It has a high thermal stability and can be used at 70 °C for three days without apparent loss of activity. [117] Also, compared to mammalian transketolase, it has a relaxed substrate scope, even further widened during years of mutational studies. [118,119,120,121,122,123] These features make it an attractive enzyme for biotechnological applications with potential, ranging from API to bioethanol production. [124,125] Recently one variant (H102L/H474S) has even been found to catalyse the synthesis of *N*-arylhydroxamic acids, using nitrosoaryls as acceptor and for example pyruvate as a donor. [126] For a deeper look into the substrate spectrum visit the discussion part of this work.

2.5.1 Structure and Mechanism

Thiamine pyrophosphate, the enzyme's cofactor is the biologically active, phosphorylated form of thiamine, or vitamin B_1. The molecule has three parts: an aminopyrimidine ring, able to undergo imine enamine tautomerization, a thiazole, forming an activated ylide during catalysis, and a phosphate anchor, tightly bound to the enzyme. Due to the ability of the thiazole to form a nucleophilic ylide zwitterion and the build-in amine base, the coenzyme functions in a variety of bond-breaking and forming enzymes (*e.g.* pyruvate decarboxylase, pyruvate dehydrogenase complex, pyruvate oxidase, pyruvate ferredoxin oxidoreductase). Upon binding of thiamine pyrophosphate (ThDP), also called thiamine diphosphate (ThDP) to transketolase, the enzyme undergoes several spatial rearrangements, leading to the formation of the active ylide form of the coenzyme, rendering one active site ready for catalysis (Figure 22). [127]

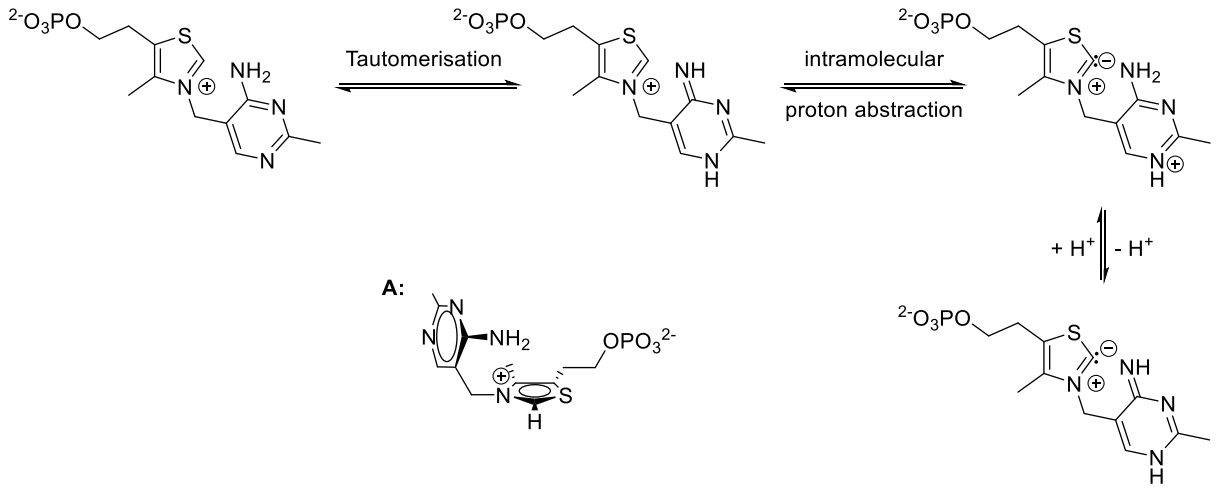


Figure 22: Thiamine pyrophosphate cofactor of transketolases and reaction cascade leading to the formation of its active ylide form. **A**: In active site V-conformation of ThDP.

Tautomerization is assisted by a glutamic acid, protonating the pyrimidine N(1) and a histidine residue, deprotonating the amine. [128]

Besides the ThDP cofactor, the enzyme also contains an octahedrally coordinated bivalent cation for each molecule of ThDP, which participates in the binding of the ThDP's pyrophosphate moiety. While in transketolase, isolated from human tissue, this cation is Mg^{2+}, the enzyme is crystallized almost exclusively with Ca^{2+}. Experiments, applying transketolase synthetically are mostly carried out with Mg^{2+}. The potency of bivalent cations to bind as coenzymes for human transketolase follows the order: $Ca^{2+} > Mn^{2+} > Co^{2+} > Mg^{2+} > Ni^{2+}$. [129] *Thermus thermophilus* HB8 transketolase is

described to preserve its activity upon subsequent treatment with EDTA and Mn^{2+}, Mg^{2+}, Co^{2+}, and Ni^{2+}, while Zn^{2+}, Fe^{2+}, and Cu^{2+} decrease the activity. [114]

The holoenzyme is a triangular, barrel shaped homodimer, forming one active site at each of the two triangular barrel tops, on the monomer interface. The L-shaped monomers of 70 – 75 kDa are subdivided into three domains: the so-called PP (pyrophosphate-binding), Pyr (pyrimidine-binding) and C-terminal domains. Some publications emphasize the flexible inter-domain loops from the three domains (PP-Pyr: 321 – 349, Pyr-C-terminal: 513 – 533, numbers given for TK_{Gst}, PDBID: 8CIP). [115,129]

These loops are believed to carry some features, leading to the high thermostability of some transketolases, besides the large inter-mer contact area (>30% of monomer surface) with many H-bonding and salt-bridging interactions. In the thermostable enzyme from *Thermus thermophilus* HB8, a long inter-domain loop, reaching from the PP to the C-terminal domain (PP-Pyr interdomain loop) contains many prolines, which is believed to prevent interdomain movement and contribute to the high thermostability. [114] In the thermostable enzyme from *Geobacillus stearothermophilus*, this proline-rich region is not present. Instead, the loop has an additional short helix and ends in a short β-strand that forms a short antiparallel β-sheet with the other interdomain region (Pyr-C-terminal interdomain region). For illustration of these differences in the two thermostable enzymes, their monomers are displayed in Figure 23, with the interdomain regions in red. To complete the picture, the transketolase dimer from *Geobacillus stearothermophilus* is also presented.

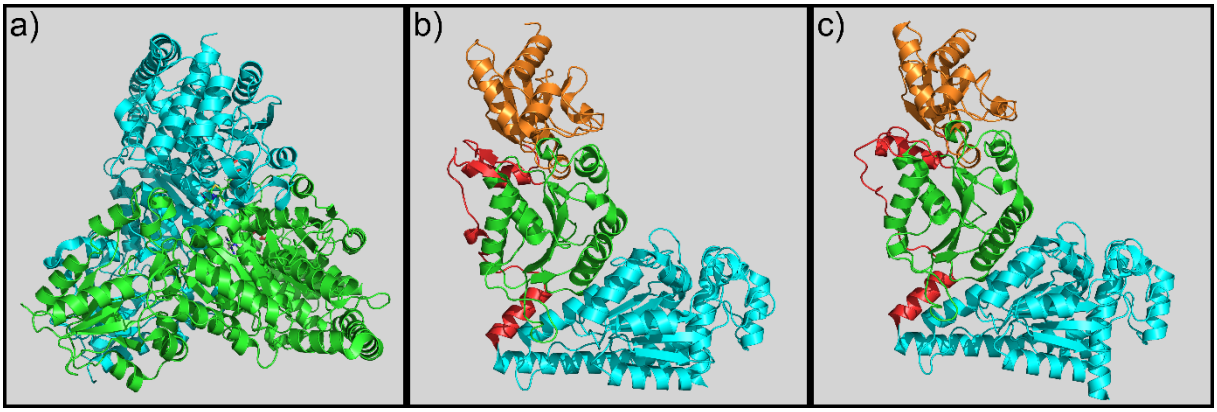


Figure 23: Transketolase pdbids: 8CIP (*Geobacillus stearothermophilus*), 7WRT (*Thermus thermophilus* HB8). [114] a) Dimer from *Geobacillus stearothermophilus* with bound ThDP and active site opening in the centre of the triangular surface. b) Monomer from *Geobacillus stearothermophilus*, with PP-domain in cyan, Pyr-domain in green and C-terminal domain in orange. Interdomain regions displayed in red. c) Monomer from *Thermus thermophilus*, coloured as b).

The so-called PP (3 – 320) and Pyr (350 – 512) domains closely resemble (α/β)-Rossmann-folds with parallel β-sheets, lined on both sides with α-helices. The C-terminal domain (534 – 651) also has a β-sheet, lined with α-helices, but one of the strands is antiparallel and the whole assembly is less tightly packed.

In the following paragraphs, numbers of amino acids in the enzyme's peptide chain are given as in pdbid: 8CIP (transketolase from *Geobacillus stearothermophilus*, 2023, unpublished) and the corresponding residues of transketolase from *Thermus thermophilus* (pdbid: 7WRT, in complex with ThDP and D-erythrose 4-phosphate) [114] are given in brackets. These residues are conserved among transketolases from various sources.

The coenzyme ThDP is bound in a high-energy V-conformation, with the methylene bridge between the thiazole and the pyrimidine as the tip of the V. This conformation is rarely occupied in solution but, conserved in ThDP-dependent enzymes. It is required for catalysis, since the pyrimidine 4-amino group approaches the thiazole C(2)-hydrogen. [115,130] Both aromatic rings are located inside a hydrophobic pocket with low dielectric constant, particularly stabilizing zwitterionic intermediates. [130] This is a common feature in ThDP containing enzymes and for example the dielectric constant in yeast pyruvate decarboxylase has been found to be 13 – 15, compared to ~ 78 in water. [131] This does not only help to stabilize reaction intermediates, but also lowers the effective pK_A of the thiazole C2 compared to the deprotonating imine. [130]

The active site is formed at the interface of the PP-domain of one monomer and the Pyr-domain of the other monomer, from residues of both domains. The PP domain holds the residues binding the bivalent cation and the pyrophosphate of the thiamine cofactor. The cation is bound by the side chains of Asp157 (159), Asn187 (189) and the main chain amide oxygen of Ile189 (191). The three remaining positions in the octahedral coordination sphere of the cation are occupied by one water molecule and one oxygen atom of each of the two phosphates of the ThDP's pyrophosphate moiety. The GDGX_{26}(N/D)(C/S)N (in TK_{Gst}: G156D157G158X_{26}D185S186N187) cation-binding motif is strictly conserved in ThDP dependent enzymes and has been used to identify ThDP dependent enzymes. [115,130]

Besides the cation, the phosphates are also coordinated by His68 (70), Asn187 (189) and His263 (263) and several water molecules. Interestingly, Ser158 also coordinating

the pyrophosphate in some transketolases is replaced through glycine in *Geobacillus stearothermophilus* and human transketolase, directing one coordinating water.

The ThDP's pyrimidine is π-stacked against Phe438' (432') and, together with the thiazole moiety, surrounded by several other non-polar residues (L118, L191, L382', V410', F435'). The 4-amino group is H-bound by the Gly116 (118) amide oxygen. Glu412' (406'), coordinating the pyrimidine N1 is an essential residue in ThDP-dependent enzymes, since it allows easy tautomerization between the imino-chinoid- and the amino-form of the pyrimidine, essential in ThDP catalysis.

Recently, this glutamine has been found to be part of a proton wire between the two active sites in human transketolase. This proton wire seems to be conserved and herein, numbers of corresponding residues are given for TK_{Gst}. The corresponding Nature publication, filled with buzz-words like "Newton's -cradle", demonstrates that low-barrier hydrogen bonds help in signal transmission between the 20 Å apart active sites of the enzyme. The often-reported cooperativity of the two active sites finally receives a worthy explanation, since protonation of the pyrimidine N1, leading to the imino-chinoid-form in one active site, induces N1-deprotonation in the other active site. The hydrogen bond network, including Glu412', Glu162, Glu167, four water molecules and the corresponding residues from the other monomer transmits this signal with a communication rate of $1.2 * 10^5$ s^{-1} in human transketolase, a rate that exceeds k_{cat} by far and mutation of any of the residues either abolishes the enzyme's activity or the cooperativity of the active site. The authors state that this is proof of the significance of low-barrier-hydrogen bonds for enzyme catalysis, a circumstance that has long been under debate. [132]

Another factor in transketolases is the flexible fitting of diverse substrates into the active site hydrogen-bonding network. For microbial transketolase with a more relaxed substrate spectrum, not restricted to sugar phosphates, ketol-donors must have a 3,4-(D)-threo conformation (except hydroxypyruvate and dihydroxyacetone) and aldehyde acceptors must have a (*R*)-hydroxide at the α-position. [129] The reason for the broader substrate spectrum is believed to lie in a broader, funnel-shaped substrate channel in microbial transketolases. [115] D-erythrose-4-phosphate for example is H-bound by His263 (263), His28 (30), Asp470' (464'), and Ser385' (379'), with additional contacts of its phosphate group to Arg358' (352'), His462' (456') and Arg521' (515'). His263 has the role of a catalytic base and is believed to be part of a catalytic "dyad" with an active

site lysine. [113] However, this lysine is not present in TK_{Gst} and the transketolase from *Thermus thermophilus*. [114]

Nonetheless His28 and His102 represent a “backstop” for ketoses entering the active site head first. The two residues strongly bind the two hydroxy groups of the developing dihydroxyethyl thiamine diphosphate (DHEThDP) and fix this important intermediate in a distorted conformation, with the newly formed C-C-bond bent out of the thiazole plane. H102L/H474S and H102T mutants have been shown to decarboxylate pyruvate and 3-methyl-2-oxobutyrate respectively, rendering these two molecules, lacking the hydroxyl group of the ketol unit, donors. [121]

Arg358’, Ser385’, H462’, and Arg521’ are the residues always contacting the phosphate group of donors and acceptors, representing “gatekeepers” at the active site entrance. Sugar phosphates of different chain length are then assuming conformations, directing their hydroxy-groups towards the different histidines, Asp470’ and Ser385’. Asp470’ always binds the (*R*)-3-hydroxy group of aldehyde acceptors. This circumstance has been used to design a L382D/D470S double mutant, where an aspartate is established at the opposite site, now accepting 3(S)-hydroxy aldehydes as substrates. [118]

The current understanding of the full catalytic cycle is displayed in Figure 24.

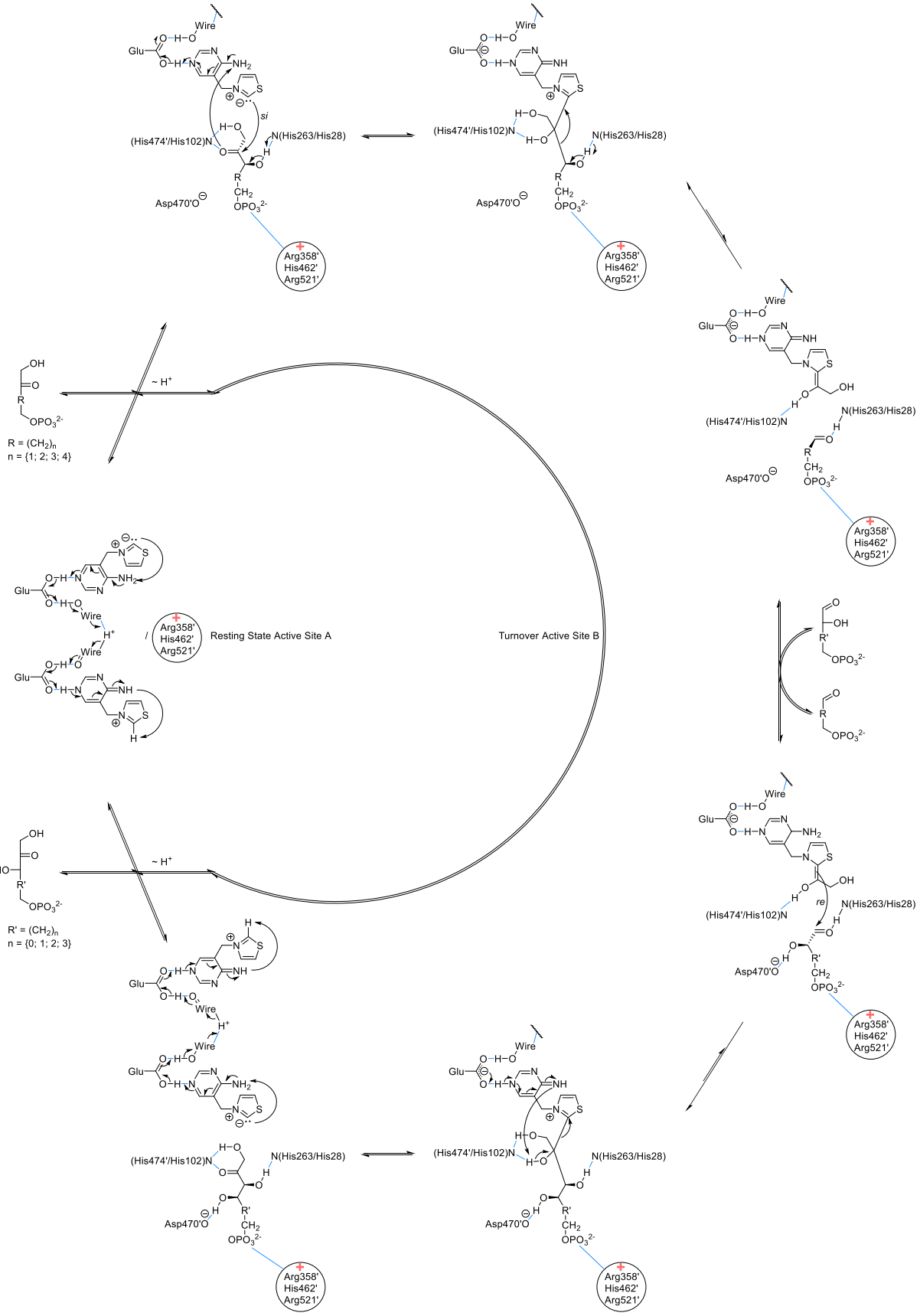


Figure 24: Molecular mechanism of transketolase. Hydrogen bonds and phosphate charge interaction displayed as blue dotted line. ThDP side chains not displayed, histidine nitrogens not in the correct ionization state. For more information about pyrimidine binding glutamates and the “Wire” see text. Based on (Yoshihara et al. 2023) [114] and (Dai et al. 2019) [132].

As mentioned before, the catalytic competence of the enzyme is enabled by the protonation of the pyrimidine N1, leading to tautomerization to the imino-chionid-form of the pyrimidine ring, deprotonating the C2 of the thiazole ring, forming an ylide, the active form of the cofactor. Upon exposing the apoenzyme to the coenzymes, these quickly bind to the peptide ($t_{1/2}$ = 250 ms). Subsequently, the enzyme undergoes slow spatial rearrangements and forms the catalytically competent active site ($t_{1/2}$ = 200 s). [129,133] For that reason, transketolase is often preincubated with the cofactor. The catalytically competent enzyme with the bound cofactors is more stable against denaturating influences, like elevated temperature. [134]

In the first step of catalysis, the active site binds the ketose substrate, forming the Michaelis complex. Due to the described hydrogen-wire, the wild-type transketolase then covalently binds the substrate in a single phase, without detectable intermediates. [132]

The exact nature of this step is under debate and can potentially proceed with the thiazolium activated as an anion (zwitterion, electron pair transfer, EPT), carbene or non-Kekulé π-diradical (single electron transfer SET). [135] In this work, the mechanism is displayed as in most publications, proceeding *via* the anion (EPT). First step is the attack of the thiazole ylide on the carbonyl and protonation of the oxo-group by the pyrimidine 4-amino group. The exergonic formation of the covalent adduct leads to a distorted conformation, with the thiazol-C2-ketose-C2 bond bent by up to 30 ° and the scissile C-C- bond almost perpendicular, both relative to the thiazole plane. This distorted conformation renders the electrons of the scissile bond readily absorbed by the thiazole π-system (maximum overlap), raises the internal energy of the tetrahedral intermediate and is additionally stabilized by H-bonding interactions of the C1 and C2 oxygens of the ketose substrate with His474’ (468’), His102 (104), and the pyrimidine 4’ amino group. The high steric strain of this intermediate represents a driving force for the C-C bond breaking, leading to strain release. [133,135,136] The formed dihydroxyethyl thiamine diphosphate (DHEThDP) enamine intermediate probably exists in a rapid equilibrium with the tetrahedral intermediate, only trapped in decarboxylation reactions. [115,129] The proton of the newly formed aldehyde, prior representing the 3-hydroxy group, is abstracted by the active site base (His263). If there is no acceptor present, the newly formed aldehyde stays in the active site and in the rapid equilibrium.

For irreversible donor substrates (hydroxypyruvate, removal of the formed aldose) and if there is no acceptor present, slow glycolaldehyde, and subsequent erythrulose formation can be observed. [129] The binding of the intermediate and subsequent fast equilibration of the enamine adduct with the dihydroxyethyl thiamine diphosphate (DHEThDP) with the equilibrium on the side of the enamine is believed to usually prevent side reactions of this high energy intermediate.

Upon exchange of the newly formed aldose through an acceptor substrate, the protonated active site base (His263) provides a proton and activates the aldehyde for the attack of the (E)-enamine. The aldehyde is held by the enzyme in a favourable conformation, reaching for a Bürgi-Dunittz trajectory on its *re*-face. The strained adduct can then be released from the enzyme and recover the ylide. A new ketose product is formed and deprotonated by the pyrimidine 4-amine. Pyrimidine 4-amine protonation triggers pyrimidine N1 proton release and N1 protonation at the other active site, leading to ketose adduct formation.

The whole molecular mechanism represents an example for a classical ping-pong mechanism with a cooperative half-of-the-sites reactivity. The net-reaction is equivalent to a benzoin condensation, which is an “umpolung reaction”. Rate determining step is probably the first C-C bond breaking or the formation of the tetrahedral, highly strained intermediate. [125,129,133]

It should be noted that most sugar substrates are available in solution in cyclic, non-reactive form (*e.g.* fructose-6-phosphate: 2.5% open form in water) and that multiple hydrogen bonding interactions in the active site probably favour ring opening. [133]

The described enzyme classes do not only present a lever to manipulate amino acids and turn them into high-value products, but also represent a gateway for the transformation of amino acids into sugars. And they are well suited for combination in one-pot, or the employment in otherwise interconnected reactions, so called enzyme cascade.

For more information about substrates and process parameters, visit the discussion part of this work.

2.6 Enzymatic Reaction Cascades

The meaning of the term enzymatic cascade has shifted quite substantially during the last two decades. In 2001 Mayer, Kroutil and Faber [137] still defined an enzymatic cascade as an enzymatic domino reaction, with the first reaction of a linear reaction sequence being catalysed by an enzyme. This first reaction would then produce a high energy, unstable reaction intermediate, which is not easily caught and analysed, and would undergo spontaneous subsequent reaction(s), for example cyclization or decomposition. Thus, the definition was quite narrow. Cascades as interconnected reactions in one-pot were rather named tandem reactions. Almost twenty years later, the term cascade has become a buzzword with a very broad meaning, nonetheless underlining the great importance and potential of these kinds of reactions. In a recent review about biocatalysis from industry perspective [138] an artificial enzyme cascade is merely described as a synthetic biochemical pathway, in other words: the connection of at least two chemical transformations, with at least one involving an enzyme. However, it should be kept in mind that most biochemists still refer to cascade reactions only, if one starting material undergoes at least two transformations.

2.6.1 Types of Cascade Reactions

Cascades can be grouped according to several properties, among them, the type of involved reactions (*e.g.* redox), the type of involved catalysts (*e.g.* purely enzymatic, combining bio- and metal- catalysts/chemical agents, involved spontaneous steps…), if the cascade is artificial, or occurs in nature, if the cascade is carried out in vivo, or in vitro, and even according to the process engineering (*e.g.* one-pot, telescoping, flow, compartmentalization, immobilization…).

Another important way to classify cascade reactions is by the mode of reaction interconnection. This is conclusively summarized by Schrittwieser *et al.* [139]

According to his review, there are in general six different ways to interconnect individual reaction steps, not all of them falling into the classic definitions of cascades: linear, orthogonal, cyclic, parallel, convergent and divergent (Figure 25).

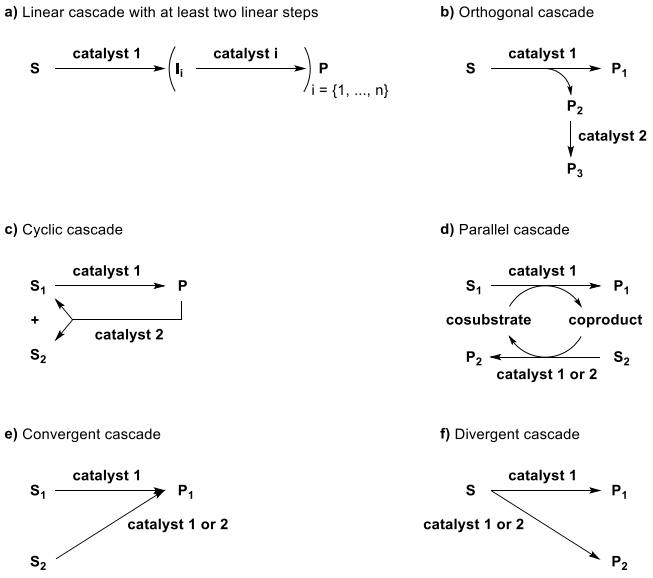


Figure 25: Types of biocatalytic cascades, concerning possible reaction interconnections. S: Substrate, P: Product, I: Intermediate Product. In general, catalysts can also be organic/metal catalysts, chemical reagents or void (spontaneous reactions), as long as at least one biocatalyst is present in the cascade. After Schrittwieser *et al.* with modifications. [139]

Some examples and explanations (Figure 25a - f) of these widely employed modes of reaction-interconnection are given hereafter:

a) Mostly associated with the term reaction cascade, linear interconnection is often applied to build up molecular complexity, avoiding intermediate isolation. There are many well-known examples presented in this thesis (*vide infra*).
b) In the oxidation of amino acids with oxidases, developing hydrogen peroxide is often quenched with catalase, or used to form dyes in colorimetric assays with horseradish peroxidase, a classic example for orthogonal cascades.
c) Many dynamic kinetic resolutions fall into the category of cyclic cascades, often applying an enzyme and an unselective chemical agent. (*e.g.* dehydrogenase/TEMPO, [140] oxidase/borohydride, [141] lipase/Pd(0)-catalyst [142])
d) Cofactor (cosubstrate) recycling is mostly achieved through a parallel reaction, for example for NADH recycling with glucose or formate dehydrogenase. The concept cascade can be applied on the cosubstrate, in its recycling process, undergoing multiple transformations.

e) Convergent reactions occur *e.g.* during building up large molecules from multiple smaller building blocks, as described for ε-caprolactam synthesis from cyclohexanone and 1,6-hexanediol employing a monooxygenase and an alcohol dehydrogenase. [143] In general, they are not classified as cascades, but are often part of cascade reactions.

f) Divergent reactions occur, when more than one product is formed. Since usually, applications are developed to be selective, the term is mostly applied to multiple individual reactions, or reaction systems, when combining multiple different catalysts on the same substrate for production of diverse products (*e.g.* for library synthesis). [144] In different context, it is more common to refer to unwanted divergent reactions as unselective reactions, producing side-products.

All these concepts can be utilized to categorize cascade reactions carried out in this work. As already mentioned, most biochemists use the term cascade only for a, b, and c.

It is clear that a lot of reaction sequences applying to the broad definition of enzymatic cascades can be found in nature and that natural cascades often inspire synthetic pathways (biomimetic, bioinspired). These natural cascades often use reaction sequences that seem unintuitive to the usual retrosynthetic thinking of a classical chemist, but also for this very reason still hold huge potential for the discovery and synthesis of useful, complex molecules. Among classical natural routes for building up complex molecules are polyketide, peptide, terpene, nucleotide and glycosyl synthesis. Many complex natural products are built up through domino reactions, triggered inside specific enzymes, where initial high energy intermediates are produced, already in specific conformations. These high energy intermediates (*e.g.* radicals, strong nucleophiles…) undergo directed multistep domino reactions and then are selectively terminated by precise quenching inside the enzyme`s active site. Such domino reactions are usually classified according to the chemical mechanism of the reaction step, initiating the domino cascade (nucleophilic, electrophilic, pericyclic, radical) and generate a large number of complex molecules. Collected secondary metabolisms of natural organisms lead to an estimate of 500.000 complex natural, often oligo- or macrocyclic products, for example isoprenoids, phenylpropanoids, alkaloids, polyketides and non-ribosomal peptides. Obviously, many of these are interesting pharmaceutically active compounds (*e.g.* antibiotics, anti-cancer drugs, hormones…)

and one challenge of the upcoming decades will be to translate some of these natural cascades into industrial applications. [145]

2.6.2 Importance of Biocatalytic Cascades

Enzymatic cascade reactions have not only become increasingly interesting for industry due to their huge dormant potential in discovering and efficiently producing active pharmaceutical ingredients with high molecular complexity. There are also several advances that become increasingly important, while transforming our industry into a more sustainable one and getting more independent from fossil feedstock.

The most obvious advancement of a linear reaction cascade is that intermediate isolation is avoided, saving labour, waste (mainly solvent) and energy.

However, circumventing intermediate isolation and conducting reactions as cascades or interconnected reactions in general also has additional benefits. For example:

- Keeping the concentration of a specific high energy intermediate low can prevent it from decomposing or undergoing unwanted side reactions. If one of the catalysts is inhibited or otherwise deactivated by an intermediate, this can also be prevented.
- Maintaining the concentration of reactive, often toxic intermediates low leads to a desired increase of safety in the process, sometimes even giving industry access to reactions otherwise not considerable.
- Directly channelling a reactive substrate into the next reaction step and the use of specific enzymes can make protection groups unnecessary, increasing atom economy.
- Certain expensive substrates can be recycled in cascade reactions, for example redox equivalents and other cofactors.
- If reaction steps of a linear cascade involve reductions and oxidations, redox equivalents can be passed between them, making the overall process redox self-sufficient, or redox-neutral. The term hydrogen-borrowing, referring to net-reaction equations, is misleading in this context, since no hydrogen molecules or hydrogen radicals are participating in the process. Instead, hydride-borrowing is suggested.

- Divergent reactions can generate large numbers of structurally related molecules, thus are suitable to build up whole libraries of compounds that can later be used for screening.
- Specifically interconnecting individual reaction steps can allow analysis of otherwise not-accessible reactions. For example, the production of a dye from the oxidase side product H_2O_2 can allow simple colorimetric screening setups. These coupled assay methodologies are often suited for in process monitoring and high-throughput applications as viable tools for process-control and protein engineering, respectively.

All these factors either save a lot of labour, energy, cost, or waste, for example in the form of solvents and high energy reactants, often making otherwise necessary process steps (*e.g.* purification of intermediates), or additional reactions (*e.g.* protection of functional groups) obsolete. One-pot cascades with reactant recycling drastically increase the atom economy of the involved reactions and high specificity of enzymes can obviate the need for auxiliary agents for reaction optimization or product isolation.

Additional to the advancements inherent in cascade reactions, enzymes themselves hold a lot of perks that become increasingly attractive during the aforementioned transformation of the traditional chemical industry. Furthermore, enzymes are perfectly suited for cascade reactions, since they are evolved to work under compatible conditions: aqueous systems, ambient temperature and pH. Benign reaction conditions are sought-after not only because they increase the compatibility of catalysts and are considered safe, but also since they are more environmentally friendly and associated with lowered energy consumption and are thought of as a greener form of chemistry.

The application of biodegradable enzyme catalysts from renewable feedstocks (*e.g.* amino acids, sugars), instead of often toxic and hazardous reagents and metal-/organo-catalysts and their preference for natural feedstocks and products also fits well into the concept of green chemistry (Table 1).

Table 1: The 12 principles of green chemistry, mnemonically summarized fitting to the word “productively”. [146]

1: **P**	Prevent waste	Prevention of waste, instead of clean-up.
2: **R**	Renewable materials	Employment of renewable raw materials and feedstocks.
3: **O**	Omit derivatization steps	Avoidance of derivatization methodologies, like protection group chemistry.
4: **D**	Degradable chemical products	Design of products, to prevent persistence in the environment. Aiming for final degradation to innocuous substances.
5: **U**	Use safe synthetic methods	Design and employment of chemical methodologies, reagents and products for minimization of toxicity to human health and the environment.
6: **C**	Catalytic reagents	Preference of (selective) catalysts over stoichiometric reagents.
7: **T**	Temperature, Pressure ambient	Acknowledgement of energy requirements and minimization of environmental and economic impacts. Preference of ambient conditions.
8: **I**	In-Process monitoring	Real-time in-process monitoring and control for prevention of the formation of hazardous substances.
9: **V**	Very few auxiliary substances	Avoidance of auxiliary substances, like solvents and separation agents.
10: **E**	*E*-factor, maximize feed in products	Maximization of incorporation of raw materials into final product and atom economy.
11: **L**	Low toxicity of chemical products	Design of products to preserve efficacy, while reducing toxicity.
12: **Y**	Yes, it is safe	Application of substances and methods that minimize the potential for chemical accidents including releases, explosions, and fires.

Additionally, enzymes are desirable catalysts, because they are usually very chemo-, regio-, and stereoselective. The latter stereoselectivity being of ever rising significance in the pharmacological, agrochemical and food industries, due to divergent effects of different enantiomers on biological organisms. This high selectivity of enzymes is considered one of the main driving forces for the development of enzyme-based industrial production.

More and more enzymes are discovered by means of proteomics, metabolomics and building up of metagenomic libraries. Also, the availability of these enzymes is always expanding as new databases are becoming available to the public, gene sequencing and DNA plasmid synthesis services are available at ever falling, low prices (Carlson curve) [147] and commercial kits of new enzyme classes are introduced into the market.

Most wild-type enzymes are not directly applicable in industrial processes, due to low turnover numbers, low stability, undesired selectivity, and the general top-down approach of industry, first choosing the product and then looking for possibilities to make it. However, the development of new methods in microbiology (*e.g.* directed evolution, [148] incorporation of non-canonical amino acids, [149] in-vitro protein synthesis [150]), bioinformatics (*e.g.* molecular dynamics simulations, [151] in-silico screening, [152] de-novo enzymes [153]) and analytical sciences (*e.g.* UHPLC, [154] droplet microfluidics [155]) has led to an immense acceleration of protein engineering, nonetheless remaining the bottleneck for the application of biocatalysts in industry. Thus, industries increasingly gain access to tailor-made catalysts on timescales satisfying their need-for-speed. The goal is to be able carry out one round of enzyme evolution within one to two days. The new swiftness in microbiology could be well-observed in the development of vaccines during the COVID-19 pandemic.

As already mentioned, enzymes bring forward new retrosynthetic disconnects and functional group interconversions (FGI, *e.g.* late-stage C-H functionalization), expanding the toolbox of synthetic chemistry.

And, last but not least, the conversion of an established process into a biocatalytic one holds the potential to generate new intellectual properties, the basis for value addition.

It should not be concealed that besides the huge benefits an enzymatic cascades-based chemical production has to offer, there are also a lot of challenges that still need to be addressed.

Besides protein engineering and the incorporation of biocatalysts into chemical processes still being slow compared to classical chemical methods, often the limited substrate scope and substrate concentration, tolerated by enzymes are being criticized. Additionally, aqueous and benign conditions can possibly hinder process development, for example because high temperatures might be desirable or due to the high energy demand of water evaporation.

But not only difficulties intrinsic to enzymes can be challenging in the development of bio-enzymatic cascades in industries. Increasing complexity of cascade reaction systems can create additional hurdles and problems arising in a late stage of development, lead to endless optimization and sometimes preventing the system from being viable at all. This can possibly lead to high investment costs, scaring off some of the players that might otherwise benefit from developments in the field. That is why implementing biocatalysis at an early stage of development is so important. Instead of switching to an enzyme catalyst, when a step in the classical chemical route fails or proves undesirable, biocatalysis should be implemented at stages as early as product design and retrosynthesis. Cascades should then be established in an iterative design (product design, retrosynthesis, cascade design)-build (enzyme selection/discovery, cascade implementation, process design)-optimize cycle (protein and process design, flux optimization). [138]

Many of the challenges arising from the use of enzymes and biocatalytic methodologies can also be overcome through process design. Prominent examples are flow-reactors, [156] immobilization, [157] telescoping, or *in situ* product removal [158] and compartmentalization techniques [159].

2.6.3 Examples of Biocatalytic Cascades

And there are already many newly developed industrial applications or methods close to industrial application with outstanding performances. [160] The most prominent example is probably the new route to Islatravir (4'-ethynyl-2-fluoro-2'-deoxyadenosine) a possible agent for the treatment and prophylaxis of HIV (Figure 26).

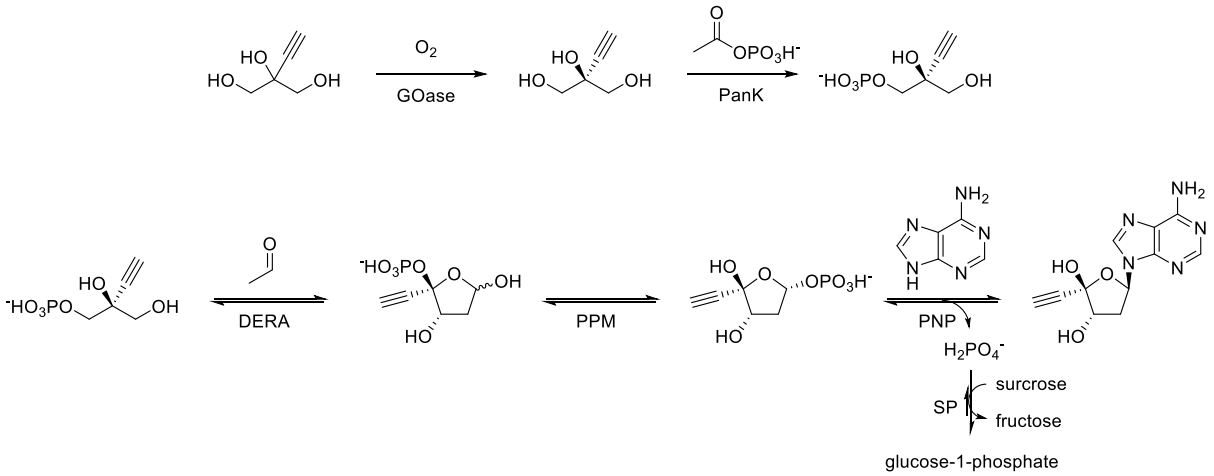


Figure 26: Biocatalytic part of the new eight step route to Islatravir. For enzyme abbreviations and description, see the following paragraph. [138]

It has been established by Merck & Co and involves two bio-enzymatic cascades: the production of 2-ethynyl-D-glyceraldehyde-3-phosphate from 2-ethynyl glycerol combining an oxidase (GOase) and a kinase (PanK) in a telescoping fashion and a four step one-pot cascade, building up Islatravir using deoxyribose-5-phosphate aldolase (DERA), phosphopentamutase (PPM), nucleoside phosphorylase (PNP) and a sucrose phosphorylase (SP) to remove the liberated phosphate and drive the equilibrium to the product. The bioenzymatic production reduces the number of necessary steps from between 12 and 18 to eight, yields the product in excellent enantioselectivity, with an overall yield of 51% in the biocatalytic steps and a single aqueous reaction stream is carried through the two bio-enzymatic cascades. Five of the employed enzymes were subjected to multiple rounds of enzyme engineering, improving their activity, stability and selectivity. The technology has already been applied for the synthesis of another antiviral agent and possible treatment for COVID-19, Molnupiravir. [138,161,162,163,164]

Another example, combining biocatalytic steps with classical chemical ones, developed by Pfizer is a multikilogram-scale synthesis of chiral intermediates of a γ-

secretase inhibitor, a possible agent for the treatment for Alzheimer's disease (Figure 27). [165]

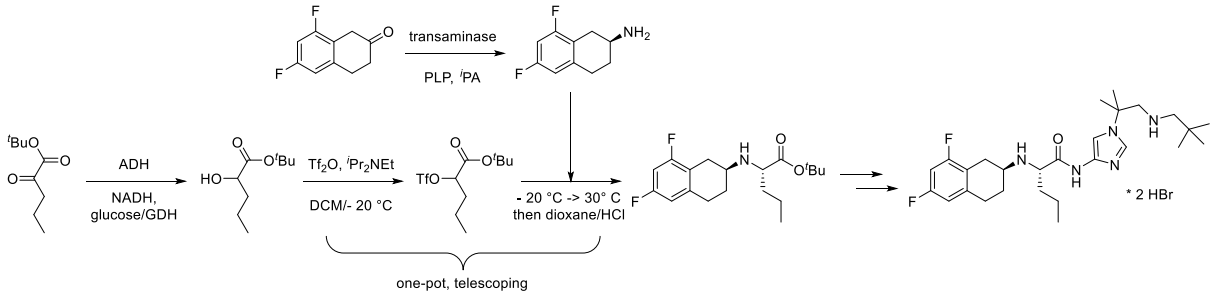


Figure 27: Route to the chiral intermediate of a γ-secretase inhibitor. ADH: alcohol dehydrogenase, GDH: glucose dehydrogenase. [165]

Even though the process is carried out in a sequential fashion, some of the individual steps can be considered cascade reactions.

The one-pot synthesis of *myo*-inositol from maltodextrin (from starch) uses four enzymes from thermophiles (Figure 28). [166]

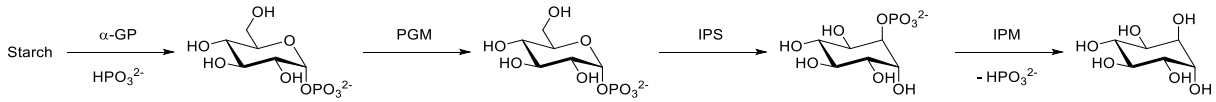


Figure 28: One-pot synthesis of *myo*-inositol from starch. α-GP: alpha-glucan phosphorylase from *Thermotoga maritima*, PGM: phosphoglucomutase from *Thermococcus kodakarensis*, IPS: inositol-1-phosphate synthase from *Archaeoglobus fulgidus*, IPM: inositol monophosphatase from *Thermotoga maritima*. [166]

So, no wonder more and more large chemical companies are discovering this biocatalytic potential and are building up new biochemical departments. Big international companies, like Merck, Pfizer, GlaxoSmithKline and Astra Zeneca are investing millions in the field. Also, German chemical firms like BASF are venturing in the area. For example, Wacker, a German company with long-standing tradition in silicone-based products is building up new biocatalytic research and development facilities in Munich, among others, for biotin, taurine, and oligosaccharide production and to be equipped and ready for future expansions into the field.

3 Research Objective

3.1 Main Objective

The goal of this work was to find new, unique, and synthetically useful amino acid racemases, extending the biocatalytic toolbox and demonstrating their usefulness in biocatalytic cascades.

3.2 Specific objectives

Main objective of the thesis was to screen a focused metagenomic library of 56 novel amino acid racemases for thermostable enzymes and unusual substrate scopes and to employ potentially useful racemases, demonstrating their synthetic potential. The enzyme panel had been constructed at Prozomix Ltd. for potential industrial applications. It became available to us in a BMBF-funded collaboration “Tralaminol”, directed at the development of novel methods for the synthesis of chiral amino alcohols.

3.2.1 Screening of the Obtained Racemase Panel

In the context of the main objective, firstly reliable colorimetric assay methods had to be tested and adapted to the chosen high-throughput multititer plate-based format for the rapid evaluation of new racemases.

Substrates are defined as interesting if

- they are usually not substrates of well-known groups of PLP-dependent amino acid racemases,
- they are not, or bad substrates of the selected model broad spectrum racemase from *Vibrio cholerae*
- they are not canonical amino acids, or
- if they possess unusual properties, for example very large, or permanently charged side chains.

Additionally, serine was deemed interesting in view of utility for envisioned coupling with transketolase.

For selected interesting substrates that are not commercially available, or too expensive ways to synthesize these in-house had to be found.

For them to be deemed synthetically useful, enzymes have to be reasonable stable, so the panel racemases also had to be evaluated for thermostability, employing modern Nano-DSF instrumentation.

3.2.2 Demonstrating the Synthetic Potential of Amino Acid Racmeases

When an amino acid racemase with interesting substrate spectrum and sufficient stability is found in a screening, its synthetic potential should be demonstrated in a dynamic kinetic resolution cascade reaction, to prove its usefulness and confirm the screening results. Enzymes, which were envisioned to potentially couple discovered racemases to are D-amino acid oxidase (DAAO), dehydrogenases and transketolase for the preparative synthesis of chiral building blocks and rare sugars. In the context of “Tralaminol”, rare sugars can subsequently be applied as substrates for the synthesis of rare amino alcohols This is meant to demonstrate amino acid racemases’ broad field of application in biocatalytic cascades.

Additionally, to be able to employ amino acid racemases in the preparation of rare deoxyfructoses, new routes and methods for the enantiospecific synthesis of the transketolase acceptor substrate, 2,4-dihydroxybutanal and equivalent protected γ-lactols were sought.

4 Results and Discussion

4.1 Colorimetric Screening of Prozomix Racemase Panel ProRCMS

The received metagenomic library of amino acid racemases shall be referred to as the racemase panel ProRCMS (Prozomix RaCeMaSe). Particular enzymes are identified with numbers from 1 to 56 (ProRCMS#1 to #56). The enzymes were obtained as cell-free extract of *E. coli* BL21DE3, in which the individual racemases were expressed, following Prozomix' in-house standard protocols with little or no optimization for expression of individual enzymes. Racemase genes were selected based on the Prozomix metaGRASP^{TM} technology, in this case with a primary metagenomic library generated, based on an *in silico* similarity search with the broad-spectrum racemase from *Vibrio cholerae* (BsrV) as a model sequence. The similarity search was conducted within the Prozomix metagenomic database, generated by assessing complex microbial communities from diverse sources, among them organisms from the deep sea.

4.1.1 Development of the Screening Method

The following chapters describe the development and principle of the employed semi-high throughput assay for the assessment of amino acid racemase activities on diverse canonic and non-canonical amino acid substrates. For those non-canonical amino acids, not commercially available, the synthesis is described. Also, the exploration of additional panel enzyme features, like thermostability and dehydratase side-reactivity, is presented. The obtained results are then summarized in an accessible colour-coded fashion.

Subsequent chapters contain the description of synthetic applications of the summarized analytical results.

For the substrate screening, a two-step telescoped, colorimetric plate reader assay was chosen, mainly due to low costs, its high throughput, possible degree of parallelization, and low requirements for analytical instruments. This type of assay is applicable on a large range of substrates, and importantly, directly on received cell-free extracts. The latter saved us a lot of time, otherwise required for the purification of small quantities of 56 enzymes.

The goal was to discover the most active racemases on individual, screened substrates, particularly unusual racemases, useful enzymes for selected subsequent applications, and general trends within the screened library. The screening was explicitly not conducted to elucidate every trace activity, or to fully characterize the contained racemases.

4.1.1.1 Assay Principle

The analytical procedure is divided into two separate steps. First step is the racemisation of L-amino acid, to generate D-amino acid, the subsequent analyte. In the second, telescoped step, a system of two coupled enzymes is applied in one-pot, D-amino acid oxidase (DAAO) and peroxidase. Here, the formed D-amino acid is oxidized, with DAAO as catalyst and molecular oxygen as oxidant. The stoichiometrically generated byproduct of the oxidation, H_2O_2, serves as a substrate for the peroxidase, which now catalyses oxidation of a chromogenic substance to form a dye, or fluorescent. Hence, the amount of generated dye, or fluorescent is proportional to the racemase and spectrophotometry delivers a measure for the racemase activity. The partition of the assay into two consecutive steps is based on a publication by Espaillat *et al.* [37] and its advantages are addressed in 4.1.1.2.

The principle of the second step of the two-step assay, to generate a dye or fluorescent through oxidation of a chromogen with H_2O_2 and peroxidase, is well known in literature and broadly applied. There is a multitude of possible chromogenic substances and mixtures. Two frequent examples are shown in (Figure 29).

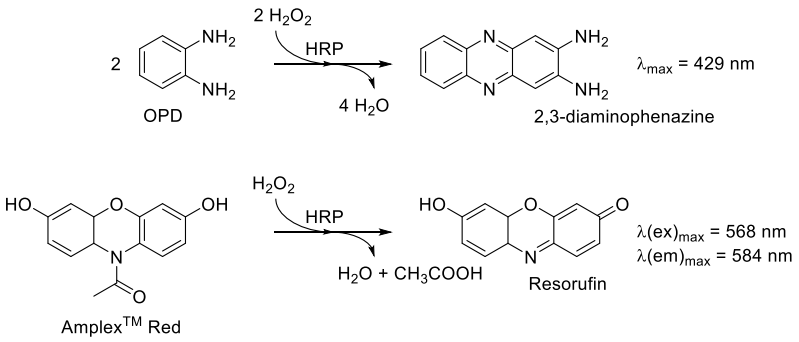


Figure 29: Dye and fluorescent formation through oxidation with horseradish peroxidase (HRP) and H_2O_2 of *o*-phenylenediamine (OPD) and Amplex™ Red, respectively.

The first example, *o*-phenylenediamine-oxidation (OPD) has been discovered early. [167,168] Many similar chromogens are applicable, mostly electron rich aromatic amines and alcohols, forming quinoid dyes upon oxidation. Amplex™ Red [169,170] is the current gold-standard in sensitivity due to the fluorescent properties of the formed

resorufin. However it is also more expensive than gold (244 € for 5 mg Ampliflu^{TM} Red; compared to < 0.5 € for 5 mg gold) [171].

The colour reaction, employed in-house for various spectrophotometric assays, is displayed in (Figure 30) [172,173], together with its successor.

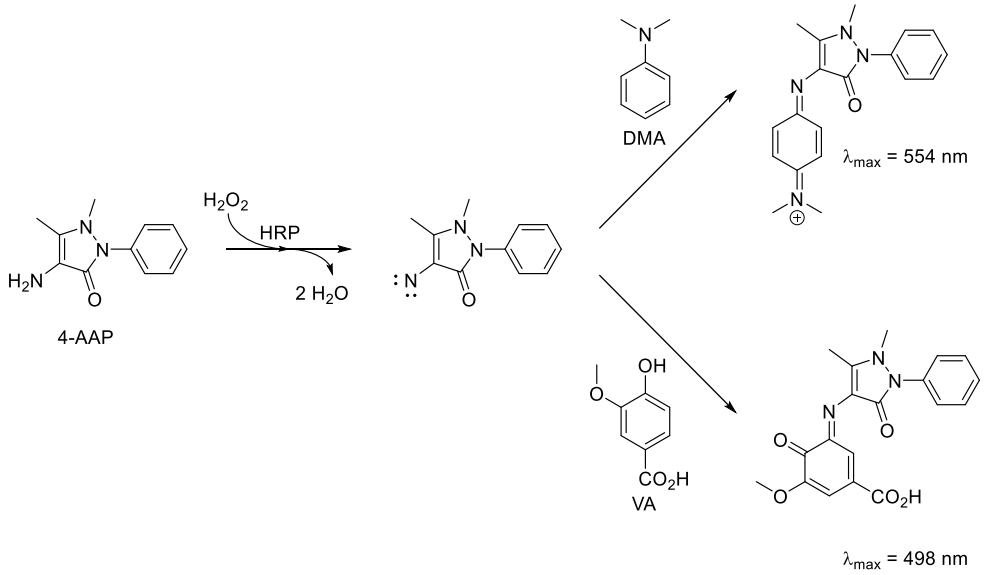


Figure 30: Dye formation upon 4-aminoantipyrine (4-AAP) oxidation with horseradish peroxidase (HRP) and H_2O_2. Displayed exemplary electron donors are dimethylaniline (DMA) and vanillic acid (VA).

The initial substrate for HRP, 4-aminoantipyrine (4-AAP) is used in many similar chromogenic systems, because it is selectively and stoichiometrically oxidized by HRP. Also, it is compatible with a multitude of different aromatic nucleophiles. Its application was initiated after its description in a publication from 1938, in which the dye-formation from 4-AAP in the presence of aromatic amines and oxidizing agents is described. [174] This has later been found to proceed very selectively through the action of peroxidase and H_2O_2 and with many electron-rich aromatic compounds as possible reaction partners for the oxidized 4-aminoantipyrine nitrene.

Based on a publication from Holt *et al.* [175] we were able to replace our previously used toxic dimethylaminopyridine, as the aromatic nucleophile, through non-toxic vanillic acid. Apprehensions that vanillic acid might, due to its analogy to benzoic acid, lead to oxidase inhibition were unfounded, since negative effects of vanillic acid as electron donor compared to other tested electron donors was not found. Additionally, the dye formed from vanillic acid and 4-aminoantipyrine was superior in terms of stability of the red colour, (data not shown) which enabled longer reaction times of the colorimetric

reaction and thus higher sensitivities, specifically for D-amino acids that were bad substrates for the DAAO (*vide infra*).

To further point out that this dye forming reaction is not only highly customizable in the formed dye, but also as a whole a versatile module in the analytical toolbox, the following paragraph will give a few more examples of its exploitation, one of them an in-house-developed four-enzyme, one-pot analytical cascade for the evaluation of novel, artificial sialoconjugates.

As already mentioned, enzymatic (horseradish peroxidase) dye formation is relatively robust and mostly generates an equimolar quantity of dye per mole of hydrogen peroxide and chromogen. Also, it takes place at benign enzymatic conditions, facilitating its versatile application in the detection of *in situ* formed hydrogen peroxide, or the detection of other essential assay components. For example, phenols (chromogen) can be detected in water, based on dye formation upon addition of hydrogen peroxide and horseradish peroxidase. [176] It is also modularly coupled to other reactions which stoichiometrically produce hydrogen peroxide *in situ* as a response to the analyte. The latter principle is often applied with various oxidases, which produce one equivalent of hydrogen peroxide upon oxidation of a substrate with molecular oxygen. Famous examples are *in situ* coupling to glucose oxidase for the determination of glucose levels in clinical samples [177], or DAAO for the enantioselective detection of D-amino acids. [178]

In a side project in cooperation with Dr. Alexander Mertsch, we exploited this modularity for the detection of sialidase activity on sialoconjugates (GM3 analogues), containing modified sialic acid derivatives. [179] Therefore, the dye forming peroxidase module was coupled with β-galactosidase *in situ*, glucose oxidase and two different commercial α-sialidases (Figure 31).

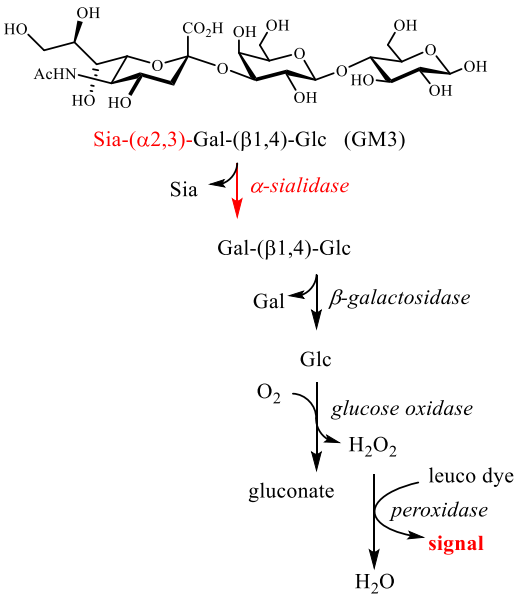


Figure 31: Assay principle: α-sialidase activity is probed by modular coupling of the peroxidase H_2O_2 detection to α-sialidase, β-galactosidase, and glucose oxidase. [179]

Likewise, bridging the peroxidase module-amino acid racemase-gap with peroxide forming DAAO, enables racemase activity detection. Starting from L-amino acids, racemized to the D-amino acid counterpart with the racemase, the DAAO aerobically oxidizes the D-amino acid product and stoichiometrically forms the H_2O_2-analyte as a side product. This principle has been exploited, for example in a stepwise fashion in microtiter plates [178] and on agar plates [180]. With purified, reasonably active enzymes, this racemase assay can also be employed in one-pot, as displayed in Figure 32.

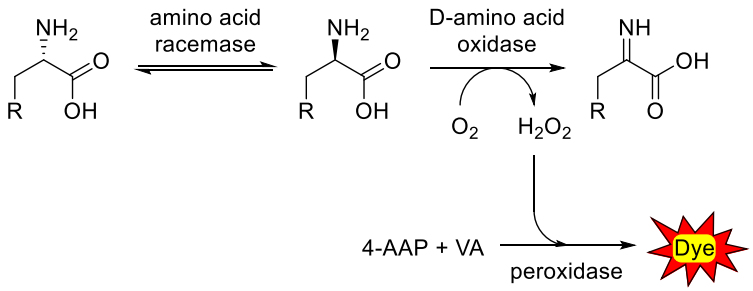


Figure 32: Principle of a directly coupled assay: Various L-amino acids can be racemized with amino acid racemase to the corresponding D-amino acid, oxidized to the imino acid with DAAO, forming the H_2O_2-analyte. Formed imino acids are hydrolyzed to the corresponding keto-acid in aqueous solutions. 4-AAP: 4-aminoantipyrin, VA: vanillic acid.

4.1.1.2 Specific Challenges, Solutions, and Scope of Application

One big challenge in the application of the assay on the racemase panel was the employment of the racemase enzymes in the form of cell-free extract. The received cell-free extract consists of lysed, freeze-dried *E. coli* cells, obtained from a culture medium without further extraction. This material contains *E. coli* enzymes, among them glucose oxidase and also *E. coli*'s own alanine racemase. [181] Besides proteins, cells and medium also hold small molecules, like sugars and small quantities of L-amino acids. Even if free amino acids would not be part of the medium and *E. coli* cytosol, they would slowly be formed from digestion of proteins and peptidoglycan through the action of *E. coli* proteases. Thus, conducting the screening with cell-free extract is expected to yield relatively large extinction readings in the substrate blanks (execution of the assay protocol without addition of L-amino acid). Also, the more active and versatile the heterologous racemase would be, the larger the expected extinction of the blank reading, due to a higher degree of racemisation of contained amino acids. This also leads to a large degree of variation within the extinction values for substrate blanks of individual panel cell-free extracts.

Also, it should be clear that due to the inherent capability of *E. coli* alanine racemase to racemise alanine, this assay is not fit to identify the best racemases for alanine racemisation due to the high background of alanine racemisation. In a hypothetic search for a racemase with high alanine turnover, complete alanine racemisation during our assay is merely an indicator for the expressed exogeneous racemase to be a candidate for in a subsequent screening with purified enzymes. However, due to the plethora of available alanine racemases and information about these and also due to relatively low interest in alanine racemisation, further examination of the alanine racemisation capabilities within the racemase panel was not conducted.

To avoid a further increase in the substrate blanks, the DAAO, known not to reach high expression levels in *E. coli* due to production of hydrogen peroxide, was employed either dialyzed, or His-tag purified. It should be noted that the assay is in principle also conductible in a directly coupled fashion in one-pot, with high loadings of purified enzymes, at least for highly active enzyme substrate combinations. Such a setup would also be suitable to determine maximum reaction rates and derive K_M and k_{cat}.

Another challenge arising from large blank extinction readings and low substrate concentrations (*vide infra*) is a resulting low sensitivity. This lowered sensitivity

prohibits the execution of the assay as an *in situ*, one-pot, directly coupled procedure. Therefore, the assay is conducted in a two-step fashion, in which the L-amino acid substrate is allowed to be transformed by the racemase for an extended period of time. In the second step, the amount of generated D-amino acid is estimated in the colorimetric reaction. Obviously, the advance of the racemisation drastically increases the sensitivity in the colorimetric reaction and thus allows us to perform the screening in a high-throughput fashion. It also enables us to identify the most active racemase candidates on a large array of substrates. However, there are also several downsides to this kind of protocol:

a) Instead of measuring the maximum velocity (v_{max}), conversions are estimated
b) There is a maximum value for the racemase activity (16,7 mU/mg), based on full conversion (racemate formation from 4 mM L-amino acid) of the substrate within the given time (30 min) and the given amount of racemase cell-free extract (4 mg/mL). There is no differentiation between racemases, reaching this upper limit.
c) Activities are underestimated if particularly high. v_{max} is only reached when the product concentration is apprx. zero, which only is the case for very low turnover.
d) Racemases are known to have different velocities in the back- and forth-reaction. Values are only obtained for the L- to D-amino acid direction.
e) The additional step provides further opportunity for experimental error and increases the magnitude of the experimental error, due to additional steps in the protocol and more pipetting.
f) The additional step takes more time, thus reducing the throughput of the assay.

It should be clear that a), b), c), e), and f) can be overcome with oxidase-mediated *in situ* product removal, application of purified enzymes and higher substrate concentrations. d) is an inherent property of the DAAO and using the same modular assay principle, it can only be overcome employing a complementary L-amino acid oxidase, for example available from snake toxins.

To accommodate for the inherent colorimetric response due to the use of cell-free extract, limit of detection (LOD) and limit of quantification (LOQ) had to be defined as

a threshold, this way filtering out false hits. Usually, the LOD is defined as the blank plus three times the standard deviation of the blank. The LOQ is defined as the blank plus between six and nine times its standard deviation. To be able to estimate the standard deviation of the blank, substrate blanks were produced in quintuplets for four different racemase cell-free extracts (ProRCMS#27, 29, 40, 55) from the panel. The standard deviations were between 2 and 5%. Thus, a collective standard deviation was defined as 5%. This value is relatively low, considering the many steps in the assay handling. Fluctuations in blank readings between individual cell-free extracts are usually much larger than 5%. Thus, for each racemase, a substrate blank was measured and values beneath “substrate blank” * 1.15 were defined as not hits (LOD defined as three times 5% of the individual blank reading of the corresponding cell-free extract). Values between “substrate blank” * 1.15 and “substrate blank” * 1.3 are defined as not quantifiable, but hits. Since this way, LOD and LOQ are solely dependent on the enzyme and not on the substrate (substrate blank), hits from substrates that are very bad substrates for the DAAO are hardly detectable.

The range of fluctuation in between individual cell-free extracts were not only large due to the presence of different racemases. To improve the assay handling, a so called “colouring solution” was priorly mixed, containing the DAAO, the horseradish peroxidase, 4-aminoantipyrin, vanillic acid and FAD. The premixed colouring stock is not stable, with time it self-stains and forms precipitates and therefore had to be prepared freshly at least twice a day. To prevent wastage of large amounts of purified DAAO and commercial horseradish peroxidase, there were always only 5 to 10 mL prepared at once. The preparation such solutions necessitates weighing tiny amounts of the constituents, thus producing 0 to 5% variation of the weight of each component in the composition of the premix. For this reason, a variation in the colorimetric response, depending on the individual colouring solution, is expected. To prevent large errors, a two-point calibration (0 mM D-amino acid, 4 mM L-amino acid and 0.4 or, 0.8 mM D-amino acid and 3.6 or, 3.2 mM L-amino acid respectively) for each individual substrate is measured with each individual colouring premix.

For two amino acids (N^{ε}-acetyl lysine, azidohomoalanine), there was no commercial D-amino acid standard available, thus no two-point calibration was included. Since absorption values with the best racemases for these substrates reached values corresponding to full conversion with other substrates and since the colorimetric

reaction responds to the universal coproduct hydrogen peroxide, formed equimolarly in the oxidation of any D-amino acid, these maximum values were set as full conversion. Other results on these two substrates were calculated accordingly. Unfortunately, this leads to an absence of the correcting factor generated from the two-point calibration, accommodating for individual colouring premixes and cell-free extract properties. Consequently, the estimated error in individual results for these two amino acid substrates are larger and values that are similar to each other should in general be compared with care. Generally, the total uncertainty of the obtained results is estimated to be larger than 10%. For result summary, colour coding is used to give a comparative overview of activities, rather than quantitative, numerical tabulation. Another benefit from the colour coding is the fast and intuitive understanding comparison of a large array of results and a simplified comparison of the properties of individual racemases.

The need to stop the racemisation reaction after the first step that is, before conducting the quantification of D-amino acids, and the way this was done is another issue that should be considered when interpreting the results. The 200 mL-reactions of the racemisation step were stopped by heating the reaction mixture to 97 °C for 5 min. This method of inactivation was chosen, because it leaves the composition of the resulting supernatant relatively intact, while efficiently denaturing most racemases. It should be noted that the most thermostable racemases from the panel may not have been fully deactivated, which was deliberately accepted. The potential error is considered to be small because (i) supernatants were cooled between the two individual assay steps and only warmed again shortly prior to the second step; (ii) the timespan in between the two assay steps was kept short. (30 min – 1 h) Also, thermostable racemases are desired catalysts for several reasons, and therefore an activity overestimation of thermostable enzymes during this assay is deemed a particularly acceptable error.

Two possible errors arising through excessive heating of the racemisation reaction mixture are concentration effects, caused by water evaporation and unspecific chemical racemization, inducible through the presence of basic amino acids and PLP. To avoid concentration effects, vials were kept tightly closed during heating and condensate was returned to the supernatant during the subsequent refrigerated centrifugation. It is known that heating of amino acids with PLP and basic reagents can cause racemisation, however the rate of this reaction is rather low and usually

autoclaving for extended durations is required. [19,20,182] So the possible error arising from heating the reaction mixture is considered very low, which was confirmed in cell-free extracts with low racemisation activity (*vide infra*).

Aside from thermal deactivation, there are multiple other methods available for enzyme deactivation. Most frequently described in literature is probably the addition of a strongly acidic or basic reagent (mostly HCl or trichloroacetic acid). However, these reagents would also possibly lead to increased unspecific racemisation and may affect the enzymes in the subsequent assay step. Neutralization would demand an extra effort and the resulting high ion strengths would lower the essential oxygen solubility.

Other methods for enzyme deactivation would also lead to large errors for similar reasons. For example, extraction with organic solvents (chloroform) may have a low reproducibility in small volumes and lead to extraction of non-polar amino acids (tryptophan). Ultrafiltration was ruled out for economic and ecologic reasons.

In reminiscence of the goal of our screening, elucidation of trace-activities of the expressed racemases is not of importance. Rather, the most active racemases on the individual substrates in the screening, unusual racemase candidates for interesting subsequent application, and the identification of general trends within the received panel were to be discovered. For this purpose, we consider the chosen method most suitable.

Categorization of the chosen assay method can be done according to Sheludko *et al.*. [183] Following their classification, the assay relies on coupled signalling by a universal coproduct. Taking a closer look makes it clear that this is only true for the D-amino acid as the substrate. Since our interest is not the measuring of the oxidase but the racemase activity, the assay represents a special case, where the reaction, which produces the universal coproduct, is part of a telescoped reaction sequence with the enzymatic reaction of interest being first step.

Other methods for detection of racemase activity, based on the enantioselective detection of D-amino acids include (i) coupling to D- or L-specific amino acid dehydrogenases with spectrophotometric measurement of NAD(P)H at 340 nm, (ii) polarographic measurement of the D- to L-amino acid proportion, or (iii) chiral separation on HPLC, often including derivatization (*e.g.* Marfey's reagent). None of these methods offer similarly broad opportunities to assess a broad substrate spectrum

with such a high degree of parallelization. However, these are alternative valuable tools for result confirmation and can be used to overcome some limitations of our protocol.

4.1.2 Thermostability of Amino Acid Racemases

Protein thermostability of the panel racemases was examined with differential scanning fluorimetry (nanoDSF), employing a Prometheus NT.48 instrument (NanoTemper Technologies, Germany).

4.1.2.1 Method Principle

In differential scanning fluorimetry, a change in the fluorescence of a dissolved molecule is examined during a continuous change in the temperature of the sample. This technique can be applied to study the unfolding of a protein that contains tryptophane and tyrosine residues, because their inherent fluorescence emission wavelengths change upon a change in the polarity of their spatial surrounding (environment-sensitive).

For example, tryptophane absorbs light at 280 nm and in a polar environment, like an aqueous solution, the maximum of the fluorescence emission lies at 305 nm. In a non-polar setting, like the hydrophobic core of a protein, the emission maximum can be bathochromically shifted up to 353 nm ($\Delta\lambda_{em} \approx 50$ nm). [184]

The Prometheus nanoDSF machine allows analysis of samples of only 10 µL protein containing samples, which are filled into glass capillaries and placed into the sample holder of the instrument. For tryptophane containing proteins, sample concentrations of less than 200 µg/mL protein are sufficient, while for proteins only containing tyrosine, the concentration should be about 5 to 10 times higher.

To satisfy our need for speed, the cell-free extract of each racemase was dissolved in potassium phosphate buffer at 1 mg/mL, equivalent to reach about 300 µg/mL of amino acid racemase. The fluorescence was read continuously at 330 and 350 nm.

For data evaluation, either the quotient of the fluorescence intensity at 350 nm to that at 330 nm was used, or if delivering a clearer signal, the shift in either of the two intensities was used.

Some samples yielded very low signals, where no clear point of unfolding could be deduced. This was probably due to bad expression, missing tryptophane residues or no sharp point of unfolding. No unfolding temperature is given in these cases.

4.1.2.2 Thermostability of Amino Acid Racemases from the Literature

As stated above, amino acid racemases are essential enzymes, present in many microorganisms. As such, their thermostability is known to be diverse. Since most terrestrial organisms are mesophilic, most racemases have their activity maximum around 35 °C and become deactivated above 45 °C. However, there are also psychrophilic and thermophilic enzymes known. [185] For example, the alanine racemase from *Bacillus stearothermophilus* has its maximal activity at 50 °C. Therefore it can be purified similarly to the same organism's transketolase by incubation of the supernatant at 70 °C for 1 h. [186] The template racemase from *Vibrio cholerae* can be deactivated by heating the containing solution to 80 °C for 5 min. [37]

4.1.3 Screening Substrates A: Canonical Amino Acids

Two rounds of screening were conducted on two different substrate panels. In the first round, the racemase panel was tested with most of the commercial canonical amino acids (Figure 33), except glycine, asparagine, isoleucine, and threonine.

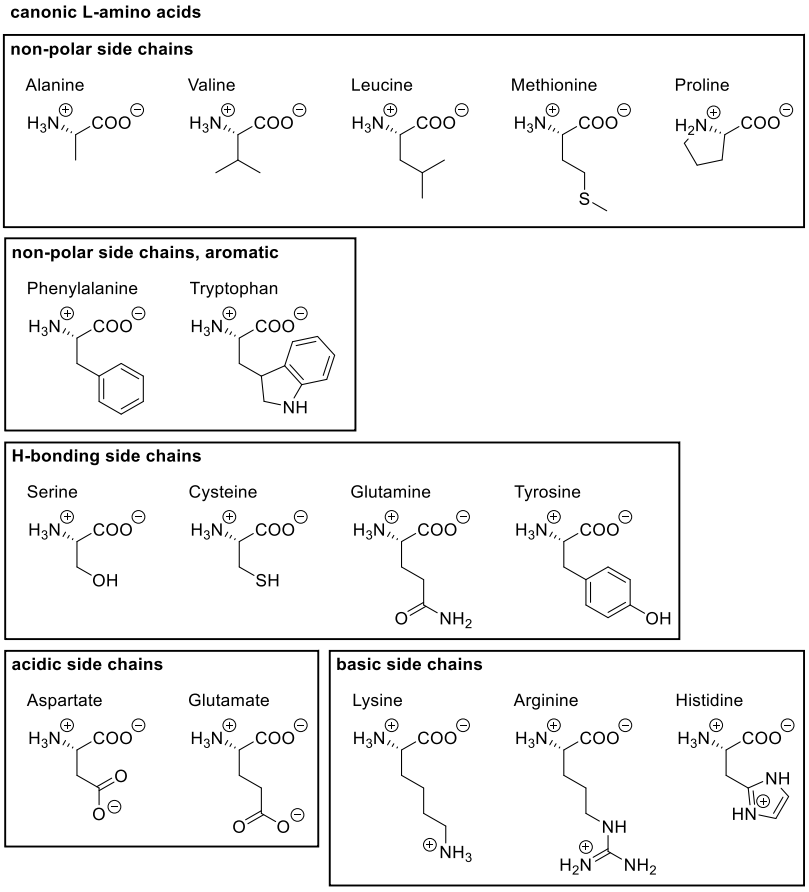


Figure 33: Canonical amino acids, included in the screening.

Asparagine and isoleucine were not immediately available in-house and thought to be close relatives to glutamine and leucine, respectively. L-Threonine would not be converted into D-threonine but, epimerized into D-*allo*-threonine.

Tryptophan has low solubility in water and its light absorption was found to overlap strongly with wavelengths used in the coupled assay. Therefore, it is not compatible to the selected assay and excluded from the screening. To evaluate this substrate in an aqueous solution with dmso as a cosolvent, direct polarographic measurement of the amino acid racemisation is suggested. Its bulky side chain is believed not to fit into the active site pockets of most racemases.

Our first round of screening already delivered interesting hits, which were confirmed in follow-up experiments. It was clear that some of the racemases had broad substrate scope and also should convert unusual substrates. For this reason, it was decided to conduct a second round of screening with non-natural, non-canonical amino acids, to probe the permissible substrate limits of these promiscuous enzymes. Some of the selected unnatural amino acids were not commercially available at reasonable prices and were thus synthesized in-house.

4.1.4 Screening Substrates B: Non-Canonical Amino Acids

4.1.4.1 Non-Canonical Amino Acid Substrate Array

Figure 34 shows the selected non-canonical amino acids and the corresponding numeric identifiers.

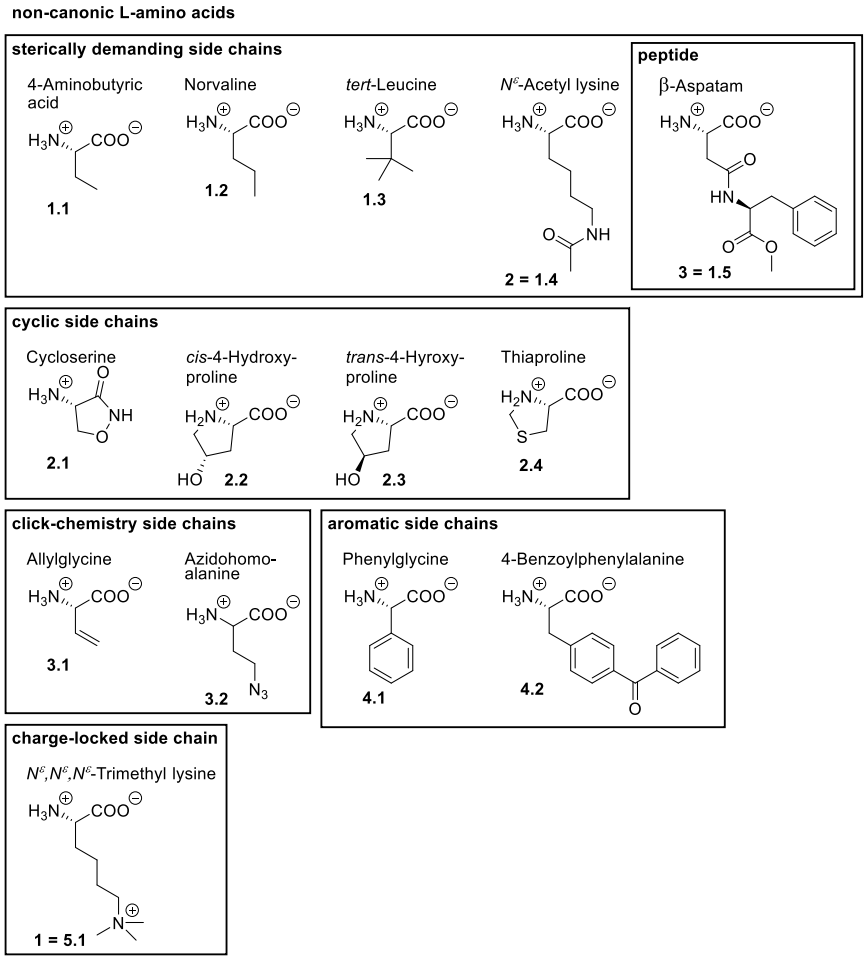


Figure 34: Non-canonical amino acids, included in the screening and their corresponding numeric identifiers. The three synthesized amino acids have two numeric identifiers, a single-number- identifier corresponding to the number to the molecules during synthesis and a two-digit identifier, corresponding to the racemase screening.

Among the selected non-canonical amino acids were sterically demanding ones, like **1.1** (L-2-aminobutyric acid), **1.2** (L-norvaline), and **1.3** (L-*tert*-leucine), and **1.4** (*N*ε-acetyl-L-lysine), the two clickable amino acids **3.1** (L-allylglycine) and **3.2** (L-azidohomoalanine), two additional aromatic amino acids, **4.1** (L-phenylglycine) and **4.2** (L-benzoylphenylalanine), and several cyclic and heterocyclic amino acids: **2.1** (cycloserine), **2.2** (*cis*-) and **2.3** (*trans*-4-hydroxy-L-proline), and **2.4** (thiaproline). **5.1** is the charge-locked analogue of lysine, trimethyl-L-lysine. The dipeptide **1.5** (β-aspartame) has its peptide linkage between the β-carboxylate of aspartate and the amine of phenylalanine methyl ester, holding a free α-amino acid functionality in the

aspartate moiety, thus representing a special case of a sterically very demanding asparagine derivative.

Not all amino acid substrates in the ProRCMS panel screening were commercially available. Some were very expensive, but considered to be easily synthesized. The amino acids, successfully produced in-house, are depicted in Figure 35.

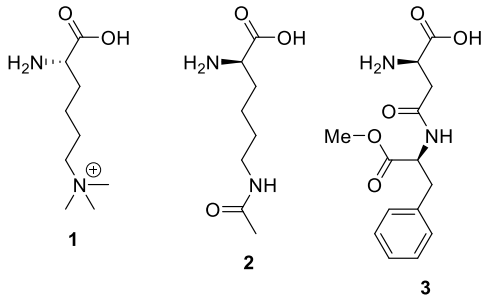


Figure 35: Amino acid substrates N^{ε},N^{ε},N^{ε}-trimethyl-L-lysine (**1**), N^{ε}-acetyl-D-lysine (**2**), and methyl D-α-aspartyl-L-phenylalanine (D-β-aspartame, **3**) for the ProRCMS panel screening, produced in-house.

4.1.4.2 N^{ε},N^{ε},N^{ε}-Trimethyllysine (**1**)

The quaternary, charge-locked trimethylated lysine derivative was produced by following a procedure of Al Temimi et al., [187] reacting the protected N^{α}-(*tert*-butyloxycarbonyl)-L-lysine with iodomethane and potassium carbonate in a mixture of DCM and methanol at 40 °C (Figure 36).

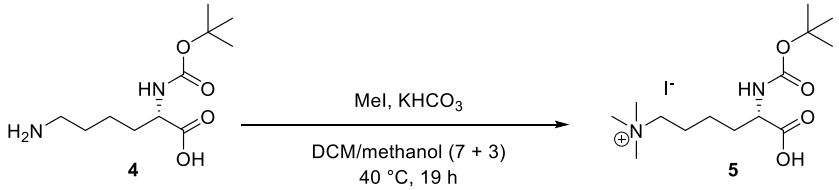


Figure 36: Preparation of quaternary ammonium salt (**5**) of Boc-protected L-lysine (**4**).

The reaction was monitored by HPLC, following the stepwise methylation, and by TLC. Dragendorff's reagent was applied to TLC plates, specifically dying quaternary ammonium salts with intense orange colour.

The reaction was set up, carefully dissolving the protected amino acid educt (**4**) by the addition of methanol to the mixture, before adding iodomethane, to precipitate the ammonium salt product (**5**). However, only little of **5** crystallized. Therefore, the crude product was purified *via* C-18 reverse phase chromatography, yielding 74% of the hygroscopic material as its iodide salt. Contrary to the original protocol, **5** was

deprotected after purification by the addition of formic acid. As a consequence, **1** was collected as the di(hydroformate) salt (Figure 37).

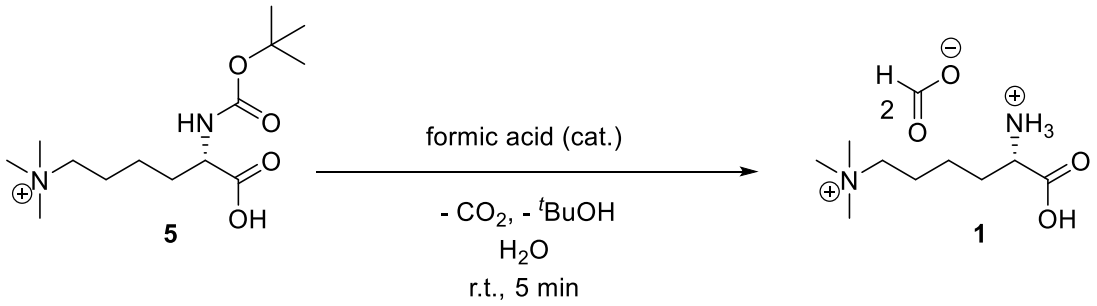


Figure 37: Deprotection of N^{α}-(*tert*-butoxycarbonyl)N^{ε},N^{ε},N^{ε}-trimethyl-L-lysine.

The final product was collected with 58% yield from the deprotection step. Losses are attributed to the purification using a normal-phase silica gel column. The highly polar compound was eluted with a mixture of chloroform, methanol and formic acid (9 + 1 + 0.1, V + V), yielding the corresponding formate salt. To reach higher yields, the product may either be isolated *via* reverse phase chromatography, or by crystallization, or employed without further purification after removal of volatiles and neutralization.

The overall yield was 43%. Losses during the methylation step are considered acceptable due to the stepwise methylation and the necessary separation from other derivatives, and the yield of 74% in this step seems reasonable. However, the low yield of 58% in the deprotection should be avoidable with a more appropriate purification strategy.

Nonetheless, the desired product was generated with good purity and suitable for application in the assay.

4.1.4.3 N^{ε}-Acetyl-D-lysine (**2**)

The protected N^{α}-(((9*H*-fluoren-9-yl)methoxy)carbonyl)-D-lysine (**6**) was available in-house. Three different catalysts (all at 5 mol% eq.) were tried for the acetylation with acetic anhydride in DCM at small scale (0.5 mL). Standard Einhorn-like (see Einhorn *et al.* - Schotten-Baumann reaction, Einhorn variant) [188] conditions with pyridine and the reaction with $AlCl_3$ both delivered full conversion within 30 min, while the reaction with CsCl only showed about 50% conversion based on TLC and HPLC/MS. To avoid pyridine separation, $AlCl_3$ was selected as catalyst for the preparative conversion. However, the preparative reaction with only $AlCl_3$ did not lead to full acetylation after

30 min, thus pyridine was added, leading to full conversion within one hour. (Figure 38).

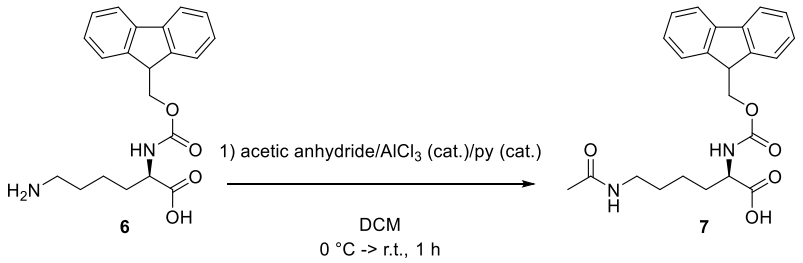


Figure 38: Acylation of **6** (Fmoc-protected D-lysine).

Acetylated Fmoc-D-lysine (**7**) was purified *via* silica column, yielding **7** in 90% yield, and deprotected on demand with piperidine at small scale to yield free **2** (Figure 39).

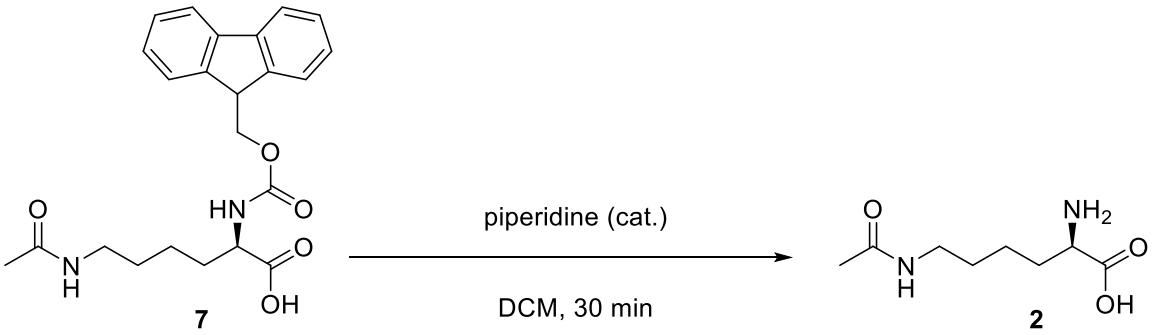


Figure 39: Deprotection of **7** (N^{ε}-acetyl-N^{α}-(((9*H*-fluoren-9-yl)methoxy)carbonyl)-D-lysine).

With this method, the desired acetylated amino acid is generated in high purity, on demand. Both steps were monitored on TLC and HPLC/MS, enabling separation and detection of the molecule masses.

4.1.4.4 Methyl D-β-aspartyl-L-phenylalaninate (**3**)

Methyl L-α-aspartyl-L-phenylalaninate is an artificial sweetener, called aspartame and present in many processed foods. It is a small dipeptide with an aromatic residue. Its β-aspartyl derivative leaves the α-amino acid part of the aspartyl moiety intact, thus theoretically could be racemized by racemases. Small peptides are considered interesting large substrates for the screening of racemases, exploring their limits in accepting such bulky substrates. Structural features of aspartame (relatively large, phenyl group, peptide, methyl ester) not only makes it an interesting substrate for racemases, but also enables relatively easy HPLC-analysis of the formed diastereomeric products of a “racemization”. The opportunity to prove suspected hits from the screening with HPLC is valuable in this case, because the corresponding

racemization product is expected to be a very difficult substrate for DAAO, thus to deliver very low signals in the screening.

The racemization of commercially available methyl L-β-aspartyl-L-phenylalaninate with amino acid racemases is expected to yield **3** (methyl D-β-aspartyl-L-phenylalaninate), the enzyme only racemizing the intact α-amino acid part of the aspartyl moiety.

A standard for application in the screening and HPLC analysis was produced from commercially available α-*tert*-butyl *N*-*tert*-butyloxycarbonyl-D-aspartate (**8**) and methyl L-phenylalaninate employing 1-[bis(dimethylamino)methylene]1*H*-1,2,3-triazolo[4,5-*b*]pyridinium 3-oxide hexafluorophosphate (HATU) as common high efficiency coupling reagent. [189,190,191]

The reagent reacts with the carboxylate, forming an activated ester (**9**) and liberating tetramethyl urea (Figure 40).

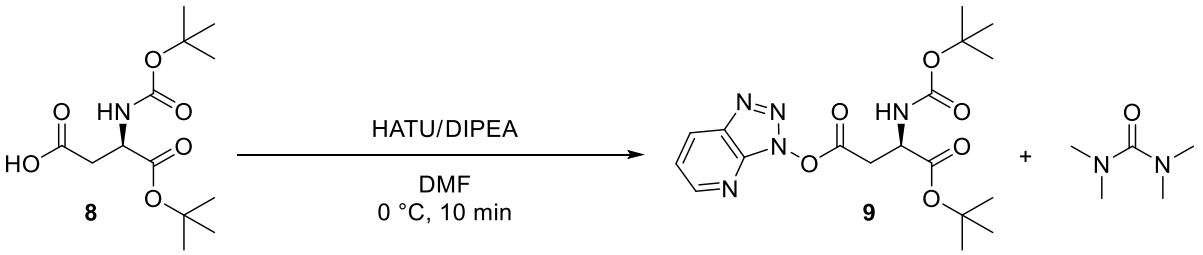


Figure 40: Carboxylate activation with HATU.

The active-ester (**9**) does not only possess an excellent leaving group, but also favors the addition of amines, which are added to the one-pot reaction sequence after the formation of **9** is complete according to TLC analysis (Figure 41).

Figure 41: Coupling of the active-ester with methyl L-phenylalaninate.

The extraordinary efficiency to activate the acid for amine acylation is attributed to the pyridine nitrogen's ability to form a hydrogen bond with an added amine's proton.

The rather apolar coupling product **10** was easily purified *via* column chromatography and deprotected with TFA in DCM (Figure 42).

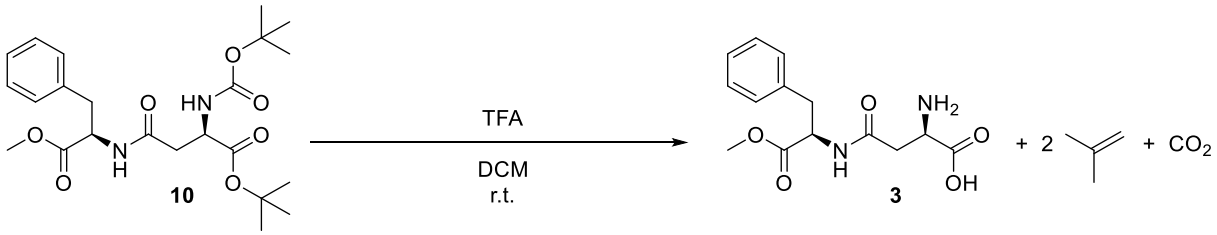


Figure 42: Deprotection of methyl α-*tert*-butyl *N*-*tert*-butyloxycarbonyl-D-aspartyl-L-phenylalaninate.

Removing all volatiles, this delivered the desired methyl D-α-aspartyl-L-phenylalaninate (**3**) in high yield (95%) and purity, ready for application in the assay.

4.1.5 Expected Results from the Screening

4.1.5.1 Substrates Not Expected to Deliver Hits

Several amino acids that are known not to be active with either racemases or oxidases, or that are considered strong inhibitors were included in the substrate panel (Table 2) as kind of a control, expected not to deliver hits.

Table 2: Amino acids expected not to deliver hits in the screening, not active enzyme and explanation.

Amino acid	**enzyme, inactive on the substrate**	**further information**
cycloserine	racemase	known irreversible inhibitor for many racemases [192]
proline (also *cis*-, *trans*-4-hydroxy-proline, thiaproline)	racemase	Not substrate for PLP-dependent racemases. Secondary amine does not form external aldimine. Separate class of PLP-independent proline racemases. All proline derivatives are expected not to be substrates (*cis*-, *trans*-4-hydroxyproline, thiaproline)
aspartate, glutamate	oxidase	known to be very bad substrates for DAAO from *Rhodotorula gracilis*
cysteine	peroxidase	L-cysteine is a peroxidase inhibitor

These compounds were still included, because (i) they are considered useful negative controls for the staged assay system, and (ii) hits among these assumed negative controls would be interesting for the following reasons:

- Conversion of secondary amines hints towards a PLP-independent racemase
- Results contradicting these assumptions would potentially deliver new insights into the assay mechanism and further expand our knowledge about the assay

- Confirmed hits within the negative controls contradict published results, hinting towards unusual molecular features of the catalyst, offering starting points for mechanistic exploration
- A racemase, not inhibited by cycloserine, potentially equips its host with a resistance against this antibiotic

4.1.5.2 Substrate Spectrum of the Employed D-Amino Acid Oxidase from Literature

Since activity detection during the assay relies on D-amino acid oxidation with DAAO, in this case from *Rhodotorula gracilis*, a requisite for racemase activity detection with a specific substrate is the ability of this oxidase to act on this substrate.

Table 3 displays a literature assortment of relative activities of DAAO from *Rhodotorula gracilis* in different preparations.

Table 3: Relative activities of different preparations of DAAO from *Rhodotorula gracilis* on the corresponding D-amino acids, measured under different conditions, with various assay techniques. [193,194,195,196,197,198,199,200]

D-amino acids	[193]	[194]	[195]	[196]	[197]	[198]	[199]	[200]
non-polar								
Glycine		3	> 0	0			3	
Alanine	100	71	100	100	100	100	100	100
Valine	86	95	127	75	33	34	60	135
Leucine	75	57		89	94	51	86	85
Isoleucine	48	56		72			60	
Methionine	127	100	82	142	114	65	113	82
Phenylalanine	97	79	99	108	102	94	94	102
Tryptophan	109	56	103	106	151	97	92	115
Proline	50	57	83	69	29	1.3	46	82
polar non-charged								
Serine	44	41		51	42	33	49	
Threonine	41	10		14	19	21	15	
Tyrosine		26		15		73	18	
Cysteine	106	1		0	30	0	18	
Asparagine	64	40		68	30	104	38	
Glutamine	70	53		86	35	52	58	

basic								
Lysine	7	5	7	6	12	0	8	7
Arginine	6	2	6	21	20	1.9	19	7
Histidine	6	58		88	68	5.7	60	
acidic								
Aspartate	0		0	0	0	2.5	3	0
Glutamate	0	4	0	0	0	5.7	3	0
non-canonical								
Phenylglycine	29							
Thiazolidine-2-carboxylic acid	17							
cephalosporin C		1					36	110
Norvaline		103						

What strikes the eye is that these literature values are quite divergent. One reason for this is that different assay methods and conditions were used.

It should also be kept in mind that some amino acids are not soluble at given concentrations under the standard assay conditions, for example tryptophan aqueous solubility is only around 2.5 mM. In some publications this is considered, in others ignored. Of course, a lower concentration might plausibly give a lower effective reaction rate in the assay.

Also, none of these assays exactly match our assay conditions. Even though some of the methods used are based on the detection of H_2O_2 through a coupled peroxidase reaction, there are differences to our method. Striking is the lack of 4-aminoantipyrine as primary electron donor, higher D-amino acid concentrations, and different temperatures and pH-values. Additionally, to elucidate the substrate spectrum of an oxidase, assay conditions are usually chosen to reveal oxidase activity differences, while in our assay, a large excess of oxidase is provided to effectively reach the end-point of this conversion at a high sensitivity. Thus, some response is expected under our conditions even with very bad oxidase substrates.

The summarized literature information clarifies that the oxidase is active on all neutral amino acids, except cysteine and glycine, has very low activity on basic amino acids and none, or at least very little, activity on the acidic amino acids aspartatic acid and glutamic acid. Hence, the screening should be suitable for a wide range of substrates. Due to a high oxidase concentration and other features of the assay conditions, it is expected to obtain hits across the whole substrate panel, even for very poor oxidase substrates. However, it is also expected that poor oxidase substrates (cysteine, aspartate, glutamate) would only yield very low signals and consequently, at very high detection limits.

The starting concentration of the L-amino acids was decided to be relatively low. The chosen 4 mM concentration for racemisation leads to a maximum concentration of 1 mM of analyte D-amino acid, considering the dilution upon mixing the reaction supernatant one to one with the assay solution. The concentration represents a compromise between sensitivity of the assay and comparability of the individual amino acids. Additional factors favouring a lower substrate concentration were the high cost or limited availability of some of the more exotic amino acids (*e.g.* azidohomoalanine) and the greater stability of amino acid solutions at lower concentrations. Increasing our confidence in the feasibility of low analyte concentrations are the low K_M values of the DAAO for the different amino acids (Table 4).

Table 4: Michaelis constants [mM] for DAAO from *Rhodotorula gracilis* on different D-amino acid substrates, measured with various assay techniques. [200,201,202]

D-amino acids	**[**200**]**	**[**201**]**	**[**202**]**
non-polar			
Glycine			
Alanine	0.83	0.8	0.83
Valine	18.89		18.89
Leucine	0.50		0.50
Isoleucine			
Methionine			0.18
Phenylalanine			0.29
Tryptophan			0.30
Proline	21.47		21.47

polar non-charged			
Serine			13.75
Threonine			
Tyrosine			
Cysteine			21.41
Asparagine			
Glutamine			
basic			
Lysine			
Arginine		18	
Histidine			
acidic			
Aspartate		18	
Glutamate			
non-canonical			
Phenylglycine			
Thiazolidine-2-carboxylic acid			1.99
cephalosporin C	1.82	5	
Norvaline			0.89
Norleucine			0.39

Also, even though these oxidases are considered very specific on D-enantiomers, it could not entirely be ruled out that large concentrations of L-amino acids may lead to unspecific reaction, lowering the assay specificity, also possibly arising from unspecific racemisation during the first assay step.

Last but not least, the assay was set up to find suitable candidates for subsequent cascade reactions with targeting complete substrate conversion, thus desiring racemases with low K_M values. Keeping the amino acid concentration in the racemisation step relatively low, logically emphasizes variants with low K_M over those with high K_M, but high k_{cat}, following the principle: “you get what you screen for.”

4.1.5.3 Substrate Scope of Different Amino Acid Racemases from the Literature

As already mentioned, the panel racemases were selected based on the sequence similarity with the broad-spectrum racemase from *Vibrio cholerae* as the model sequence. Due to this fact and the origin of the metagenomic library, all panel enzymes are expected to be structurally similar microbial racemases. Also, the different fold-types of racemases were already discussed in the theory part of this work. Since the target reference sequence of the metagenomic library, the broad-spectrum racemase from *Vibrio cholerae*, is a fold-type III PLP-dependent racemase, all panel enzymes are expected to fall into this category. The category contains six known enzyme families, grouped according to their substrate specificity and origin: bacterial alanine racemases, [203,204,205] bacterial serine racemases, [206] eukaryotic aspartate racemases, [207,208] bacterial lysine racemases, [209,210] bacterial histidine racemases, [211] and broad-specificity racemases [37,212.].

Table 5 displays characteristics in the substrate specificity of these racemase families, according to literature reports.

Table 5: Relative activities of different amino acid racemase classes on D-amino acids. (ser.: serine racemase, asp.: aspartate racemase, lys.: lysine racemase, his.: histidine racemase, bsr.: broad specificity racemase). [203,204,205,206,207,210,211,212]

	alanine racemase			ser.	asp.	lys.	his.	bsr.
D-amino acids	[203]	[204]	[205]	[206]	[207]	[210]	[211]	[212]
non-polar								
Alanine	100	100	100	100	0	0	0	6
Valine	13	2	4				0	0
Leucine	0	2	1			0	0	10
Isoleucine	2	2					0	0
Methionine	0	1	2			0	0	66
Phenylalanine	3	12	1			0	0	0
Tryptophan		8						0
Proline	3	1	2				0	0
polar non-charged								
Serine	0	2	28	23	0		0	10
Threonine	2	2					0	3

Tyrosine	6	4	0				0	0
Cysteine	0	2						0
Asparagine		10				0	0	12
Glutamine	0	5				0	0	54
basic								
Lysine	25	6	5			100	0	100
Arginine	50	7	6				0	79
Histidine	25	20	1			0	100	6
acidic								
Aspartate	0	1			100			0
Glutamate	0	2			0		0	0

From the literature data, it should be clear that most racemases are very substrate specific. For some alanine racemases, a few other amino acids than alanine are also substrates and they may be active especially on serine and basic amino acids. Broad specificity racemases may be distinct from others in their high activity on methionine and glutamine.

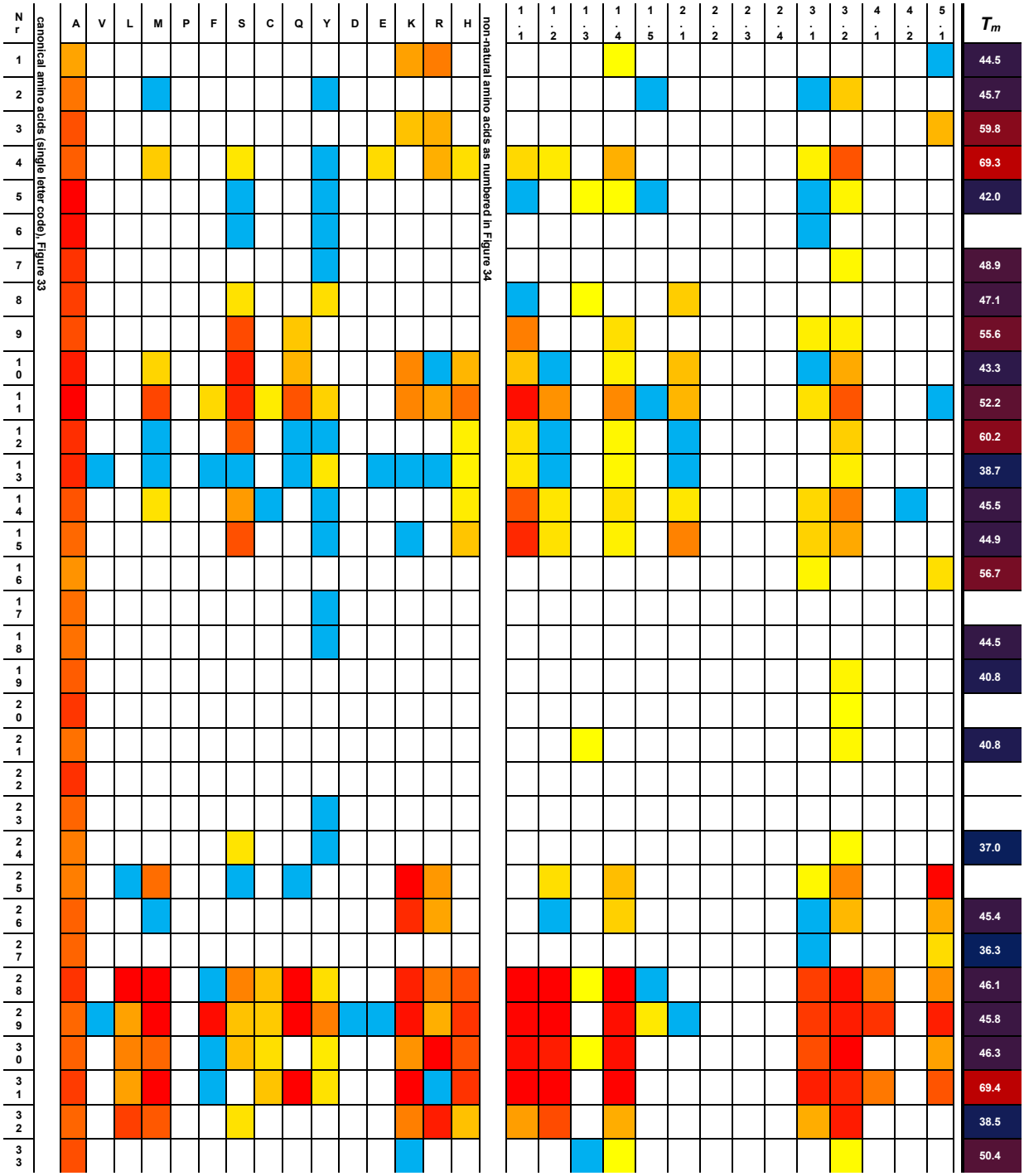
There are also enzymes that have sequences and structures similar to broad specificity racemase from *Vibrio cholerae*, but are not amino acid racemases. A rcsb.org-search gave several examples: glycerol-3-phosphate dehydrogenase (pdbid: 1YJ8, to be published), cystathione β-lyase (pdbid: 8DUY, to be published), serine dehydratase (pdbid: 3ANU and others) [213], D-threonine aldolase (pdbid: 4V15), [214] β-hydroxyaspartate aldolase (pdbid: 6RQA, 6QKB) [215], aspartate aminotransferase (pdbid: 1TOE and others) [216], L-methionine γ-lyase (pdbid: 3VK2 and others) [217], prephenate aminotransferase (pdbid: 5WMI and others) [218], cyclase OrfR (pdbid: 4M23 and others) [219], C-S lyase (pdbid: 6QP1 and others) [220], sugar aminotransferase (4ZAS, to be published), kynurenine aminotransferase I (pdbid: 1W7N and others) [221], and cysteine desulfurase (pdbid: 6O11 and others) [222]. Thus, depending on the nature of the conducted sequence similarity search leading to the identification of panel-associated enzymes, it can be expected to include enzymes without intrinsic racemisation activity and enzymes with significantly different inherent activities.

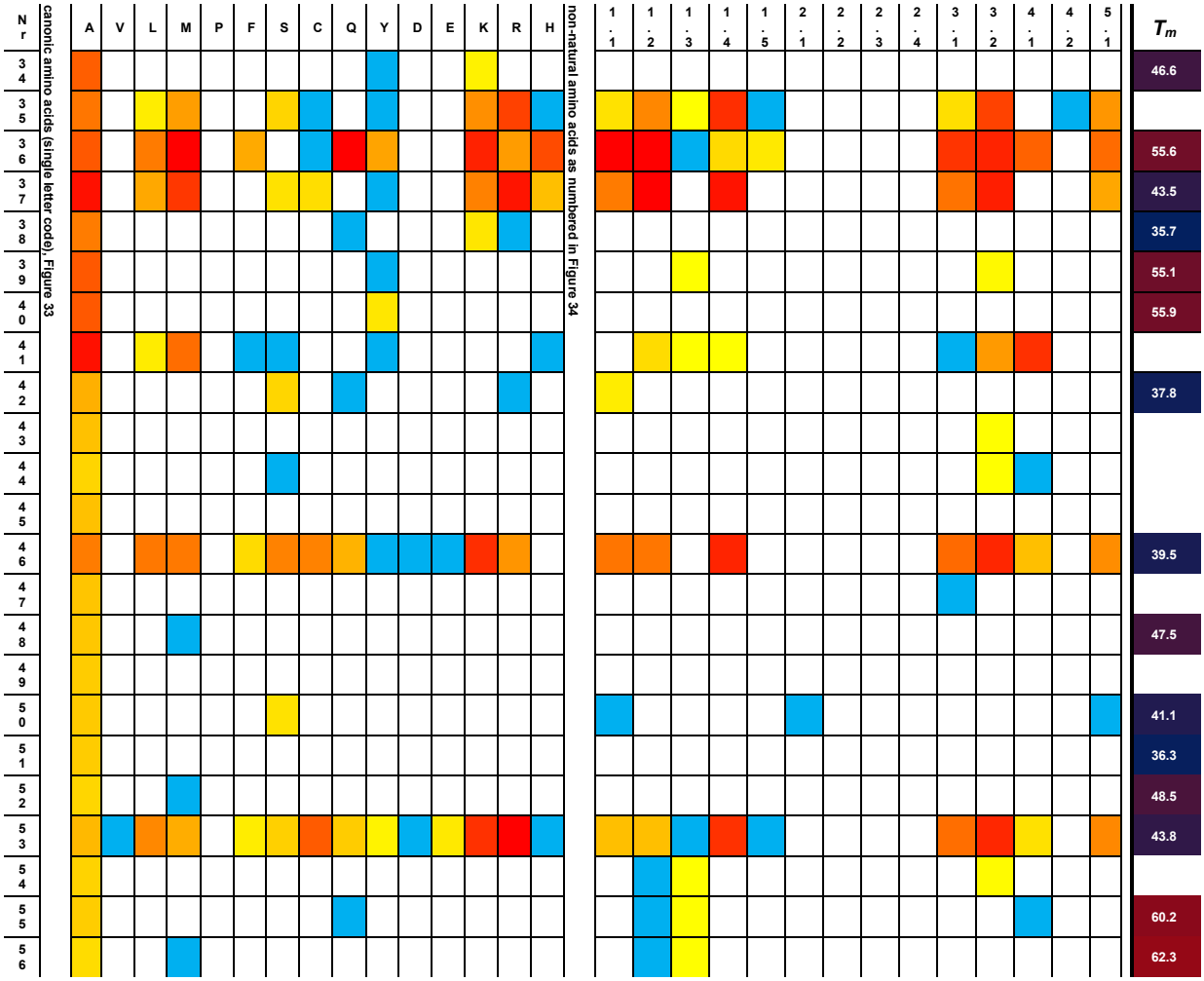
Finally, due to the Prozomix-approach, some racemases may be expressed at rather low concentrations, thus not delivering hits, or underestimating their activity. Corresponding SDS pages are provided in the Appendix.

4.1.6 Screening Results

Table 6 summarizes the collected data, as colour coded relative activities and colour coded melting points.

Table 6: Results of the coupled substrate screenings and DSF melting point determination. Data is color-coded with a linear gradient. Substrate screenings: white: no response above the limit of detection; blue: above limit of detection, but below limit of determination; values above the limit of determination, ranging from 1% conversion (yellow ♦) to 100% conversion (red ♦). Melting point determination: white: no distinct melting point could be observed from our nano-DSF experiment; values ranging from 35 °C (dark blue ♦) to 70 °C (dark red ♦). Nr: Racemases are numbered as by Prozomix Ltd. (ProRCMS#). Canonical amino acids are given in one letter code and non-natural amino acids are numbered according to Figure 34.





Firstly, it should be kept in mind that expression levels vary over the panel range, as shown in the SDS analysis (see Appendix). The found protein bands range from apprx. 35 to 50 kDa, fitting well to the expected mass range of about 40 kDa. ProRCMS#19, 43, 45, and 51 display no overexpressed enzyme and no activity aside from the expected alanine racemisation was not found.

Surprisingly, some of the racemases exhibit strong activities, even though only mediocre overexpression was detected in the corresponding SDS page, *e.g.* ProRCMS#10, or 32, indicating potential for optimization by enhanced expression levels.

4.1.6.1 Interesting Property-Groups within the ProRCMS Panel

From these results, the panel racemases can be assigned to certain categories and properties of which some are displayed in Table 7.

Table 7: Assortment of properties and classifications within the racemase panel.

Not active, not expressed, or alanine racemase	6, 7, 17 – 23, 39, 40, 43 – 45, 47 – 52, 54 – 56.
Promiscuous	4, 10 – 15, 28 – 32, 35 – 37, 41, 46, 53
Lysine racemase	1, 3, 10, 11, 25, 26, 28 – 32, 35 – 37, 46, 53
High activity on serine	9 – 12, 14, 15, 28, 46
High thermostability	4, 31
Active on a range of large substrates	5, 28, 35, 36, 53
Active on unusual substrates	4, 16, 25, 28 – 32, 53
Hits on cycloserine	8, 10 – 15, 29

The results show that the screening was suitable to discover a wide range of interesting properties in our racemase panel and as such represents a valuable tool for racemase evaluation.

One interesting property in the panel should be noted, namely that enzymes seem to be grouped in terms of their numbering, aka position in the panel. For example, the range from ProRCMS#28 to #32 represents an island of promiscuous, highly active racemases, or ProRCMS#47 to #52-racemases seem to be inactive, or alanine racemases. One could easily assume that this points towards a flaw in the screening, as the experimenter discovers certain properties on certain days of experimenting. However, to gain additional certainty in our results, the order of the screened racemases was randomized with the exception that racemase ProRCMS#42 to #56 arrived later and thus was not available in the first portion of the first round of screening. As such, it is assumed that these islands of enzymes with certain properties within the panel stem from the Prozomix *in-silico* screening algorithm and that the clustering is due to the phylogenetic organization of the panel candidates, which may group enzymes with similar properties.

Another factor boosting our confidence is the finding of activities in certain racemases on substrates similar to each other, even though some of these results were produced with long timespans in between. Examples are lysine and acetyl lysine, or phenylalanine and phenylglycine.

4.1.6.2 Broad-Specificity Racemases

Many of the panel racemases show a broad substrate spectrum, which is expected considering the reference for the *in-silico* screening also was a broad specificity racemase. Nonetheless, the abundance of promiscuous racemases is still surprising and makes these enzymes pristine candidates for further applications in the biotechnology field. There are also a lot of panel enzymes capable to racemise our first target serine, and some also racemise lysine and even histidine.

4.1.6.3 Thermostable Enzymes

Two enzymes were found, possessing a high thermostability (ProRCMS#4 and #31 with T_m = 69.3 and 69.4, respectively) with #31 as a very active and promiscuous enzyme even at 30 °C, thus holding a lot of potential for application and further enzyme evolution. For this reason, ProRCMS#31 was also selected for sequence evaluation and modelling (*vide infra*).

4.1.6.4 Unusual Substrate Scope

Some of the enzymes have very unusual substrate scopes. Some are able to racemise very large, or acidic amino acids, which was discovered even though the conversion to the corresponding oxo-acid through the peroxide producing DAAO is almost negligible. Therefore, such results, especially those with conversions below the limit of quantification, should be looked at with caution and require further confirmation.

4.1.6.5 Cycloserine

Besides β-aspartame, (*vide-infra*) the activity on cycloserine, which was found in a panel island (ProRCMS#10 – #15), seems especially unusual, since cycloserine forms an adduct with the PLP cofactor of racemases and as such represents an irreversible inhibitor. This is well known and studied, and due to this property cycloserine is even used as a broad-spectrum antibiotic agent. [223]

The current understanding of racemase deactivation by cycloserine is based on the formation of an irreversible aromatic isoxazole adduct with the PLP cofactor (Figure 43). [26]

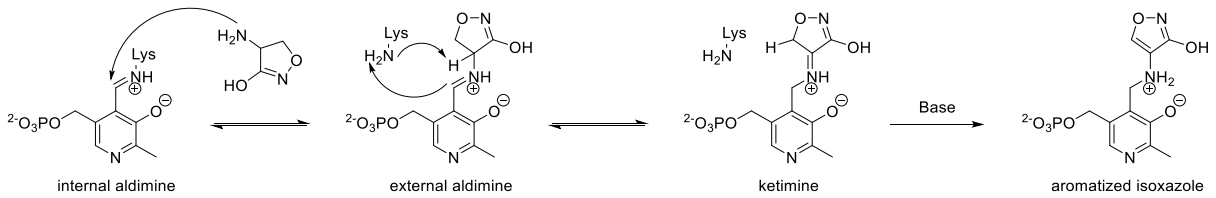


Figure 43: Transamination of cycloserine, leading to racemase inactivation. Last step is the irreversible aromatization and isoxazole formation of the adduct. [26]

Hence, for the valid occurrence of a hit in our assay, the inactivated PLP needs to be either restored or replaced and an oxidizable substrate needs to be formed for the DAAO. Since PLP was included in the reaction mixture, replacement is possible. Also, conversion of cycloserine into another amino acid prior to the assay reaction cannot be ruled out.

It was found that *Mycobacterium tuberculosis* exhibits alanine racemase activity after exposure to cycloserine and recent findings suggest that the isoxazole formation is not really irreversible due to imine-enamine tautomerization in aqueous solution and that a second reaction takes place upon exposure of PLP to cycloserine in buffer, which leads to ring opening and oxime formation (Figure 44). [192]

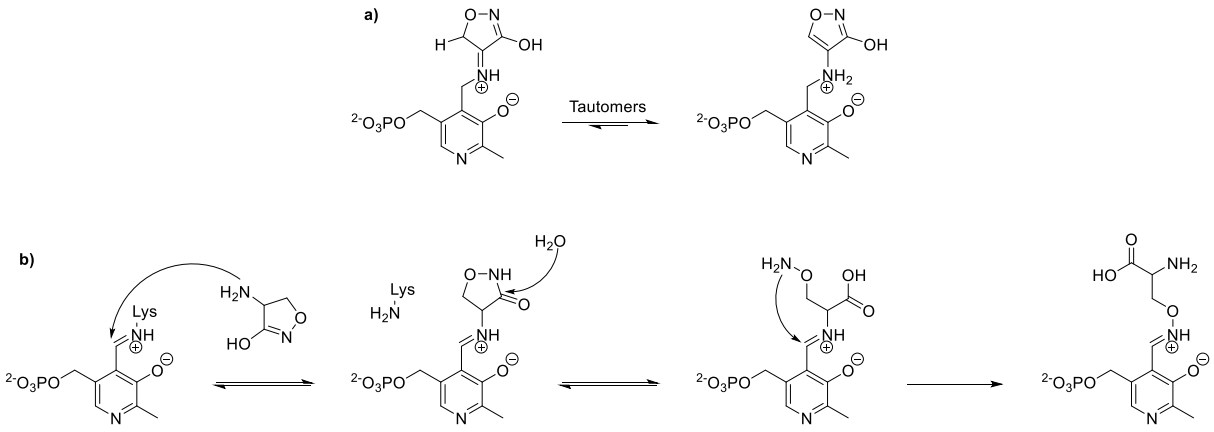


Figure 44: **a**: Reversibility of the inhibition due to possible imine enamine tautomerization; **b**: Probable pathway to the second, oxime product between PLP and cycloserine. [192]

The oxime product is an amino acid and as such might be a substrate for racemisation and oxidation in the assay. Another possible reason for the assay-hits with cycloserine is that a small portion of cycloserine becomes racemised due to tautomerization.

Nonetheless, the suggested explanations seem rather unlikely with respect to the low concentrations of still active racemase and the possibly formed alternative substrate or formed D-cycloserine. Consequently, there is no straightforward explanation for this finding and its correctness needs further confirmation.

4.1.7 β-Aspartame Racemisation, Control Experiment

One example of a very unusual substrate is the dipeptide β-aspartame as a very large amino acid derivative. Since our initial screening did not yield any hits above the limit of quantification, but some hits above the limit of detection, these results were put to the test. For this reason, L-β-aspartame was incubated overnight with the candidate

racemases. After microfiltration, the reaction solutions were checked with HPLC/UV-Vis/MS for conversion, using commercial β-aspartame (β-L-aspartyl-D-phenylalanine methyl ester) and the synthesized **3** (β-L-aspartyl-D-phenylalanine methyl ester) as standards. Since racemisation is only possible on the α-amino acid group of the aspartate moiety of the dipeptide the racemisation leads to diastereomer formation, improving a chromatographic separation. The first developed HPLC method only enabled detection of the Fmoc-derivatized diastereomer molecule masses in positive and negative single ion measurement mode, due to peak overlap with other reaction components. Therefore, a second method was developed, to enable diastereomer separation and detection without derivatization, increasing retention times of the diastereomer peaks, enabling UV/Vis and mass spectrometric detection without single ion measurements. With this second approach, the results from the first HPLC separation method were confirmed. Only either full, or no racemisation was found due to the reaction conditions (Table 8). Full racemisation corresponds to a racemase activity of > 11.5 mU/mg of racemase cell-free extract.

Table 8: Results of the incubation (18 h, 25 °C) of β-aspartame (methyl L-aspartyl-β-phenylalaninate, 45 mM) in potassium phosphate buffer (50 mM, pH 7.5, 0.1 mM PLP), employing Prozomix Racemases (3.6 mg/mL, cell-free extract). ProRCMS#: Number of the corresponding Prozomix Racemase. Active: If yes, racemization of the substrate methyl L-aspartyl-β-phenylalaninate was found, if no, no racemization was found.

ProRCMS#	2	5	11	28	29	33	35	36
Active	No	No	Yes	Yes	Yes	No	No	Yes

Even though some of the hits from the screening proved to be false hits, some of the reaction solutions contained fully racemised dipeptide (Table 8). Racemisation of this bulky substrate indicates the ability of the corresponding racemase to racemise bulky amino acids. Either the substrate β-aspartame (methyl L-aspartyl-β-phenylalaninate) fits into the active pocket of the enzymes as whole, or the active site pocket entrance is flexible enough to accommodate the aspartyl residue, while the rest of the bulky substrate does not enter the active site pocket and remains in the solvent phase. Similar behavior has been reported for the broad spectrum racemase from *Vibrio cholerae* that is able to bind non-canonical peptidoglycan.[55] This would enable racemisation of large amino acids. It seems like a methylene, or methylcarbonylene linker in between the amino acid moiety and the bulky rest of the molecule as in β-aspartame (methyl L-aspartyl-β-phenylalaninate) is enough to render such amino acids substrates.

4.1.8 Side Reactivity of Panel Racemases: Serine Dehydration

It is known from the literature that some PLP-dependent racemases possess serine dehydratase activity, and some serine dehydratases are structurally close relatives to fold-type III PLP-dependent racemases. [31,213]

Other racemases can still eliminate serine, when the hydroxy group is activated. For example, mouse serine racemase converts L- to D-serine in mouse brain for NMDA-receptor activation and is not reported to dehydrate serine. However, it has been found to eliminate serine-*O*-sulfate with a rate exceeding serine racemisation 500-fold. This has been used to detect serine racemase activity in mouse tissues, *via* coupled lactate dehydrogenase activity on the resulting pyruvate and monitoring of NADH consumption at 340 nm. [224]

One of our priorities was to find a racemase for efficient dynamic kinetic resolution of serine at preparative scale, in which case dehydration represents an undesired side-reactivity lowering the final yield. For this reason, serine dehydratase activity was sought to be ruled out before application tests. To find dehydratase activity in our candidates for serine racemisation, an additional experiment was conducted on a selection of most suitable candidates for preparative serine racemisation (ProRCMS#9 – #12, #15, #46, #53).

4.1.8.1 Dehydratase Assay Principle

The experiment is based on specific pyruvate detection, which is the product of serine dehydration (Figure 45), formed from the initial enamine *via* imine hydrolysis.

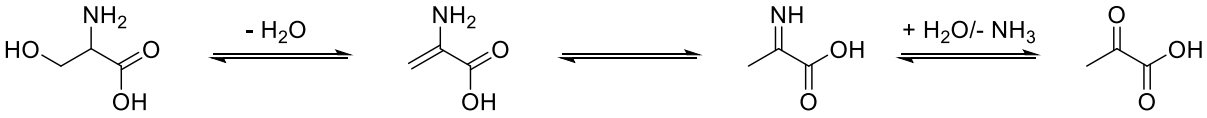


Figure 45: Reaction sequence after primary dehydration of serine, leading to the final product pyruvate.

Determination of pyruvate concentrations can be achieved through specific aldol condensation of pyruvate with salicylic aldehyde under alkaline conditions (Figure 46), leading to an orange dye.

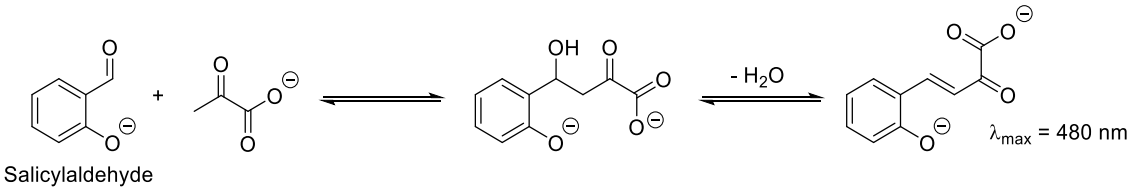


Figure 46: Formation of a conjugated dye, by aldol addition and condensation.

Other oxo acids do not react under these conditions, which was confirmed through the addition of α-keto glutarate to an additional calibration on pyruvate (data not shown). Serine forms an imine with salicylaldehyde under alkaline conditions, leading to a weak orange background. [225]

The experiments were conducted overnight at 400 µL-scale (18 h) in duplicates. After the incubation with the panel enzymes, the reaction mixtures were made alkaline and salicylaldehyde was added. A calibration with pyruvate was also prepared and found to be linear (Figure 47).

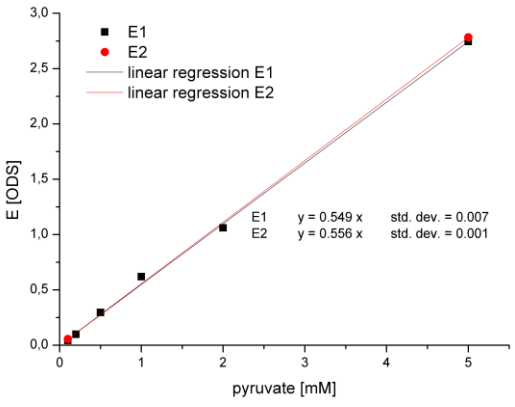


Figure 47: Calibration on pyruvate after reaction wit alkaline salicylaldehyde. Detection at 480 nm. E: Extinction; E1, E2: duplicates.

The long reaction time was chosen to also find low amounts of dehydration and to accommodate for the relatively low temperature (25 °C). The dehydration was determined relative to the corresponding racemisation. For this purpose, relative amounts of cell-free extracts were adjusted according to the racemisation activity found in the substrate screening, ranging from 1.0 to 1.5 mg/mL.

4.1.8.2 Amino Acid Dehydratases within the ProRCMS Panel

Dehydration activity was found in ProRCMS#9, #46 and #53 but not in #10, #11, #12, and #15 (Figure 48).

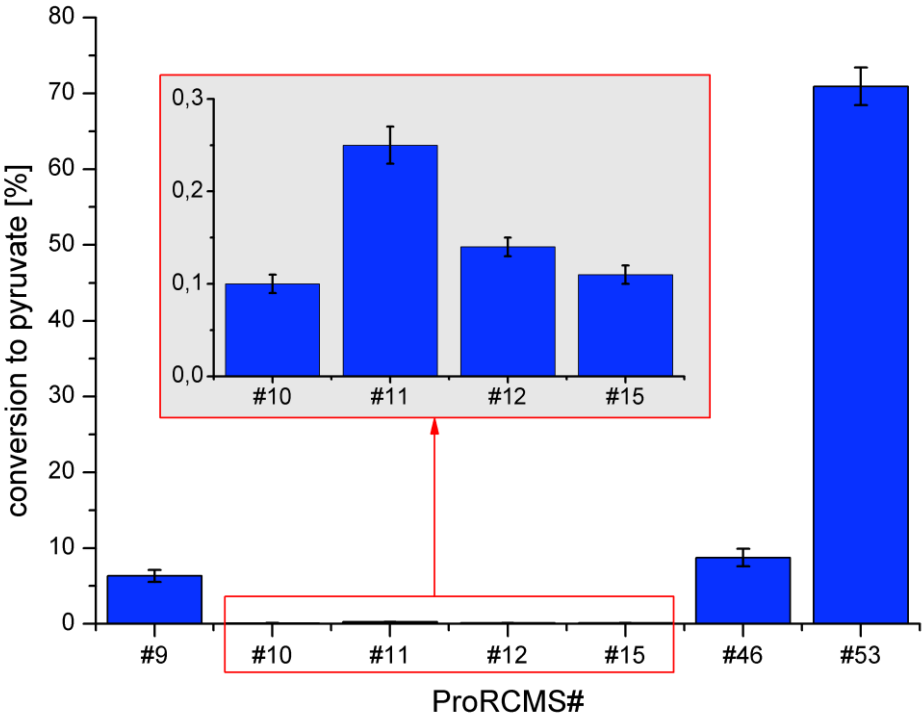


Figure 48: Extent of serine dehydration with the corresponding racemase under the given conditions.

While ProRCMS#9 and #46 only exhibited serine dehydration level of 10 to 20% under the reaction conditions, ProRCMS#53 achieved with 70% the highest conversion to pyruvate by far. This value is surprising, considering the fact that *E. coli* cell-free extract is known to decompose pyruvate, which has been confirmed in a separate experiment (data not shown). So, even though the corresponding activity value of 6.5 µU for the ProRCMS#53 cell-free extract seems small, dehydration activity can lead to significantly less product in the targeted cascade reactions, which are run over up to two days with high loading of cell-free extract.

ProRCMS#10, #11, #12, and #15 were determined to contain less than 50 µM of pyruvate. Such values can originate from the cell-free extract and lie below the limit of quantification (46 µM).

Consequently, ProRCMS#53 was excluded from later preparative experiments requiring serine racemisation. However, ProRCMS#10, #11, #12, and #15 show negligible amounts of dehydration only and as such were especially suited for serine racemisation over longer periods of time.

4.1.9 Reversible and Irreversible Unfolding of ProRCMS#9 and #31

To confirm the diverse melting points of ProRCMS#9 and #31 and to assess at what temperature the two enzymes begin to irreversibly unfold, a special protocol was applied to the differential scanning fluorometry measurements, as suggested by Svilenov *et al.*. [226] The protocol includes alternating heating and cooling cycles (25 °C) , as described in the experimental part of this work.

4.1.9.1 Method Principle

With the peak temperature of the heating periods increasing by 1 °C each cycle, fluorescence measurements at the peak temperature correspond to the reversible unfolding, as measured in a standard ramping differential scanning fluorometry experiment.

The fluorescence measurement at the minima temperatures correspond to the irreversible unfolding, because a fluorescence change only persists at the temperature minimum, if the enzyme is not refolded during the cooling.

Figure 49 displays the obtained fluorescence/temperature curves for ProRCMS#9, #31, and the DAAO from *Rhodotorula gracilis*.

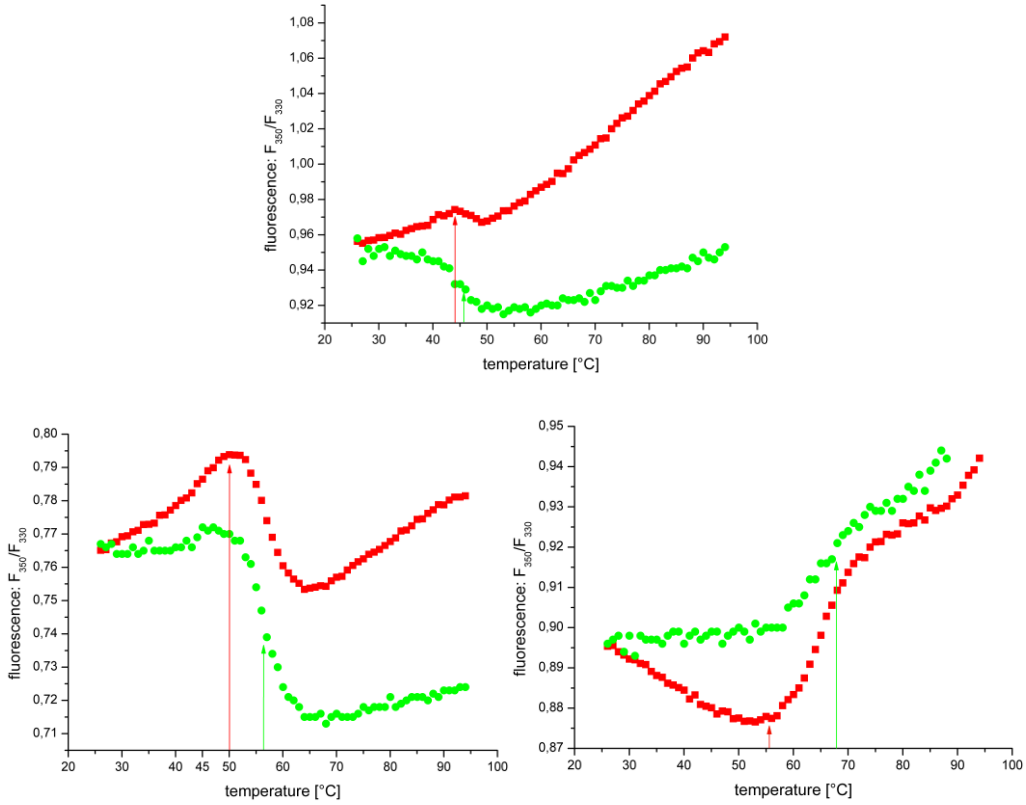


Figure 49: Fluorescence/temperature curves for DAAO from *Rhodotorula gracilis* (top) ProRCMS#9 (left), and #31 (right). The fluorescence is displayed as the ratio between the fluorescence emission 350 nm to that at 330 nm. Green circles: emission at 25 °C, after refolding, red: emission at temperature peaks. The green arrow indicates the obtained infliction point of the reversible unfolding, hence T_m. The red arrow indicates the onset of the irreversible unfolding, $T_{on/irr}$.

4.1.9.2 Results of the Reversible and Irreversible Unfolding Experiment

Table 9 displays the derived reversible and irreversible onsets of unfolding.

Table 9: Apparent melting point (T_m), determined *via* the infliction point of the reversible unfolding curve and onset of the irreversible unfolding ($T_{on/irr}$) of DAAO from *Rhodotorula gracilis*, ProRCMS#9, and ProRCMS#31.

	T_m [°C]	$T_{on/irr}$ [°C]
DAAO, *Rhodotorula gracilis*	46	44
ProRCMS#9	56	50
ProRCMS#31	68	56

Firstly, it should be noted that the experiment, employing purified enzymes, yields the same result for reversible unfolding temperature as the previous measurements with the standard DSF protocol, employing cell-free extract. This confirms that measurement of cell-free extract is valid and shows the improved sensitivity when employing purified enzyme.

The DAAO from *Rhodotorula gracilis* was measured to possess an apparent melting point of T_m = 46 °C and an onset of irreversible unfolding of $T_{on/irr}$ = 44 °C on. Compared to the value of T_m(Lit.) = 35 °C, provided by Pilone *et al.* [200,227] and the experienced allowed process temperature of below 25 °C, these temperatures are astonishingly high. Pilone *et al.* obtained this result from heating the dissolved enzyme for 30 min, before determining the residual activity. Clearly, the reason must be found in the method, since the results were confirmed in multiple repetitions and when conducting stability experiments with conditions similar to Pilone *et al.*'s, (see 4.2.3.2) very similar results were obtained. The deeper cause for the obtained high melting point is unknown. It may be found in the short period in which the enzyme is held at the peak temperature of the heating cycle, or in the high surface to volume ratio of the Prometheus capillary tubes.

The melting points of ProRCMS#9, and ProRCMS#31 were confirmed to lie at T_m = 56 °C, and 68 °C, respectively (compare 55.6 °C, and 69.4 °C, respectively, calculated with the Prometheus firmware). While the ProRCMS#9 melting point lies in a range, typical for microbial alanine racemases, the ProRCMS#31 was confirmed to be relatively thermostable.

The apparent onsets of irreversible unfolding of ProRCMS#9, and ProRCMS#31 lie at 50 °C, and 56 °C, respectively. The onsets of the reversible and irreversible unfolding strongly overlap in both enzymes. Hence, reversible and irreversible denaturing cannot be clearly distinguished with these experiments.

4.1.10 Discussion of Racemase Sequences: ProRCMS#9 and #31

Prozomix Ltd. provided two primary sequences from the ProRCMS amino acid racemase panel for further discussion. The sequences selected for further investigation were obtained for ProRCMS#9, a racemase employed in many serine cascade experiments due to its high activity on serine, and ProRCMS#31 due to its

broad substrate spectrum and the exceptionally high thermostability (T_m = 69 °C) were selected for further investigation.

To compare the sequences and find similarities and unique structural features, 3D-models were created from both racemase sequences using the Swiss-Model online tool. As templates for modelling, alanine racemase from *Geobacillus stearothermophilus*, (pdbid: 1BD0, external aldimine with D-alanine and N_ε-carboxylated Lys129) and broad- specificity racemase from *Pseudomonas putida* (pdbid: 4DYJ, internal aldimine) from the protein data base (pdb) rcsb.org were selected, based on their maximum sequence similarity.

Swiss-Model is an automated protein homology-modelling server that models a three-dimensional protein structure, based on a template structure, *via* an iterative least squares algorithm. Template structures can be either uploaded or chosen from the large database. Also, the Swiss-Model server provides functions for comparison of database sequences with the uploaded sequence, based on sequence homology. [228]

4.1.10.1 ProRCMS#9 Sequence

ProRCMS#9 is active on alanine and serine and, with lower activity, on glutamine. For non-canonical amino acids, racemization of N_ε-acetyl lysine, allylglycine and, with higher activity, 2-aminobutyric acid was detected.

Therefore, the overall substrate spectrum is limited and close to an alanine racemase (only alanine and serine). Additionally, found substrates (alanine analogues with a main chain extension of one carbon, such as allylglycine, 2-aminobutyrate) are usually not tested as a standard with other alanine racemases, thus may be substrates for these in general. The activity on glutamine and N^ε-acetyl lysine seems a little off, since these two moieties are much larger and the activity on N^ε-acetyl lysine might indeed be a misreading, because it was below the limit of quantification.

These results are quite well represented in the structure model which has a high sequence similarity to alanine racemase from *Geobacillus stearothermophilus* (sequence identity 42.0% with a global model quality estimate of 0.76). In comparison, the sequence similarity to BsrV is lower compared to this (sequence identity 26.7% with a global model quality estimate of 0.65). Neither an N-terminal extension, nor any of the 16 residues (*e.g.* chloride binding N174 in 4BEU) that are postulated to constitute

a molecular fingerprint [37] of broad-spectrum racemases, could be found in the ProRCMS#9 sequence.

What catches the eye at multiple occasions is that there are less prolines (16 vs. 23) and more glycines (31 vs. 26) in the structure of ProRCMS#9 compared to the template alanine racemase. This fits well to the expectation of an enzyme with properties similar to an alanine racemase but a somewhat extended substrate scope. Looking at available cysteine residues, alanine racemase from *Geobacillus stearothermophilus* carries six, but in the crystal structure of 1BD0, none of them seem to be available to form disulphide bridges. In ProRCMS#9 there are even eight cysteines, two of them on the enzyme surface. The remaining 6 however, seem to be close enough to form disulfide bonds in the holoenzyme, even though this cannot be deduced with confidence from the model structure.

Looking more closely at the aligned structures reveals 7 distinct sites with larger differences that can be grouped into 4 categories:

- **A:** Loops, connecting the barrel's α-helices and β-folds on the opposite site to the cofactor binding pocket.
- **B:** Residues around the active site entry.
- **C:** Residues on the enzyme surface
- **D:** The C-terminus

These areas of the primary sequence are marked red in the following schematic representation of the arrangement of the enzyme's secondary structures (Figure 50).

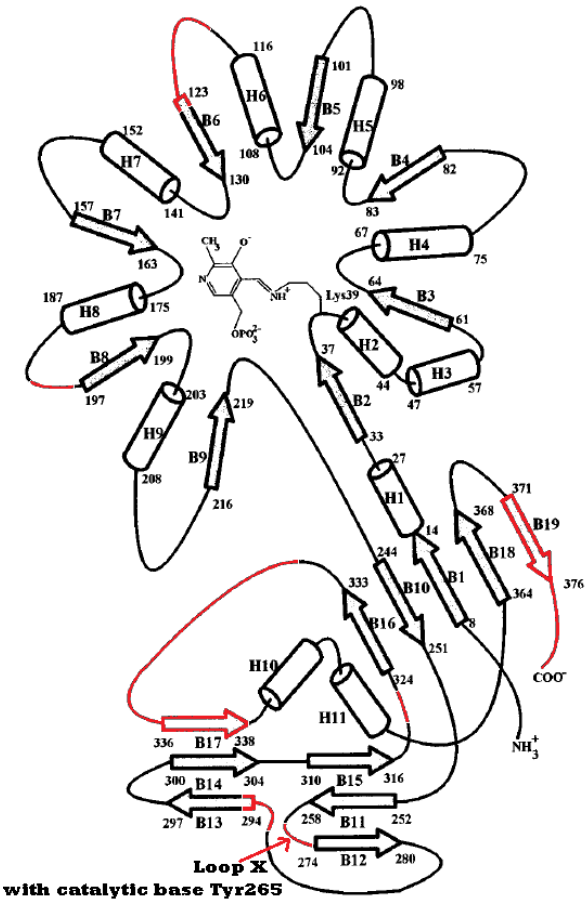


Figure 50: Schematic representation of the arrangement of secondary structures in alanine racemase from *Geobacillus stearothermophilus*. Structural elements that do not align well with ProRCMS#9 are colored red and discussed below. With slight modifications from (Shaw 1997) [33].

A: As already mentioned in Chapter 2.2.1, fold-type III racemases contain an eight stranded α/β-barrel with two openings. On top of the first opening resides the PLP cofactor with many polar contacts to the loops connecting the α-helices with the β-folds. Two of the eight loops at the opposite second opening contain larger alterations in ProRCMS#9 as compared to the template alanine racemase. The loop constituted by D118 – L126 is two residues longer than in the corresponding alanine racemase (Y118 – I124). Also, while the template alanine racemase contains two prolines in this loop (P121, P123), ProRCMS#9 has none. Interestingly, the same is the case for the second altered loop (R195 – I201). This loop has 3 more residues and only one proline close to it (P202, equivalent to P197 in ProRCMS#9), while there are three (P193, P196, P197) in the template sequence.

B: B: The entry site's external surface and the entry channel are composed of residues from both enzyme monomers, specifically two loops from each subunit. The monomer that also binds the same active site's coenzyme through interactions with its α/β-barrel, has one of these entry site loops (172 – 176, 235 – 241 in 4BEU) altered. The varied

loop has been reported to be shorter in broad spectrum racemases than in alanine racemases. Part of the altered loop (A236 – V242) is one amino acid shorter in ProRCMS#9 than in the template alanine racemase and contains only two prolines (P234, P247), instead of four (P229, P231, P235, P238).

Additionally, one of the entry site loops from the other subunit between β-sheet 2 and α-helix 2 (Figure 50), which also carries the catalytic base Tyr269 (Tyr265 in 1BD0), is modified. While there are no inserts or deletions in this loop, the surface stretching from the entry site opening to solvent exposed areas (T271 – Y278) seems completely altered, with one less basic, one additional acidic and 2 more aromatic residues.

C: Also close to the entry site, between the Tyr269 carrying loop (*vide infra)* and the monomer interface, there is another altered loop. While in the template alanine racemase the 4 amino acid long part carries two histidines, in ProRCMS#9 there is an additional aspartate, one lysine and no histidine. A little further away from the entry site, also close to the base-carrying (Tyr269) loop (*vide infra)*, there is yet another loop (Q325 – I334), which is 6 residues longer than in the alanine racemase template, extending into the bulk. This loop and the following β-sheet 17 carries 6 aspartate residues on the enzyme surface (EEAEEIQVGEE), where the template sequence does not even have a single acidic amino acid. Such repetitions of charged residues may belong to recognition elements. Additionally, the region has no proline, while the template structure has two (P320, P324).

D: Lastly, the C-terminus of ProRCMS#9 is 5 residues shorter.

4.1.10.2 ProRCMS#31 Sequence

ProRCMS#31 has a very broad substrate spectrum, contrary to ProRCMS#9. It is active on neutral amino acids, like leucine threonine, cysteine and methionine. Astonishingly, even though threonine was racemised, its activity on alanine and serine does not stand out, compared to the racemases that are available in the *E. coli* cell-free extract at 30 °C. As reported in the literature for the similar broad specificity racemase from *Pseudomonas putida*, ProRCMS#31 is able to racemise basic amino acids, [229] like lysine, histidine and was even tested positive on arginine-racemisation, however below the LOQ. Surprisingly, it exhibited high activity on aspartate, which is unique in the ProRCMS panel and to the best of my knowledge has not been reported for broad specificity racemases from *Pseudomonas putida* (Bar, template enzyme 4DYJ) [209,229] and *Vibrio cholerae* (BsrV) [37]. Even though there was no hard evidence

collected in the form of HPLC/MS or NMR spectra, to prove the viability of the assay results, the aspartate racemisation is reported here with high confidence, since also additional TLC analysis of coupling of ProRCMS#31 with DAAO from *Rhodotorula gracilis* displayed appearance of a new spot. Also, decarboxylation attributed to the β-keto acid oxidation product, has been observed upon acidifying this reaction mixture. However, glutamate racemisation has not been observed in the screening results. Additionally, aromatic amino acids (phenylglycine, phenylalanine, tyrosine) are racemised.

Looking at the results of the screening for activity on non-canonical amino acids, also reveals a broad substrate spectrum. The 4-carbon analogues of alanine, allylglycine and 2-aminobutyric acid are good substrates, and also the C5 homolog norvaline was racemised. Both lysine derived amino acids, N_ε-acetyl lysine and the bulky, charge-locked $N_\varepsilon,N_\varepsilon,N_\varepsilon$-trimethyllysine are racemised with good activities. The enzyme is more active on phenylglycine than on phenylalanine, and the click-chemistry building block azidohomoalanine was racemised with high activity. However, no activity was detected on β-aspartame.

It should be kept in mind that the enzyme has a high melting point and some activities might still be hidden, since the racemisation reaction was performed at 30 °C. Also, the whole assay is limited by the oxidase substrate spectrum [230] and rather unexpected results, like the activity on aspartate, require further confirmation due to the nature of the screening procedure and the limited time.

Attempts to fit the ProRCMS#31 sequence to the two discussed template structures 1BD0 (alanine racemase from *Geobacillus stearothermophilus*, sequence identity 25.14% with a global model quality estimate of 0.65) and 4BEU (BsrV, sequence identity 52.16% with a global model quality estimate of 0.82) always caused the Swiss-Model tool to shift to another template, due to the higher similarity. This preferred template is a crystal structure of a broad-specificity racemase from *Pseudomonas putida* (Bar, pdbid: 4DYJ, Wu et al. 2012). [209] The presence of a broad specificity racemase in *Pseudomonas putida* has already been known since the 1980s. [229] This enzyme has a sequence identity of 72.58% with ProRCMS#31, leading to a very high global model quality estimate of 0.91. Alignment of the generated ProRCMS#31-model to the crystal structure 4DYJ in PyMol leaves few parts (17) identified as not well

aligned, most of them single (13) or double (3) amino acids alterations and only one five amino acid long part.

Aligning the ProRC#31-model to BsrV also reveals high structural similarity. Most amino acids that define the active site and most molecular fingerprint residues that define broad spectrum racemases are conserved.

One unique feature, present in the BsrV structure, is the mode of stabilization of a substrate carboxylate binding arginine. As described in the theory part, this arginine is conserved among PLP-dependent racemases (R173 in 4BEU, R151 in ProRC31, R174 in 4DYJ, R136 in 1BD0). However, the secondary stabilization of this arginine differs between alanine racemases, which have a N^{ε}-carboxylated lysine (KCX129) for this purpose, and broad-spectrum racemases, which stabilize this arginine through contacts to an incorporated ion, suggested to be Cl^-. This ion in turn is stabilized by a neighbouring asparagine residue (N174 in 4BEU), which is conserved in broad spectrum racemases. And even though no chloride ion has been modelled into the template 4DYJ structure, the features needed for the described mode of chloride and arginine stabilization are also present in Bar (PDBID 4DYJ) and ProRCMS#31. In turn, a lysine residue in bonding distance to Arg151, like in alanine racemases (KCX129), is not available in the modelled ProRCMS#31 structure.

Most other features, present in the BsrV active site, are also present in ProRCMS#31, for example, the two catalytic bases (K74, Y299' in BsrV, K52, Y278' in ProRCMS#31) and the residues stabilizing the pyridoxal ring (H204, S245, R260 in BsrV, H182, S223, R238 in ProRCMS#31). A minor change is present in the binding of the PLP's phosphate group. In BsrV, Y208 donates a hydrogen bond to one of the phosphate oxygens, while in ProRCMS#31, this residue is replaced with E186, which lies well within bonding distance to the same phosphate oxygen in our model. Such variation may increase the acidity of the phosphate group, due to preferred localization of the aspartate proton at the phosphate, compared to a tyrosine proton. In the template broad-specificity racemase (Bar) from *Pseudomaonas putida*, this position is also occupied with an aspartate. This may be a functional variation between different broad-specificity racemases.

Most features fit nicely to the reported molecular fingerprint of BsrV. However, from the 16 postulated fingerprint residues of the group of broad-spectrum racemases, [37] six

are varied in ProRCMS#31, when compared to BsrV. In the following, the numbers are given as in 4BEU (BsrV crystal structure) and residues replacing these in ProRCMS#31 are given in brackets. The six altered residues are P206 (A184), Y208 (E186), K216 (G194), Y246 (F224), T349 (L329) and P391 (A370). Nonetheless, these residues still fit, since most fingerprint-positions are not limited to single amino acids. All alterations in ProRCMS#31 have already been found in other members of the broad-spectrum racemase group, with the exception of leucine instead of T349. This renders ProRCMS#31 unique, but in my opinion not to be excluded from this group of racemases.

Comparing the sequences of 4DYJ (Bar from *Pseudomonas putida*) and ProRCMS#31 also seems interesting, since their sequences are so similar and in the few differences, the origin for aspartate racemisation and the high thermostability of ProRCMS#31 may be found.

However, comparing amino acid counts does not really yield any pattern. They have the same amount of basic amino acids and even though ProRCMS#31 has six more acidic amino acids (48 vs 42), the altered positions are not found to form any relevant new bonds in the model. All three compared promiscuous racemases (BsrV, Bar, ProRCMS#31) have two cysteines at the same position of their sequence, forming a disulphide bridge between β-sheet 2 and 3 (Figure 50), which are the first two β-sheets of the barrel motif. The loop in between these two barrels carries the catalytic lysine. This might be another conserved feature in promiscuous racemases, considering all examined enzymes of this type bear this disulphide bond, whereas all examined alanine racemases do not. However, to confirm this, a lot more sequences should be compared. Interestingly, in the sequence of the alanine racemase from *Geobacillus stearothermophilus* there are six cysteines, but none of them seem to be in a distance to form disulphide bridges in the crystal structure 1BD0, even eight in ProRCMS#9 (*vide supra*).

Examining the individual changes in charged amino acids in the peptide chain led to the discovery of a potential additional inter-mer hydrogen bond between the N-terminus of α-helix 6 and the loop between β-fold 15 and 16 of the other monomer, close to the C-terminus of β-fold 15 (Figure 50). In our model, the γ-amide nitrogen of Gln122 is in hydrogen bonding distance with the side chains of Gln340’ and Glu341’. In the template racemase, the peptide chain carries an aspartate instead of Gln122, a

glycine instead of Gln340' and a lysine instead of Glu341'. Even though polar interactions might be possible between lysine and aspartate side chains, the available crystal structure does not show such interactions and the lysine might even be bent back through a neighbouring Asp361. (Figure 51)

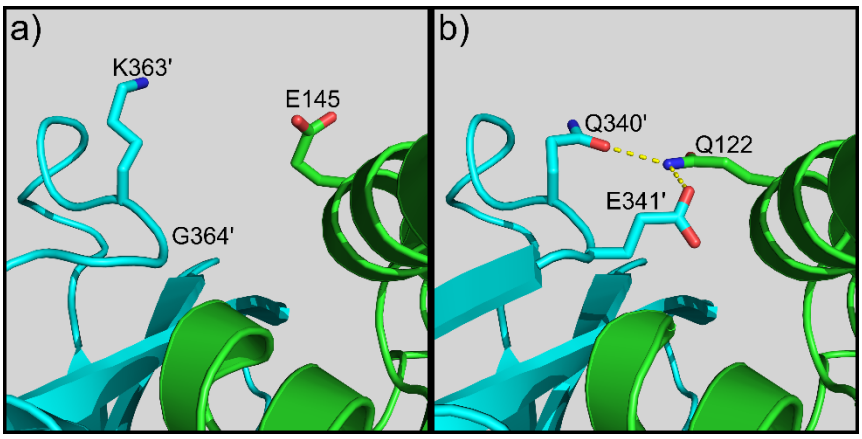


Figure 51: Potential additional inter-mer H-bridges. a) Crystal structure of broad-specificity racemase from Pseudomonas putida (pdbid: 4DYJ) [209], b) model of ProRCMS#31 with additional H-bonds.

Considering that dimer dissociation is believed to be the first step in irreversible enzyme deactivation, [231] the additional inter-mer hydrogen bonds may be a feature contributing to the high thermal stability of ProRCMS#31. To test this hypothesis, thermostability of wild-type broad-specificity racemase from *Pseudomonas putida* should be compared with an E145Q/K363Q/G364E triple mutant.

To evaluate another contributor of enzyme stabilization, the surface electrostatics of the two similar enzymes (broad-specificity racemase from *Pseudomonas putida*, pdbid: 4FS9, crystal structure with modelled external aldimine and ProRCMS#31-model) were calculated, using the APBS electrostatics plugin of PyMol, which is based on the Poisson-Boltzmann method. [232]

Computing the partial charge sum yields -6 for the crystal structure of the broad-specificity racemase from *Pseudomonas putida* and -12 for our model of ProRCMS#31. This may indicate a higher hydrophilicity of ProRCMS#31, leading to a higher solvation energy, or demonstrates the limit of our model. However, due to the high sequence similarity of both primary structures and because little new stabilizing interactions were found in the comparison of both structures, this might be a major factor for the higher melting temperature.

Additionally, close examination of a loop between β-fold 7 and α-helix 8 (Figure 50), which generates a “finger” pointing into the active site opening, led to the finding of an additional potential polar interaction between ProRCMS#31 and potential substrates. Since no structural changes were found inside the active site, this might be a reason for the observed acceptance of aspartate as a substrate in ProRCMS#31. At the tip of the “finger” sits a negatively charged glutamate (E186 in ProRCMS#31), possibly able to repel negatively charged amino acids. In the ProRCMS#31-model, this finger is shorter, due to an additional hydrogen bond between Tyr183 and Thr220. In the compared broad-specificity racemase from *Pseudomonas putida* equal positions are occupied by the two non-polar amino acids Phe206 and Ala243. (Figure 52)

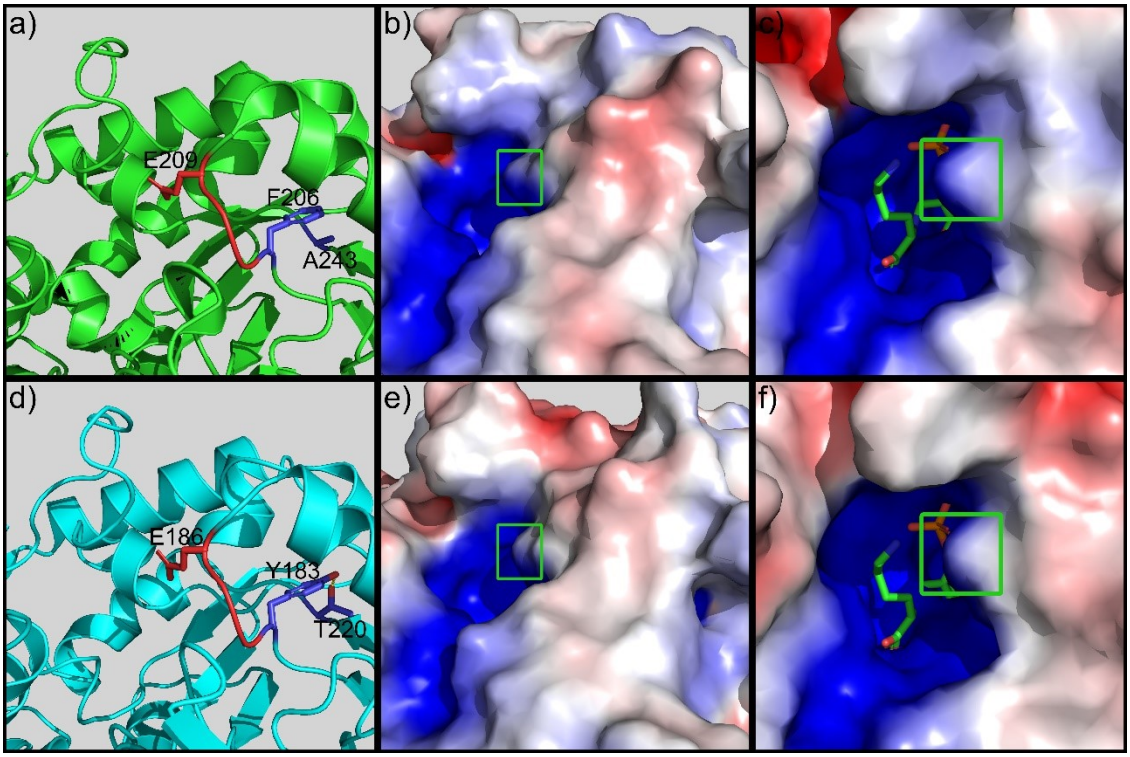


Figure 52: Active site entrance finger. a, b, d, e and c, f; pictures taken from the same angle. top row (a, b, c): broad specificity racemase from *Pseudomonas putida*, bottom row (d, e, f): model of ProRCMS#31. b, c, e, f: Electrostatic potential surface with a cut-off of -5 (red) and 5 (blue). a, d: main-chain of “finger” with aspartate side chain depicted in red, new hydrogen bonding interaction in ProRMS#31 with participating residues depicted in blue. b, e: side view of the electrostatic potential surface with the “finger” framed green. c, f: top view of the electrostatic potential surface with the “finger” framed green and the external aldimine.

To test this hypothesis, a Y183F/T220A mutant of ProRCMS#31 could be checked for activity on aspartate and if mutation of Glu209 to alanine opens the substrate spectrum of Bar to acidic amino acids.

However, ProRCMS#31 proves to be yet another example for how fine-tuned enzymes are during an evolutionary process, and that properties like thermostability cannot

simply be deduced from patterns in the molecular structure, at least not without years of research or the abilities of sophisticated AI's.

Why ProRCMS#31 displayed aspartate racemisation in our experiments cannot be deduced from the active site model. All residues that are in a range of 3.5 Å to the PLP-lysine external aldimine in the 4FS9 (lysine reactant, modelled into the active site of Bar) structure are preserved in our model of ProRCMS#31. So, the reason for this result must be found somewhere else, maybe in features at the entry site, within the entry channel, or to other experimental errors, or due to the static model. Anyways, aspartate racemisation should be repeated and confirmed, for example *via* chiral HPLC.

4.1.10.3 Conclusion of the Sequence Alignment

Even though very limited by the static model of a highly dynamic reality, the sequence comparison via the generated models yielded interesting results. Both selected enzymes are at the same time very representative and also unique members of their specific groups.

ProRCMS#9 is a classical alanine racemase, but also bears high reactivity and a higher flexibility than the compared racemase from *Geobacillus stearothermophilus*. Also, comparing it with structures from the pdb, the best matching sequences still have relatively low similarities to ProRCMS#9 (1BD0, sequence identity 41.98% with a global model quality estimate of 0.76).

ProRCMS#31 clearly belongs to the group of broad-spectrum racemases, even though one of the fingerprint-residues is missing and it lacks the N-terminal extension. It has extraordinary catalytic features, like aspartate racemisation and a high thermostability, not present in the very closely related Bar (4DYJ, sequence identity 72.58% with a global model quality estimate of 0.91). This allows to propose some parts in its sequence that might be, to a certain extent, responsible for some of its features and inspires ideas for mutational studies.

4.2 Coupling Racemase and Oxidase: Process Optimization

Operating transketolase cascades (*vide infra*) at elevated temperatures is desired, mainly because at 25 °C the employed transketolase from the hyperthermophile *Geobacillus stearothermophilus* has less than 20% of its maximum activity. At 70 °C, where it reaches its maximum activity, it is stable over three days. [117] Additionally, other catalysts are compatible to elevated temperatures as well. One of the amino acid racemases from the screened panel has a melting point of 69 °C, as determined in our experiments. Catalases are known to be relatively stable, low-cost commercial preparations, and can be continuously added during reactions. Also, heat-resistant, alternative, inorganic reagents for H_2O_2-quenching, for example MnO_2 are available in every lab.

Overall, the DAAO from *Rhodotorula gracilis* represents a serious bottleneck for setting the process temperature. Compared to the high thermostability of the cascade enzymes, available DAAOs for serine deamination rapidly loose activity above 25 °C, [233] which limits the maximum process temperature to 25 °C, more than 20 °C below the desired temperature of > 50 C. It should be kept in mind that the fungal oxidases from *Trigonopsis variabilis* and *Rhodotorula gracilis* already are relatively stable, compared to for example the commercially available oxidase from bovine liver. [234]

Because the relatively low thermostability of the oxidase represents such a prominent bottleneck, possibilities to enhance the operational stability of the chosen enzyme from *Rhodotorula gracilis* were explored in greater detail.

In theory, protein denaturation follows a variety of chemical and physical routes. Physical unfolding leads to desolvation and aggregation. The main contributor to conformational stability of proteins is long range hydrophobic interaction within the protein core, counteracting the destabilizing loss of non-local conformational entropy. The fine-tuned packing in native proteins leads to an exclusion of solvent molecules from the protein interior. Flexible charged residues (Glu, Lys, Arg), located on the protein surface, help maintain the protein solvation shell. A measure for the conformational stability is the free energy change ($\Delta G_{f->u}$) upon denaturation (transition from folded to unfolded state) of the protein. Most unfolded proteins are insoluble and form aggregates.

The melting temperature of a protein is defined as the temperature at which 50% of the protein molecules are unfolded ($\Delta G_{f->u}$). It usually lies between 40 and 80 °C and is dependent on protein hydration. The melting temperature increases when protein hydration decreases, due to destabilization of the unfolded state. Depending on the analytical method, and in accordance with the diverse unfolding routes, proteins can have more than one melting point. For example, differential scanning fluorimetry measures a change in the fluorescence upon a change of temperature. This is attributed to a change in the chemical environment of tyrosine's and tryptophan's aromatic moieties upon conformational movement of the protein backbone. Since protein denaturation often occurs in several distinguishable steps, of which some are reversible, melting temperatures can be measured for each of these steps.

The thermal stability of proteins can be enhanced through molecular features of the protein itself (*e.g.* additional hydrogen bonds, disulfide bridges, *vide infra*) and favourable interactions with so-called thermoprotectants. These can be either cellular components, or accumulated low-molecular-weight compounds, decreasing the contact area of hydrophobic fragments of the protein with water, thus decreasing the energy of the folded state. Also, dissolved ions influence protein stability by non-specific (Debye-Hückel) electrostatic shielding and by specific binding. [235]

Some proteins adsorb to hydrophobic air/water interfaces, which can induce denaturation. [235] Protein denaturation during aeration is usually attributed to foaming and shear-stress, which in turn is caused by the movement of the bubbles through the liquid. [236,237] However, the reason for deactivation is sometimes attributed to a changing microenvironment, when certain proteins adsorb to and align at the hydrophobic surface of a gas phase. [238] There are several reports that bubbling reaction mixtures with air, or oxygen deactivates contained oxidases. [239,240] One well-documented example of oxidase deactivation was described by Findrik *et al.*, reporting the operational stability decay rate constant for the snake venom L-amino acid oxidase from *Crotalus atrox* to be $1.53 \cdot 10^{-4}$ min^{-1} without aeration and $3.63 \cdot 10^{-3}$ min^{-1} with aeration of 10 L/h in their 50 mL fed-batch reactor at 30 °C. [241] During early experiments conducted for this work, such phenomena were noted as well. Low conversions were attributed to a loss in oxidase activity and foaming was observed when the reaction mixture was bubbled with oxygen and when HCl (0.1 M) was fed into the reactor for pH-shift compensation during transketolase reactions, liberating

HCO_3^- from hydroxypyruvate. Such acid additions potentially cause a local pH-minimum, contributing to enzyme destabilization.

However, bubbling represents the most frequently applied method and effective measure to quickly saturate reaction mixtures with oxygen and replace it, when used up during reaction. An efficient measure to maintain the reaction mixture saturated with oxygen is important, considering the high K_M of oxidases for molecular oxygen, which is in the case of *Rhodotorula gracilis*, 2.6 mM. [39] This value is quite high, considering the low aqueous solubility of molecular oxygen (*vide infra*).

The fact that the oxygen partial pressure is a critical parameter in the reaction velocity of our amino acid racemase/DAAO/transketolase reaction cascades was well observable during one of our experiments. In this experiment, molecular oxygen was provided to the reaction mixture inside a reactor that was not air tight, *via* a submerged silicon semipermeable tubular membrane (*vide infra*). The oxygen bottle, providing the oxygen to the tubing was shut in the evening and only opened again the next morning, due to an error. The reaction velocity of the reaction was monitored as the titration curve of the pH-stat hydrochloric acid addition to the reaction mixture. The amount of added hydrochloric acid is equivalent to the hydroxypyruvate conversion, because it compensates a pH-shift caused by $NaHCO_3$-liberation (*vide infra*). Figure 53 shows the titration curve, displayed on our auto-titrator, exhibiting a continuous velocity decrease after the oxygen provision was shut off and a sharp increase, after the oxygen bottle was reopened.

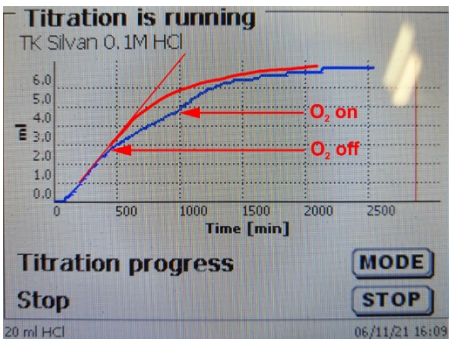


Figure 53: Picture of the display of the auto-titrator, displaying the continuous addition of 0.1 M HCl to an amino acid racemase/DAAO/transketolase one-pot cascade reaction for the conversion of L-serine and 2,4-dihydroxybutanal to 5-deoxyfructose. Later additions to the picture are displayed in red, including arrows marking the points, where the oxygen bottle was shut and reopened and an estimated extrapolation of the initial velocity and an uninhibited reaction curve.

From the slopes, shortly before and after the oxygen bottle was reopened, it can be estimated that the overall cascade reaction was about 1.4 times faster, when pure

oxygen was provided *via* the silicone tubing, compared to when only air diffused into the reactor.

In general, oxygen-solubility mainly depends on the oxygen partial pressure, salinity, and temperature. At 25 °C, under 1 atmosphere of air, the maximum oxygen solubility in pure water is 256 µM and with 1 atm of pure oxygen, it is 1.2 mM, according to Henry's law. Reactions catalysed by DAAO were carried out in buffers with concentrations ranging from 25 mM to 250 mM and additional salts in low concentrations (*e.g.* $MgCl_2$). Such salinities translate into a minor reduction of the oxygen concentration of less than 10%. [242,243]

Concluding, there was a strong interest in improving the performance of the overall cascade reaction through widening of the oxidase-bottleneck. Therefore, the thermostability of the enzyme was sought to be improved without lowering of the oxygen concentration.

Known factors, **decreasing** the operational stability of enzymes at fixed temperatures are:

- Reactive components in the solution (*e.g.* H_2O_2), leading to disruptive alteration of the enzyme's functional groups (*e.g.* oxidation of thiol, amine groups), or primary structure (hydrolysis)
- Shear forces from turbulent agitation of the reaction mixture and
- Adsorption to heterogenous surfaces adjacent to the reaction mixture, for example surfaces of the reaction vessel, or liquid gas interfaces can lead to unfolding and agglomeration of proteins

Known factors **increasing** the operational stability are

- Beneficial molecular features of the protein
- Codissolved thermoprotectants

Therefore, four strategies to achieve higher operational stabilities of the oxidase were considered: improvement of H_2O_2 dismutation-efficiency, stabilization of the DAAO through the addition of stabilizing additives, or multipoint covalent immobilization, and alternative designs to efficiently introduce molecular oxygen into the reaction mixture.

4.2.1 Process Optimization A: Choice of Catalases

Efficient H_2O_2 dismutation is important during DAAO catalysis. It does not only prevent undesired side reactions of substrates and products, increasing the overall yield, but also increases the stability of the enzymes, which can otherwise be deactivated by the highly reactive peroxide. When it comes to the choice of catalyst for H_2O_2-decomposition, mainly commercially available catalases from Sigma-Aldrich (Merck) were considered, due to their high turnover, known compatibility to other enzymes in the system, and due to good experience when combining it with the oxidase from *Rhodotorula gracilis*: the catalases from bovine liver, *Micrococcus lysodeicticus*, *Corynebacterium glutamonicum*, and *Aspergillus niger*.

In 2002, catalases from different origins (*Aspergillus niger*, *Micrococcus lysodeicticus* and bovine liver) were compared in their potential to improve the DAAO catalysed conversion of phenylalanine to phenylpyruvate. [244] The authors state that 100 mM racemic phenylalanine was fully (50 mM, D-phenylalanine) converted within 55 min, when immobilized DAAO from *Trigonopsis variabilis* (4.34 nkat/mL) and catalase from *Micrococcus lysodeicticus* (43.4 nkat/mL) were added. On the other hand, employing an equivalent amount of the catalases from *Aspergillus niger* or bovine liver delivered only about 80% of product, even at longer resident times and even though the residual activity of these catalases was higher. The authors reasoned this to be due to the lower K_M of the catalase from *Micrococcus*.

Another publication from 2002 compared kinetic parameters and thermostability of several catalases from different origins, [245] which are displayed in Table 10, together with the aforementioned results from the pyruvate conversion.

Table 10: Kinetic parameters [K_M: Michaelis constant for H_2O_2, v_{max}: maximum rate, given in (mmol H_2O_2)/(µmol heme · sec)] and thermostability ($t_{1/2}$: half-life) of different catalases, [245] compared with results of a DAAO catalysed phenylalanine conversion (conv.) with determined residual activity (res. act.) of the corresponding catalase. [244]

	K_M [mM]	v_{max}	$t_{1/2}$ (65 °C) [min]	conv. [244]	res. act. [244]
A. niger	465	227	14	83% (70 min)	88%
bovine liver	93	212	0.3	81% (60 min)	66%
M. lysodeicticus	152	313	14	100% (55 min)	55%

Comparing the kinetic constants of the used catalases contradicts the authors' statement that the reason for the higher conversion is the lower K_M. [244] Half-lives at 65 °C were determined to be 14 min for the two microbial catalases and 0.3 min for the enzyme from bovine liver, underlining the higher stability of the microbial catalases, compared to the mammalian one.

Even though the literature comparison did not yield a coherent conclusion, the reported higher stability and phenylalanine conversion led us to replace the in-house available catalase from bovine liver with a commercial preparation of the catalase from *Micrococcus lysodeicticus*.

4.2.2 Process Optimization B: Stabilizing Additives

The most straightforward and frequently applied method to stabilize an enzyme is the addition of thermoprotectants, small molecules that stabilize proteins against thermal inactivation. It should be kept in mind that adding more components to the reaction mixture does not only make the system more complex to evaluate, but also complicates workup and represents a trade-off with the oxygen concentration, which is decreasing with higher additive concentrations.

The stabilizing effect of different ions and small organic molecules at concentrations between 0.1 and 1 M is quantified in the Hoffmeister series, which divides these cosolutes into kosmotropic (stabilizing, salting out) and chaotropic (destabilizing, salting in) cosolutes. [246,247] However, the general trend that kosmotropic cosolutes stabilize enzymes is largely dependent on the protein's molecular structure and surface and can even be reversed. [248] Additionally, proteins can react very specifically with certain ions or molecules and these effects may outperform the general trends of the Hoffmeister series by far. For example, there is a specific binding site for bivalent cations in transketolases, and binding of cations has a large influence on their catalytic activity, conformation, and stability and most crystals of transketolases could only be obtained with bound Ca^{2+}. [116,129] Another example is lactate dehydrogenase from different sources, exhibiting a specific fructose-1,6-bisphosphate binding site that largely influences the equilibrium between the dimeric and tetrameric states of the enzyme, effecting its activity and stability (allosteric activation). [249,250]

Kosmotropic ions and small molecules, expected to be most stabilizing according to the extended Hoffmeister series, are fluoride, phosphate, sulfate, different ammonium

ions, trimethylamine-*N*-oxide, glycine betaine, and urea, with their effects being mostly additive. [247] Examples for other small molecules, reported to potentially stabilize enzymes, are sugars, sorbitol, glycine, glutamate, cyclic 2,3-diphosphoglycerate, 2-*O*-β-mannosylglycerate, and di-*myo*-innositol-1,1'(3,3')-phosphate. [235]

Due to the observed dependence of the cascade reaction velocity on the oxygen concentration, a threshold for the salt concentration was defined. Only salt concentrations that lower the oxygen concentration by less than 5% were considered negligible and tolerable. At 25 °C, this threshold is reached with a conductivity of 15 mSi/cm, corresponding to a KCl concentration between 0.1 and 0.2 M. [242,243,251] The threshold does not represent a hard limit and in most cascades, for example in those with ammonium formate as the NADH regenerating cosubstrate, this concentration range is exceeded. Nonetheless, these considerations provide a guideline for what concentration levels may outweigh the beneficial effects of additives by their negative effect on the oxygen concentration.

Concluding, additives are only worth considering, when stabilizing the oxidase for at least 1 - 2 °C, at concentrations below 100 mM. The probability of finding an additive, especially at such low concentrations, that has a large stabilizing effect on the oxidase is low and finding the one co-solute, specifically binding to the oxidase is like finding a needle in a haystack. However, since differential scanning fluorimetry is available in our group as an analytical tool to quickly estimate the effect of the additive on the enzyme's melting temperature, the influence of a few common, frequently used additives at low concentrations (100 mM) was determined anyways.

For this purpose, the DAAO from *Rhodotorula gracilis* (2 mg/mL, cell-free extract) was dissolved in bicine buffer (200 mM, pH) with or without known kosmotropic additives at 100 mM concentration, and the melting temperatures of the oxidase in these solutions was compared with the melting temperature of the same catalyst, dissolved in deionized water. In a second experiment, comparing the reaction rates of the oxidation of D-serine to hydroxypyruvate in solutions containing the DAAO and the respective additives, no influence of the additives on the oxidase activity was found at the given concentrations (data not shown).

Figure 54 displays the effect of the chosen additives ($(NH_4)_2SO_4$, sorbitol, glycine betaine, urea) on the melting temperature of DAAO from *Rhodotorula gracilis*.

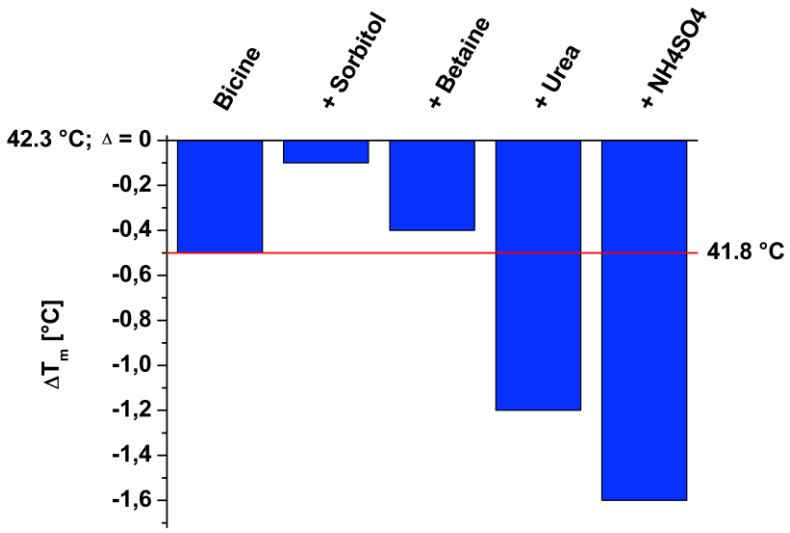


Figure 54: Effect on different additives (100 mM) on the melting temperature of DAAO. Melting temperature of the catalyst in deionized water is 42.3 °C. Additives are added including bicine buffer. Betaine is glycyl betaine.

A remarkable effect on the oxidase melting temperature was observed, even at 100 mM concentration. However, the stabilizing effects of sorbitol and glycyl betaine are too low (< 0.5 °C) to be considered a serious option to increase the enzyme's operational stability. The bicine buffer has a small destabilizing effect (ΔT_m = - 0.5 °C). Since the effects are known to be additive (*vide supra*), sorbitol and glycyl betaine are stabilizing the protein (ΔT_m = + 0.4 and + 0.1 °C, respectively). Surprisingly, addition of the most kosmotropic additives ($(NH_4)SO_4$) and urea led to the most prominent destabilization (- 1.1 and -0.7 °C, respectively). An influence of the additives in the given concentrations on the rate of D-serine oxidation was not observed.

As expected, the effect of the best stabilizer sorbitol is only small. Therefore, it was decided not to add it to preparative reactions with DAAO, to avoid purification problems

In the context of small molecule stabilizers, it should be noted that amino acid oxidases generate H_2O_2 during turnover, representing a major contributor to its inhibition. For this reason, reagents protecting the enzyme against reaction with H_2O_2, like antioxidants (*e.g.* ascorbic acid), can potentially increase the lifetime and operational stability of the enzyme. However, it was already decided to apply the most suited catalase to protect the oxidase against H_2O_2. Also, the effects of such additives on the protein can only be examined in solutions containing the enzyme and H_2O_2 and such systems are dependent on multiple variables.

4.2.3 Process Optimization C: Immobilization of D-Amino acid oxidase

Another effective measure to protect enzymes against unfolding and resulting deactivation is immobilization. Industry often uses immobilized enzymes not only because they are more stable, but also because such catalysts can be reused as demonstrated for example in the synthesis of Islatravir. [161]

Immobilization can stabilize enzymes for multiple reasons. In our case:

- Binding the oxidase to a solid carrier protects it from macroscopic physical stress, like excessive movement of non-liquid phases within the solution, leading to shear stress, association to the surface of gas bubbles and aggregation of multiple enzyme molecules.
- Coimmobilization of amino acid oxidase and catalase promotes H_2O_2-channeling from the oxidase to the catalase, thus reduces residence time of H_2O_2 and opportunity to react with other components of the reaction mixture.
- Multipoint attachment, meaning the formation of more than one covalent bond between the carrier and one enzyme dimer can prevent dimer dissociation and interdomain movement.

To understand, why multipoint attachment can stabilize enzymes, it is useful to look at a simplified model for stepwise thermal inactivation of dimeric enzymes. In a first step, oligomeric enzymes reversibly dissociate into monomers and only after that, unfolding of the tertiary structure, starting with interdomain movement, leads to irreversible deactivation. [252]

Figure 55: Simplified mechanism of thermal denaturation. A: thermal activation of the dimer, B: dimer dissociation, C: irreversible monomer unfolding. Green: active forms, red: inactive forms. [252]

In reality, enzyme deactivation, accompanied by macroscopically observable enzyme aggregation is far more complex. For example, the D-amino acid oxidase (DAAO) from the yeast *Trigonopsis variabilis* has been shown to deactivate at higher rates, when the FAD cofactor is dissociated. In turn, this FAD dissociation itself is dependent on the oxidation state of a loop cysteine (Cys108), converted to cysteine sulfinic acid in a reaction with the oxidase coproduct H_2O_2, or under thermal stress. Likewise, addition of free FAD to the solution can stabilize the oxidase. It has also been shown that the

oxidase is more stable at higher oxidase concentrations, which is in agreement with the simple two step mechanism. Interestingly, the externally added FAD has a lower effect on enzyme stability at higher enzyme concentrations. [253] However, the enzyme concentration that leads to increased stabilization in the aforementioned example, leads to decreased stability probably due to favoured dimer formation, in other cases, caused by its influence on aggregation kinetics. [71,253]

Despite the complexity of realistic models for enzyme deactivation through unfolding, it should be clear that factors that prevent domain movement, dissociation of coenzymes and dissociation of enzyme oligomers have the potential to prevent one of the reversible steps of the simplified model and drastically stabilize enzymes.

One possibility to prevent dimer dissociation besides, for example, the preparation of fusion proteins, is immobilization of both monomers to the same support.

4.2.3.1 Immobilization Protocol

For immobilization of the DAAO from *Rhodotorula gracilis*, a protocol of Guisán *et al.* was followed. In their paper from 1988, they state that oxidation of carrier glyceryl groups leads to a surface density of 17 aldehyde groups per 1000 $Å^2$ and that an enzyme of 90 kDa molecular weight exhibits a surface area of 20000 $Å^2$. Assuming the shape of the enzyme would be a simple cube, this translates into more than 50 available aldehyde groups per enzyme. This fact, together with the later installed and optimized flexible linker, enables the establishment of multiple covalent bonds between one enzyme molecule and the carrier (multipoint covalent immobilization). [254,255,256,257,258]

The oxidase from *Rhodotorula gracilis* exhibits patches with a high density of reactive lysine residues on its surface, as displayed in Figure 56.

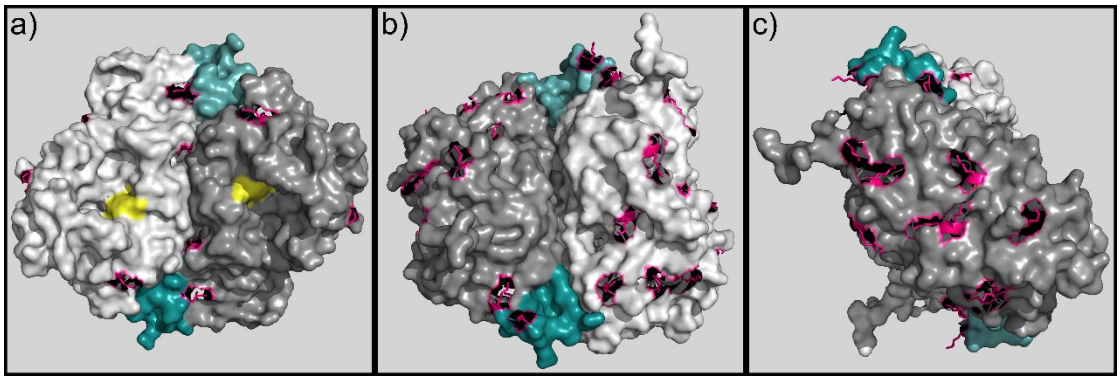


Figure 56: Surface of DAAO from *Rhodotorula gracilis*. Residues important for substrate binding (Y223, Y238, R285) [259,260,261] in yellow; C-terminal regions important for dimer association [262] in cyan and the surface lysines as red sticks, a: front, b: backside, c: side view.

It should be noted that none of the reactive lysine residues is close to the active site entry, where covalent bonds to the resin have a high probability to hinder necessary domain movement, or substrate entry. Instead, the oxidase has a patch of adjacent lysines at the lateral side of the enzyme.

As a carrier, Sepharose CL-6B was used. The solid beads (45 – 165 µm) consisted of a cross-linked (with 2,3-dibromopropanol) agarose matrix that has a high surface area of 25 m^2/mL, carrying a high density of active (70 µmol/mL) surface hydroxy groups. [254] In a series of subsequent chemical steps, a relatively long, flexible, polar linker was installed on the beads' surface (Figure 57).

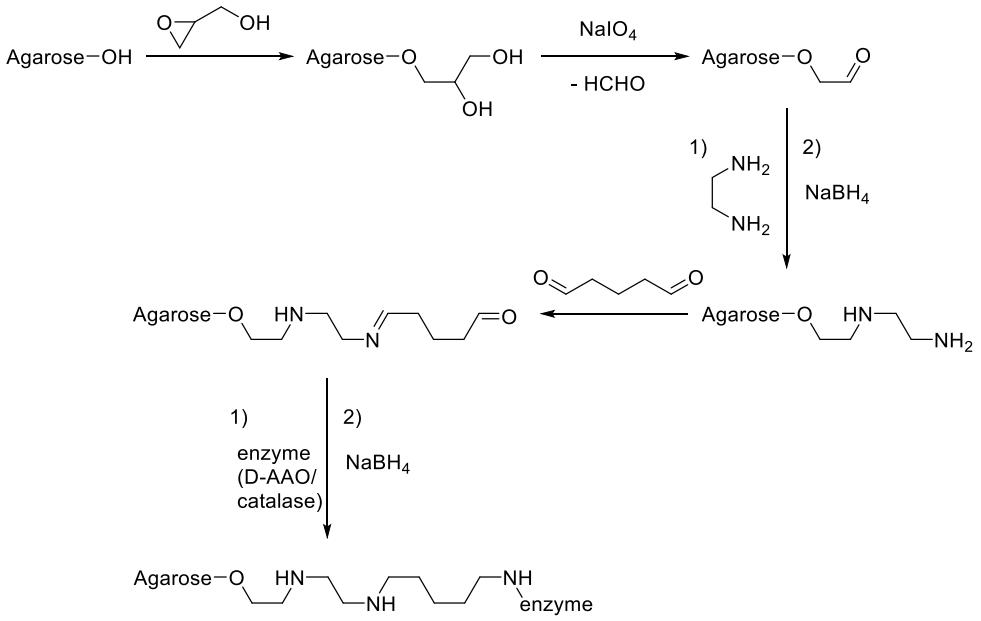


Figure 57: Reaction sequence for binding the DAAO from *Rhodotorula gracilis* (D-AAO) and catalase from bovine liver to the agarose carrier.

In the first step, glyoxal is added. The Sepharose's hydroxy groups react with the glyoxal's epoxide to establish an efficiently oxidizable group. After oxidation of the glyceryl vicinal diol to a carbonyl group, the carrier is reductively aminated with ethylene diamine. At this point, the properties of the modified material significantly change compared to the unmodified, such that it swells more when excess water is available and that the individual beads have a greater tendency to agglomerate. The free terminal amino groups are then allowed to form imines with glutaraldehyde. The high polarity, and the intramolecular repulsion of the charged secondary amines of the now installed linker facilitates solvation, and prevents collapse onto the surface of the carrier. At the same time, the low hydrophilicity of the glutaraldehyde prevents removal of too much of the bound glutaraldehyde during washing steps. After removing excess glyceraldehyde through washing with water, the beads were slightly brownish. Now, the enzymes (catalase from bovine liver, Sigma-Aldrich (Merck) and DAAO from *Rhodotorula gracilis*, purified *via* Ni-NTA chromatography) were added to the suspended beads. To allow the enzymes to slowly form multiple glutaraldehyde diimines by coupling the amino groups on the beads and the enzymes' lysines ε-amino groups, the mixture was gently stirred at r.t. for 2 h. The imine groups were then gently reduced with borohydride at pH 10, converting them to stable secondary amines and binding the established flexible linkers permanently to the enzymes.

In general, multipoint attachment of enzymes is enabled by the large internal surface of the carrier, a high superficial density of reactive groups on the carrier and the enzyme, minimal steric strain on the linker, allowing a high degree of conformational flexibility, and high compatibility of the enzymes' and carrier's reactive surface groups leading to a permanent and stable bond. To achieve a stable and useful catalyst, after the immobilization is finished, there should not be left over reactive surface groups, possibly interfering in the application of the catalyst.

Individual parameters of the procedure, already highly optimized by Pedroche *et al.* [257], can be found in Chapter 5.4.2.

4.2.3.2 Assaying the Thermostability of Immobilized Enzymes

To verify the successful immobilization, to estimate the attained stabilization, and the thermal compatibility to the heat resistant ProRCMS#31, the different enzyme preparations were incubated at different temperatures (24, 30, 40, 50, 60, 70 °C) in potassium phosphate buffer for 2 h. After the incubation, the different catalysts were

employed to generate hydroxypyruvate from serine. In the case of oxidase preparations, hydroxypyruvate is directly formed from D-serine, in the case of ProRCMS#31, it is produced with the help of an excess of DAAO from L-serine (Figure 58).

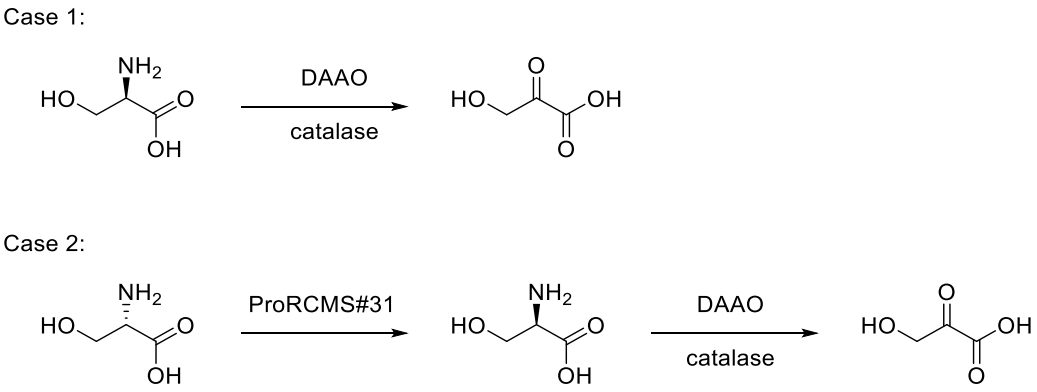


Figure 58: Reaction catalysed by the preincubated enzyme preparations. Case 1: preincubated DAAO preparations are incubated with D-serine. Case 2: Preincubated ProRCMS#31 is incubated with L-serine and an excess of DAAO.

The amount of formed hydroxypyruvate was measured with WST-1, which forms a blue dye specifically from the reaction with hydroxypyruvate Figure 59. [263]

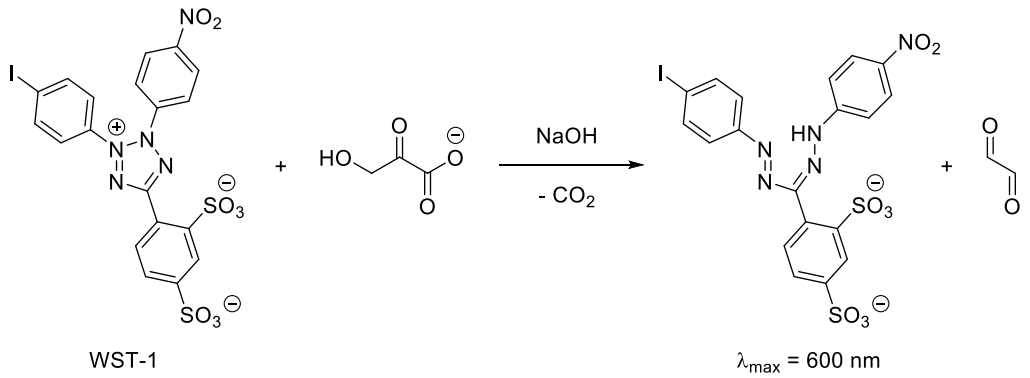


Figure 59: Specific redox reaction between WST-1 and hydroxypyruvate, leading to the formation of a blue dye.

4.2.3.3 Results of the Stability Measurements of Immobilized Preparations

Figure 60 shows the conversion of D-, or L-serine to hydroxypyruvate, catalysed by the preincubated enzyme preparations, relative to the conversions of those enzyme preparations, preincubated at 24 °C.

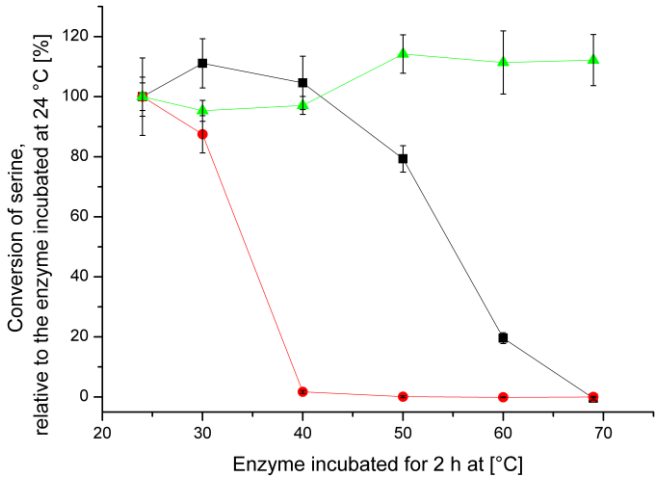


Figure 60: Residual activities of enzyme preparations after incubation at different temperatures. Preincubated DAAO preparations are incubated with D-serine, ProRCMS#31 is incubated with L-serine and an excess of DAAO. Circles: DAAO cell-free extract, Squares: immobilized DAAO, Triangles: ProRCMS#31.

The first astonishing result of this experiment is thatProRCMS#31 does not loose activity during 2 h at 70 °C. This temperature has to be considered very high, despite the high melting point, determined to be 69 °C.

D-Amino acid oxidase from *Rhodotorula gracilis* looses more than 10% activity when warmed to 30 °C for 2 h, and after 2 h at 40 °C the enzyme is almost completely deactivated. Compared to that, under the same conditions, the multipoint immobilized preparation of the same enzyme was stable at 40 °C, and looses only about 20% of its activity, at 50 °C. Even after 2 h at 60 °C, there is still some activity left.

This shows that the immobilization was successful and fits well to the results from Fernández-Lafuente *et al.*, exhibiting a drastic stabilization of this specific oxidase with the same method of immobilization. [256]

In a second experiment, it was shown that the immobilized material can be recovered from the conversion of D-serine to hydroxypyruvate, with no apparent loss of activity in a second run of the reaction under identical conditions (data not shown).

However, in the one-pot cascade reaction, when combining the immobilized oxidase with ProRCMS#31 and transketolase from *Geobacillus stearothermophilus*, there was

no 5-deoxyfructose product detected by TLC both at small scale (1 mL) and during an attempted scale-up (20 mL). Also, the beads with the immobilized enzymes showed signs of degradation. In the small-scale experiment, the beads had an increased tendency to agglomerate after the reaction and in the scale-up experiment, the beads changed their colour to dark brown. The inhibition of the catalyst can be attributed to the presence of transketolase, while the change of colour is to be attributed to an iron scaffold that was used to hold the silicon tubing in place, used for oxygen introduction into the reaction mixture (*vide infra*). This scaffold was visibly oxidized to iron oxide after 24 h in the reaction mixture probably due to corrosion from H_2O_2, O_2 and HCl. For this reason, it was replaced through a Teflon scaffold in later experiments.

Since the cascade reaction unfortunately did not show conversion, also at small scale, the immobilized material was not employed for subsequent experiments. Nonetheless, two options to stabilize the oxidase were retained in later cascade reactions: (i) exogeneous FAD was added at low concentrations (0.1 mM); (ii) the overall concentration of oxidase was kept relatively high (> 1 mg/mL cell-free extract) during preparative reactions, both to improve the oxidase stability.

4.2.4 Process Optimization C: Oxygen Supply

Since bubbling gas through the reaction mixture can lead to inhibition of amino acid oxidases, as a third possibility, alternative designs for oxygen-provision were tested to increase the operational stability of the enzyme.

Strong foaming and deactivation of the DAAO had also been observed during initial transketolase/DAAO/amino acid racemase cascade reactions at the 20 mL-scale with bubbling for aeration. Unfortunately, the tested immobilization technique proved to be incompatible with the cascade setup. For this reason, a milder method to provide oxygen to the reaction mixture was sought and found in the use of a semipermeable membrane tubing, internally pressured with 2 bar of pure oxygen.

4.2.4.1 Oxygen Supply *via* Air-Permeable Membranes

To assess the efficiency for introducing oxygen into a solution of different tubing materials, the oxygen concentration of an initially degassed solution after exposure to the tubing was measured. To measure the oxygen concentration, a photometric protocol was employed. The underlying chemical principle is that of the classical Winkler titration (Figure 61). [264,265]

Figure 61: Redox equations of Winkler's titration for the determination of dissolved oxygen.

During the Winkler protocol for the determination of dissolved oxygen, in the first step $MnCl_2$ is added to a basified sample, leading to quantitative oxidation of the manganese ions to MnO_2 by oxygen. Acidifying the solution stops the reaction and prevents further oxidation of Mn^{2+}. In the second step, KI is added, returning the manganese-(IV) back to its original redox state (II) and forming molecular I_2, colouring the solution pink. The classical method involves back-titration of I_2 with thiosulfate and endpoint detection through de-coloration of an added solution of starch.

However, it has been shown that the amount of formed I_2 can be directly determined photometrically on a large linear range. This simplifies the method and enables faster determination of dissolved I_2, further minimizing the loss of Iodine due to exposure to the atmosphere. [266] For efficiency-estimation of different oxygen-introduction methods, where large differences are expected and the goal is not to quantify the total concentration of dissolved oxygen but to qualitatively compare the different materials in their potency to introduce oxygen into a reaction mixture, this method is considered sufficient.

The fact that the absorption of dissolved iodine is linear over a large concentration range enables semi-quantitative comparison of different methods, with bubbling of oxygen and exclusion of oxygen as the two benchmarks (Figure 62).

Figure 62: Visible difference in the amount of formed MnO_2 during the modified Winkler protocol. One tube was stirred and bubbled with oxygen (O_2), one was stirred, while a semipermeable silicon membrane, internally pressured with 2 bar of oxygen, was submerged (Si) and the third one was stirred while sealed.

Three different tube materials were tested (Figure 63):

- Teflon: fluoroethylene propylene (FEP), internal diameter 1 mm, wall thickness 0.30 mm, MZ-Analysetechnik GmbH
- Teflon: polytetrafluoroethylene (PTFE), internal diameter 1 mm, wall thickness 0.30 mm, MZ-Analysetechnik GmbH
- Silicone: Silastic™, internal diameter 0.64 mm, wall thickness 0.28 mm, Dow Chemical Company

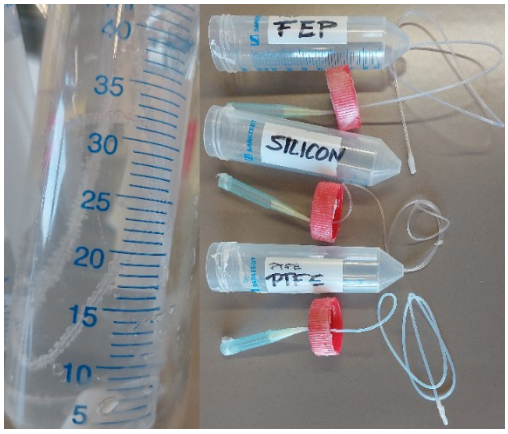


Figure 63: Setup of the oxygen dissolution tests with stationary gas bubbles visibly formed on the submerged tube (left).

36 cm of the tubing was pressurized with 2 bar oxygen and submerged into 50 mL of stirred deionized water for 5 min. Beforehand, oxygen was purged from the deionized

water by sonication and bubbling with argon for 8 min. Controls with bubbled oxygen and no introduction of additional oxygen were also prepared. For details of the procedure, see Chapter 5.4.4.

The results are displayed in Figure 64.

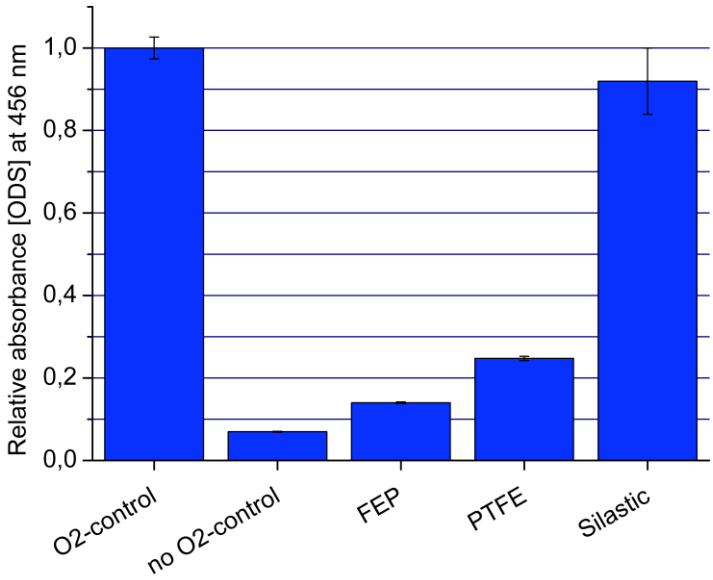


Figure 64: Relative absorbance of solutions of dissolved I_2, prepared after the method of Winkler, corresponding to different methods of introducing O_2 into water. O2-control: liquid bubbled with oxygen, no O_2-control: no oxygen provided, for the other three datapoints, O_2 was provided through a tubular membrane of the corresponding material.

The value for the control without additional oxygen lies below 10% of the value for the control, where oxygen was directly bubbled through. This is likely due to oxygen contamination those phases, where the tubes need to be opened and before the solutions are acidified. The value is considered low and shows that the experimental setup is well-suited to provide a semi-quantitative comparison, which was the goal of this experiment. For total quantification of the oxygen concentration, electronic oxygen sensors are recommended, however such were not available in-house. The total values are largely dependent on experimental parameters (mainly time of exposure and mixing) and the general trends were confirmed in a repetition (data not shown).

The second lowest value lies at 14% of the oxygen-control and belongs to the FEP tubing. This value is only 6% higher than the control without additional oxygen supply. The PTFE tube lies at 25% of the value of the oxygen control.

The absorption of the solution supplied with the silicon tube reached 92% of the oxygen control.

Concluding, it is very easy to introduce oxygen into stirred solutions, even unwantedly, however the efficiency range is broad. Tested PTFE tubes can introduce oxygen, however with very bad efficiency and are thus not suited for our reactions. The Silastic™ silicon tubing is well suited to provide oxygen to solutions, visible as tiny gas bubbes, forming on the surface of the tube. This was expected, since the tube is frequently used in HPLC setups with an external vacuum, to purge air from eluents running through the tube. However, the high efficiency in the reverse direction, reaching 92% of the bubbled oxygen control, was a pit surprising, especially since the bubbling oxygen represents an additional element of mixing in the experimental setup. As such, the Silastic™ silicon tubing is well suited for reactions with DAAO with regard to oxygen supply.

It should be noted that the formation of precipitated MnO_2 during the experiment lead to the browning of the tube, due to Mn^{2+} ions, penetrating the silicone and being oxidized to brown MnO_2 within the tubing material. MnO_2 is an efficient and thermostable quencher for H_2O_2 which is desired during most cascade reactions with DAAO. Additionally, silicones are known to possess active hydroxy groups on their surface or can easily be modified to do so. As such, this kind of tube represents an interesting carrier for coimmobilization of DAAO and MnO_2 similar as presented in Chapter 4.2.3 for Sephadex™ carriers. Such modified tubes would provide oxygen to where it is used up, channelling developing H_2O_2 to an inorganic thermostable quencher for employment in batch, or tubular reactors, depending on which surface is modified.

4.2.4.2 Optimized Scaffold for In-Flask Membranes

Additionally, the Silastic™ tubing was tested, fixed on a Teflon scaffold for reproducible submersion, as employed during one of the 20 mL transketolase/DAAO/amino acid racemase cascade reactions (Figure 65).

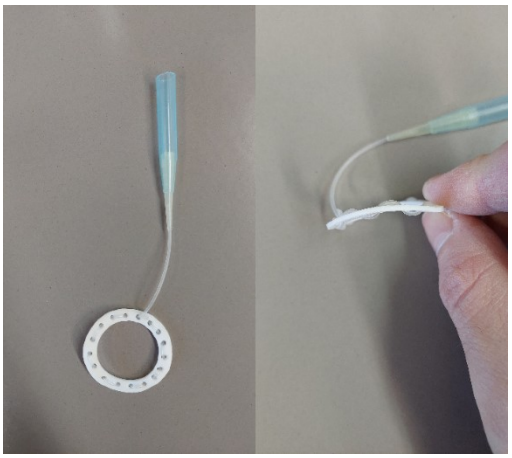


Figure 65: Top and side view of the employed Teflon scaffold with the interwoven silicone tubing.

This scaffold was employed to sink the pressurized tubing to the bottom of a round bottom flask, as displayed in Figure 66 holding it in place, while enabling stirring with a flying stirring bar, preventing the free tubing from wrapping around the stirring bar.

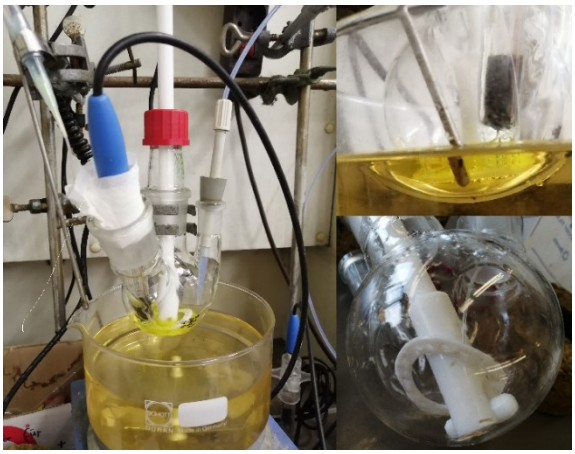


Figure 66: Teflon scaffold with interwoven silicon tube as employed in early cascade reactions. During the reaction (top right) gas bubbles visibly formed, stationary on the surface of the tube.

In addition to the technology, developed for 20 mL-scale, a second strategy for small scale was developed. For micro-scale (250 µL – 2 mL) reactions in Eppendorf vials, the standard procedure for testing oxidase reactions is to leave the vials open, or to puncture the caps of the vials. This facilitates intrusion of air into the reaction mixture *via* the surface of the liquid phase. A disadvantage of this method is that at higher reaction temperatures and during longer reaction intervals, evaporation of the aqueous

phase can become a significant factor. Evaporation leads to concentration of the reaction mixture, which in turn can have a negative effect on reaction components and makes reliable quantification during analysis more difficult.

4.2.4.3 Micro-Scale Oxygen Supply

Many reactions require an exchange of the gas phase, for example the transketolase/DAAO cascade with serine as a substrate, where the oxidase requires fresh oxygen and carbon dioxide should be able to leave the mixture. However, in some reactions, for example the oxidation of amino acids with amino acid oxidase, oxygen gas is only required to enter the reaction as a substrate, but no gas is required to exit the reaction vial.

For these applications, a 2 mm thick, self-adhesive ethylene-propylene-diene/styrene-butadiene rubber foil can be glued onto the surface of the caps of the Eppendorf vials, turning it into a septum. Oxygen can then be introduced through a needle into the reaction mixture without continuous bubbling (Figure 67).

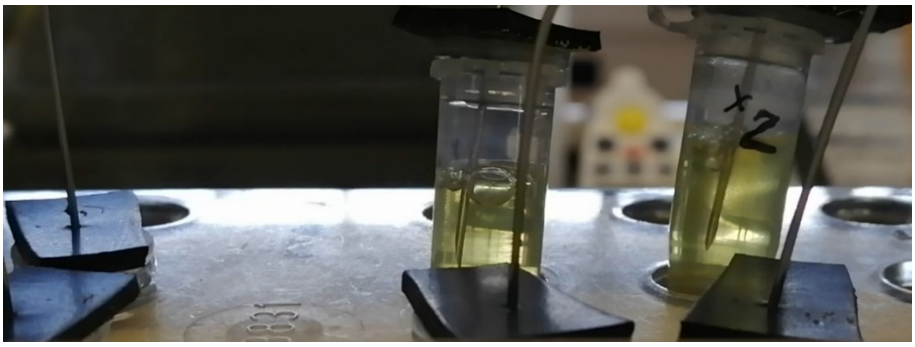


Figure 67: Small-scale reactions with oxygen supply through a septum.

The method enables parallelization of small-scale reactions with defined oxygen provision with minimum evaporation of the aqueous phase and measurement of the reaction velocity through measurement of the volume of introduced oxygen while avoiding enzyme deactivation through continuous bubbling.

4.2.4.4 Facile Oxygen Supply Setup for Synthesis Parallelization

It should be noted that these methods provide specified solutions for specific demands. They offer advantages over conventional methods and are interesting over a broad range of scales. However, they are also more demanding with regard to material and the time needed for setup to guarantee reproducibility. For this reason, in most cases we applied punctured caps of Eppendorf vials at small scale, while for 20 mL-scale a

continuous stream of pure oxygen was supplied above a reaction mixture in a tilted tube, mounted in a shaker at 600 rpm (Figure 68).

Figure 68: Facile and effective oxygen supply to cascade reactions at 20 mL-scale.

This system enables purging of formed CO_2 during cascade reactions with transketolase and supplies fresh oxygen through the large agitated surface of the reaction mixture, while also being easy and fast to operate. For cascade reactions where sampling in short intervals was required, a small window was cut into the side of the tube, easy to open and close. This way, the reaction vial can remain fixed inside the shaker and connected to the oxygen supply.

4.3 Application of Panel Racemases A: Transketolase Cascades

Transketolase is a central metabolic enzyme from the pentosephosphate pathway, naturally transferring the terminal ketol-unit from a C5 - C7 (3S)-ketose ketol donor to a C3 - C5 (2R)-aldose ketol acceptor in a reversible equilibrium reaction (see theory part). Already in the course of its purification in the 1950s it has been found out that transketolases accept hydroxypyruvate as a ketol donor, liberating the carboxy-group as CO_2. [267] The decarboxylation is practically irreversible and drives the transketolase reaction equilibrium to the side of the elongated ketose product. [268] Exploiting this driving force, even less reactive aldehyde acceptor substrates can turn out to be synthetically feasible. As such, the finding that hydroxypyruvate is a transketolase donor enables a lot of new, synthetically useful C-C-bond forming reactions, elongating diverse aldehydes by a two-carbon atom fragment with high stereoselectivity.

This methodology is superior to classical chemical methods, including hydroxide substitution with cyanomethyltrimethylphosphonium iodide, Wittig olefination, and Masamune-Sharpless reaction, as discussed by Wohlgemuth, since all of these methods require stoichiometric amounts of activating high energy reagents. [269]

It also sparked interest in a lot of mutational studies about transketolase, ever extending the scope of acceptor and donor substrates (*vide infra*).

The irreversible nature of the CO_2-formation leads to a shift in the pH, enabling universal colorimetric assaying and high throughput screening of transketolases by applying pH indicators (Figure 69). [270]

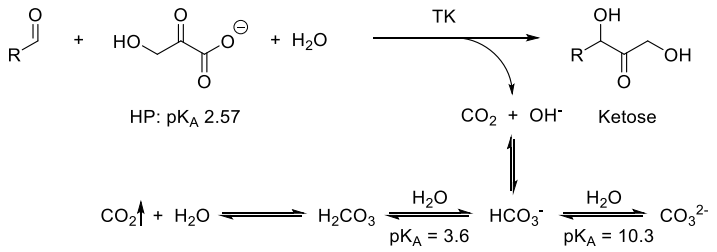


Figure 69: Reaction equations underlying the apparent pH shift, initiated by the transketolase (TK) catalysed coupling of aldehyde and hydroxypyruvate (HP), releasing CO_2 and consuming a proton from the medium.

The liberation of CO_2 into the aqueous solution leads to the formation of a carboxyl anion. Recombination with an acceptor aldehyde ultimately consumes a proton, due to the high pK_A of the formed alcohol group. The resulting solution contains carbon dioxide and a hydroxide ion, leading to complex carbonic acid buffer equilibria. The

loss of CO_2 from the solution drives the equilibrium. Starting the assay at a pH near the transketolase optimum at low buffer concentration leads to hydroxypyruvate consumption to be accompanied by a pH increase. Even though reaction dynamics of such a complex pH-based assay is depending on a lot of variables and thus, is hard to control at pH values above ~4, it represents a highly valuable, universal tool for the skilled experimenter to evaluate the transketolase acceptor substrate spectrum. The assay is even extendable to other 2-ketocarboxylic acids (*e.g.* pyruvate, 2-oxobutyrate, 2-oxovalerate) [121,271] as donors, which are no, or only very bad substrates of wild-type transketolases, but increasingly important for raising variant diversity.

Controlling reaction stoichiometries in synthetic reactions with hydroxypyruvate is challenging, mainly because it is not stable in aqueous solutions. Decomposition velocity and fate vary depending on many variables, for example temperature, pH and the availability of oxygen (Figure 70). [272]

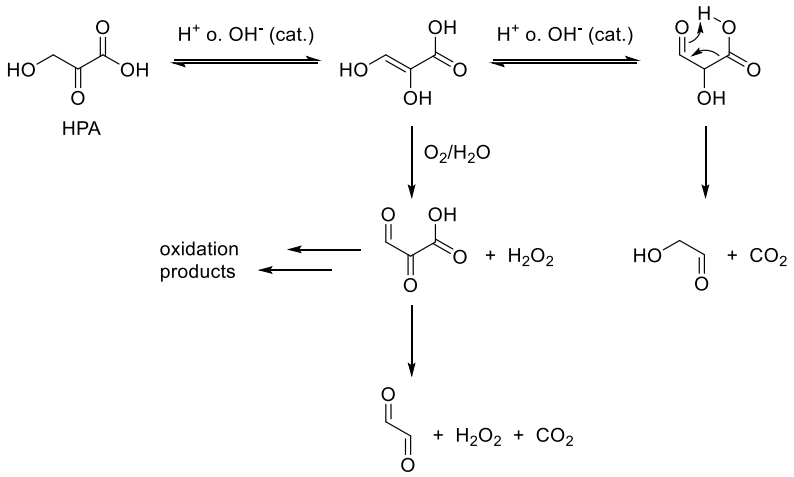


Figure 70: Different pathways of the chemical decomposition of hydroxypyruvic acid (HPA). [272]

Additionally, hydroxypyruvate can decarboxylate enzymatically, for example catalysed by decarboxylases and transketolase itself, even in the absence of an acceptor aldehyde. In the latter case, transketolase can form L-erythrulose from hydroxypyruvate and the primary decomposition product glycol aldehyde. [121,273,274,275]

For these reasons, reaching full conversion with one equivalent hydroxypyruvate, added at the start of the reaction, at extended reaction durations is not likely. Thus, *in situ* generation from cheap precursors and decreasing residence time of hydroxypyruvate is desired, increasing the total yield relative to the pricey commercial substrate. Therefore, employing cheap serine as a cheap precursor for *in situ*

generation of hydroxypyruvate is advantageous. However, it should be kept in mind that also serine itself may be dehydrated to pyruvate in reaction media containing cell-free extract (*vide supra*). Such *in situ* generation of hydroxypyruvate from serine has been achieved in biocatalytic cascades, deaminating L-serine with transaminases, or D-serine with DAAOs (Figure 71). [233]

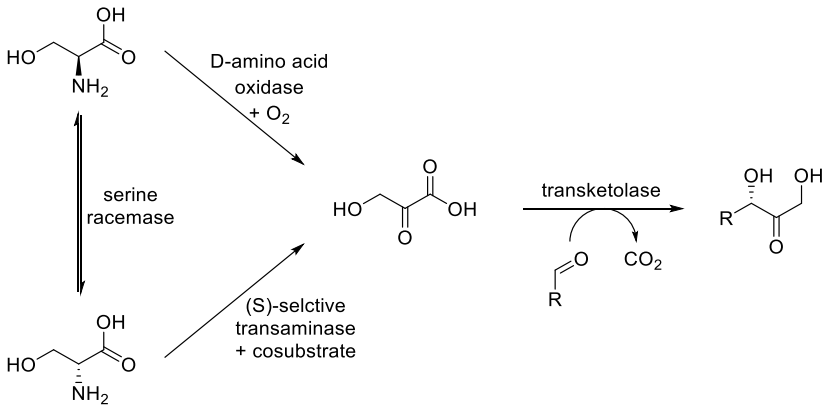


Figure 71: Complementary serine deamination with transaminase and DAAO and subsequent ketol transfer. [233]

It should be noted that amino acid dehydrogenases can also be applied for hydroxypyruvate generation, however this exceeds the scope of this work. Transaminase and transketolase have recently been synthetically coupled for *in situ* hydroxypyruvate generation by the groups of Ward, Hecquet, and Fessner. [276,277,278,279] However, already in 1999 transaminase has been applied for *in situ* production of hydroxypyruvate for the synthesis of ^{14}C-labeled heptuloses [280] and in 2006, the authors of a publication about the detection of *N*-phenylglycolohydroxamic acid in spinach leaf homogenate reasoned the metabolite being formed from nitrosobenzene, pyruvate, and serine through the action of transaminase and transketolase. [281] Also, transaminase cannot only cooperate with transketolase in providing hydroxypyruvate, but also in the conversion of the ketol product into a chiral amino alcohol (Figure 72), [282] or both transformations simultaneously in one-pot. [276]

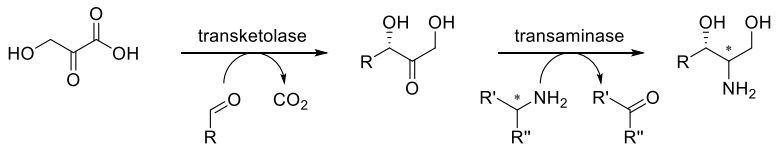


Figure 72: Transamination of transketolase products for the generation of chiral amino alcohols. [282,276]

In situ production of hydroxypyruvate with mostly (S)-selective transaminases enables the employment of the generally cheaper L-serine enantiomer, even though racemic serine is even less expensive. Another advantage of transaminases is that there are

relatively stable enzymes available, which combine well with the thermostable transketolase from *Geobacillus stearothermophilus*. However, transamination of serine requires a ketone cosubstrate, serving as an amine acceptor, decreasing the atom economy and increasing the process costs. Frequently used amine acceptors are α-ketoglutarate and pyruvate, the latter being less expensive, but often found to be less efficient and not compatible with transketolases for pyruvate conversion. The equilibrium of the transamination is driven by the subsequent irreversible decarboxylation of the hydroxypyruvate during ketol transfer.

D-Amino acid oxidase has also been used in many examples of *in situ* hydroxypyruvate production in transketolase reactions. [280] Contrary to transaminases, it does not require any other cosubstrate than molecular oxygen, leading to a high atom economy. Additionally, it has its own equilibrium driving force, because the oxidation reaction is practically irreversible.

4.3.1 Acceptor Substrates for Transketolases

In recent years, the already wide transketolase substrate spectrum has been further diversified, mainly through rational design and directed evolution. [117,118,119,120,121,122,123,124,125,126,270,283,284,285,286,287,288,289] Non-natural sugars are chemically challenging to produce and transketolase products can be converted into enantiopure amino alcohols through the action of transaminases and reductive aminases (RedAm). This potential renders the enzyme a valuable tool for the exploration of the chemical space of this pharmaceutically important class of compounds. For example, transketolase variants suited to convert aromatic acceptor aldehydes hold potential in the synthesis of chloramphenicol (broad spectrum antibiotic, Figure 73) analogues. [288]

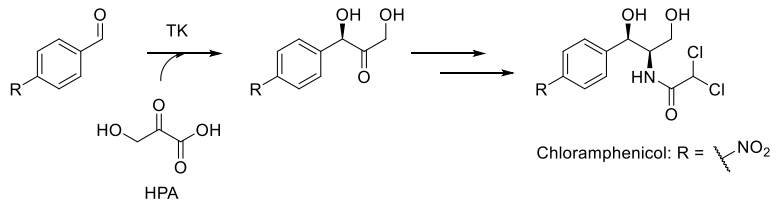


Figure 73: Net-reaction equations of transketolase assisted synthesis of Chloramphenicol analogue precursors from various benzaldehydes with hydroxypyruvic acid (HPA) as a transketolase donor. [288]

Figure 74 gives a non-exhaustive overview over the diverse acceptor substrate scope, reported in the literature. [118,120,122,126,270,283,284,285,286,287,288,289,290,291,292,293,294,295,296,297,298,299]

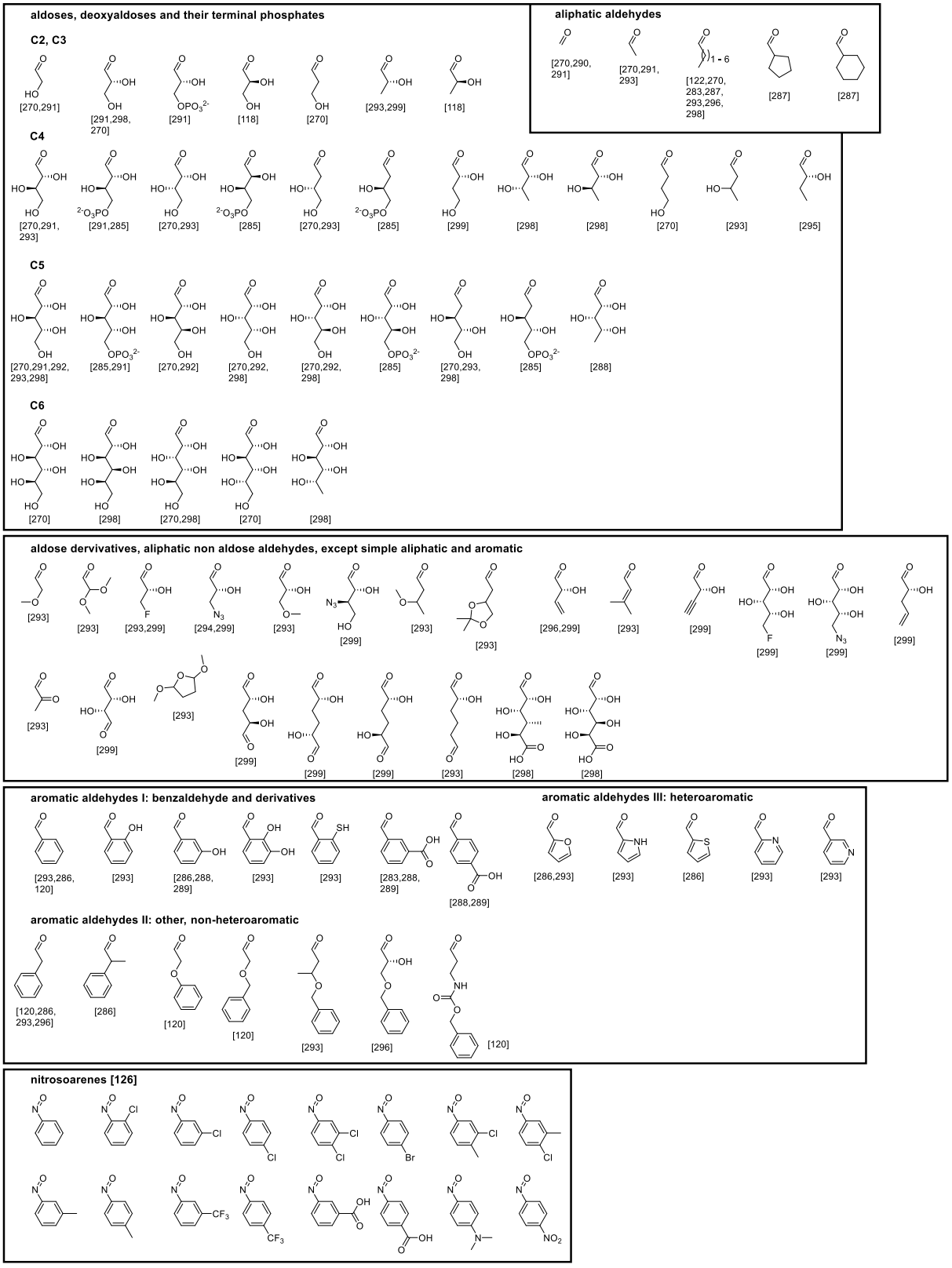
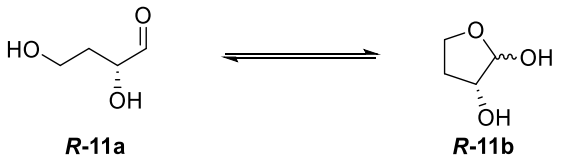


Figure 74: Transketolase acceptor substrates. [118,120,122,126,270,283,284,285,286,287,288,289,290,291,292,293,294,295,296,297,298,299]

To demonstrate the compatibility of enzymes from the screened ProRCMS amino acid racemase panel with transketolases and *in situ* hydroxypyruvate production, we

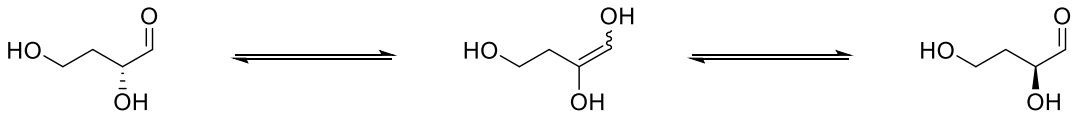
selected 2(*R*),4-dihydroxybutanal (***R*-11a**) as an acceptor substrate, in aqueous solutions present in an equilibrium with its furanose-form (***R*-11b**) (Figure 75).



***R*-11a** ***R*-11b**

Figure 75: Acceptor substrate 2(*R*),4-dihydroxybutanal (**11a**) in equilibrium with its hemiacetal form (**11b**).

From NMR analysis in D_2O the equilibrium can be estimated to lie on the side of the furanose-form (~ 97%). Additionally, the aldehyde can possibly tautomerize to the enol, which can lead to a slow racemisation of the hydroxy group at C-2 (enol peaks not found in NMR).



***R*-11a** (E/Z)-enol ***S*-11a**

Figure 76: Keto-enol tautomerism of 2,4-dihydroxybutanal (**11a**).

The velocity of the tautomerism is dependent on pH, temperature and the position of the equilibrium, and has been observed in our group to be slow during cascade reactions under mild conditions.

The unprotected aldehyde is of limited stability in aqueous solutions upon concentration and has been observed to form a colourless amorphous precipitate, probably due to oligo-, or polymerization. For this reason, the substrate was produced and stored either as the racemic *syn*-3-hydroxy-2-methoxytetrahydrofuran (**12**), or the 2,3(*R*)-diacetoxytetrahydrofuran (***R*-13**) and freshly deprotected and prepared as aqueous solution prior to the biocatalytic reactions (Figure 77).

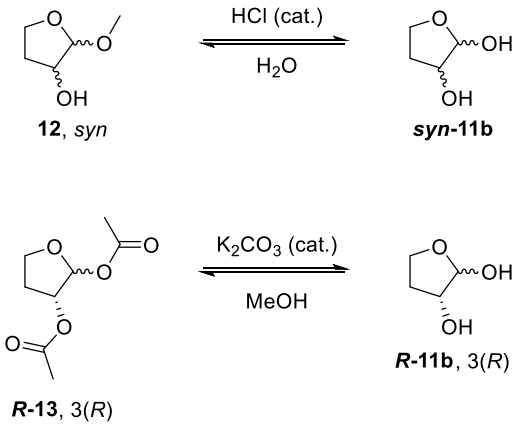


Figure 77: Deprotection of stored derivatives of ***R*-11b** (3-deoxy-L-*glycero*-tetrose).

Since the total amount of dissolved acceptor could not be reliably deduced from simple stoichiometric calculations, the concentration of such solutions was determined *via* NMR, employing an authentic reference material (*vide infra*).

Four synthetic routes to the substrate (**11**) have been explored. **12** (*syn*-3-hydroxy-2-methoxytetrahydrofuran) has been produced from 2,3-dihydrofuran *via* dihydroxylation and ***R*-13** (2,3(*R*)-diacetoxytetrahydrofuran) reductively from 2(*R*)-hydroxy-γ-butyrolactone, described in the following chapter.

4.3.2 2,4-Dihydroxybutanal Acceptor Synthesis – Different Routes

The four different synthetic routes to the aldehyde acceptor have been explored (I – IV) as displayed in Figure 78.

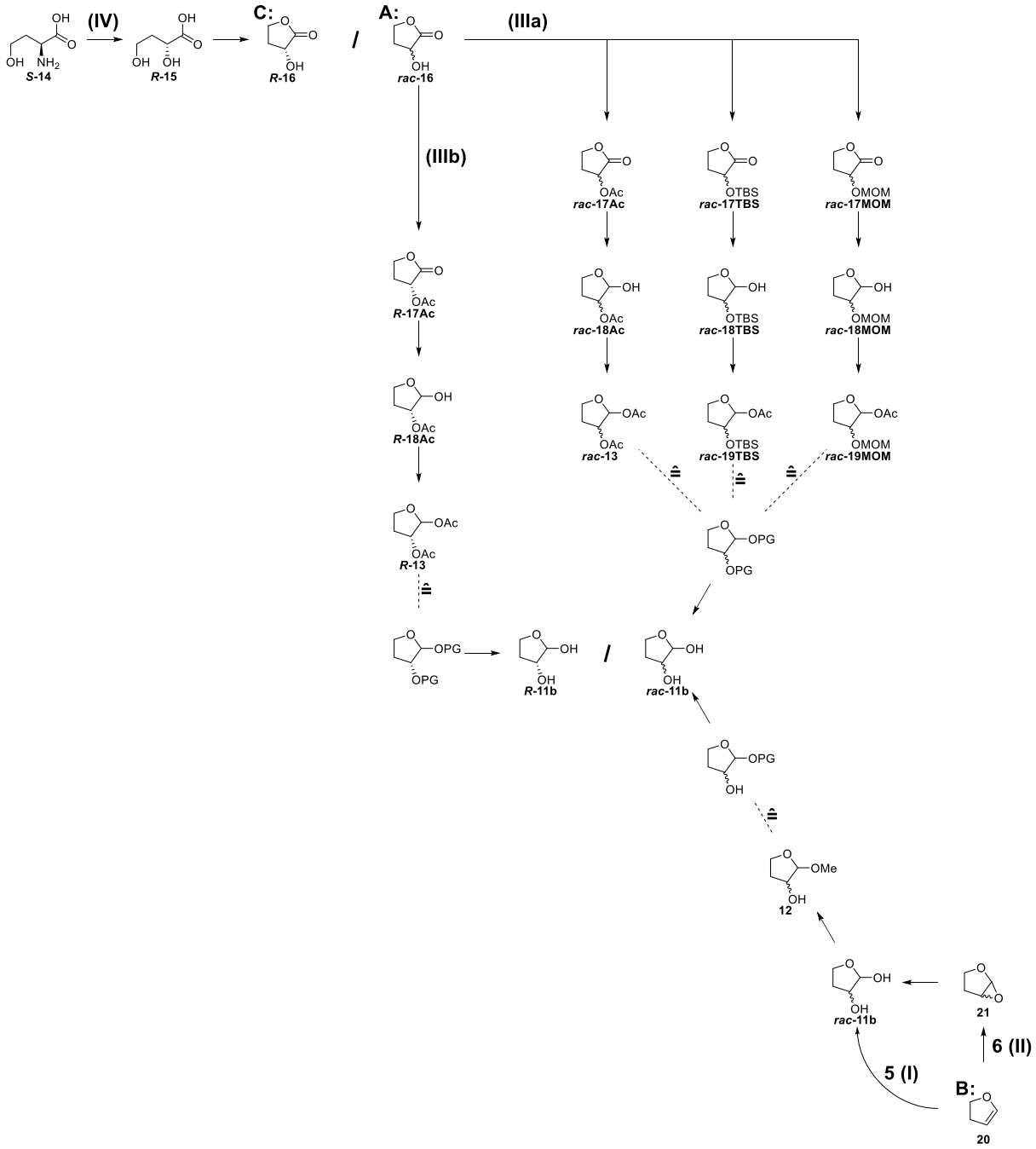


Figure 78: Exploration of different routes toward the transketolase acceptor substrate **11** (2,4-dihydroxybutanal). PG: protecting group, TBS: *tert*-butyldimethylsilyl, MOM: methoxymethyl, Ac: acetyl.

4.3.2.1 Routes **(I)**: *syn*-Dihydroxylation

The original route **(I)**, established at large scale in our group by Zimmermann, [299] and originally based on a publication by Kobori *et al.*, [300] is the dihydroxylation of 2,3-dihydrofuran using hydrogen peroxide, and OsO_4 as a catalyst with methylation of the crude product for purification and storage (Figure 79).

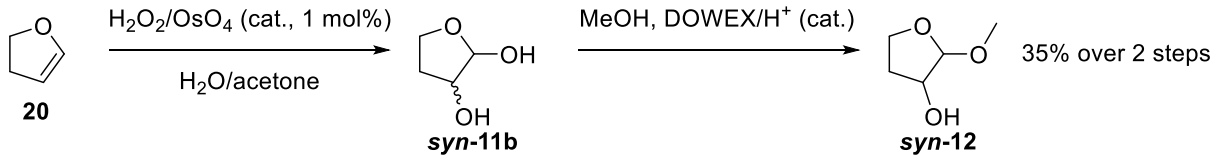


Figure 79: Route (I): OsO_4-catalysed dihydroxylation of 2,3-dihydrofuran (**20**) and subsequent acetalization.

The yield is particularly low, mainly due to low catalyst loading. The given protocol of Zimmermann describes the conversion of 5 g of **20** (2,3-dihydrofurane) to racemic ***syn*-11b** (3-deoxy-D/L-*glycero*-tetrose) employing 0.5 mL of OsO_4 (5%, aq.) including direct column chromatography with 84% yield. The protocol states that 5 g of **20** (2,3-dihydrofurane) is equal to 7.13 mmol, which is incorrect. In my opinion, it is unlikely that the product was produced this way due to the following reasons. Firstly, the given catalyst loading (0.5 mol%), calculated for 5 g of **20** (2,3-dihydrofuran) is particularly low, compared to the literature. Even though there are protocols with catalyst loadings as low as 0.2%, those protocols achieving high yields usually employ between 5 and 10 mol% OsO_4. [301,302] Interestingly, the literature protocol from Kobori *et al.*, published 5 years before the Zimmermann thesis, does not give any catalyst loading, but instead the term “catalytic osmium tetroxide in acetone/water (20 mL, 1:1)” is used. [300] In their discussion, they even state having used *N*-methylmorpholine-*N*-oxide as an additional peroxide conductor and that their protocol “gave, after silica gel chromatography, aldehyde **23** [corresponds to 2,4-dihydroxybutanal (**11**)] in high yield”. In the experimental section of this publication, they describe the conversion of 10 g of **20** (2,3-dihydrofuran) with 88% yield and there is no mention of *N*-methylmorpholine-*N*-oxide. In our experiments, the use of *N*-methylmorpholine-*N*-oxide as in a protocol from Nicolau *et al.* did not increase the yield, but complicated workup. [302] Presumably, Kobori *et al.* used 10 mL of commercially available aqueous 4 – 5% solution of OsO_4, diluted (1/1) with acetone to dissolve the 10 g of **20** (2,3-dihydrofuran) and then started the addition of H_2O_2, corresponding to a catalyst loading of 4 – 5%.

The experimental protocol from Kobori's paper also does not include the direct column chromatography of the unprotected lactol, described in the discussion part of the same paper and the Zimmermann thesis, leading to the second reason for doubting the application of the protocol. During our initial experiments, following the original Zimmermann protocol, column chromatography of the crude dihydroxylated hemiacetal product was attempted, but led to unprecedented adsorption to the silica gel and smearing of the product from the column. This culminated into a drastic loss of product and almost no yield at all.

Thus, there is good reason to believe that the original protocol, applied by Zimmermann and Kobori, included approximately 5 mol% of OsO_4 and direct acidic acetalization of the crude hemiacetal with methanol and some water absorbent, and column chromatography of the protected acetal. By following a similar procedure, but employing only 1 mol% of OsO_4 we obtained 35% of product acetal and it is plausible that an increase of the catalyst to 5% would lead to shorter reaction times (*vide infra*), decreased free hydrogen peroxide and thus, increased yield.

The electron rich olefin (**20**) has a tendency to polymerize, as observed in the NMR-spectra of an old batch and in a H NMR-sample of a new batch, after storing the sample in $CDCl_3$ at r.t. over the weekend (Figure 80).

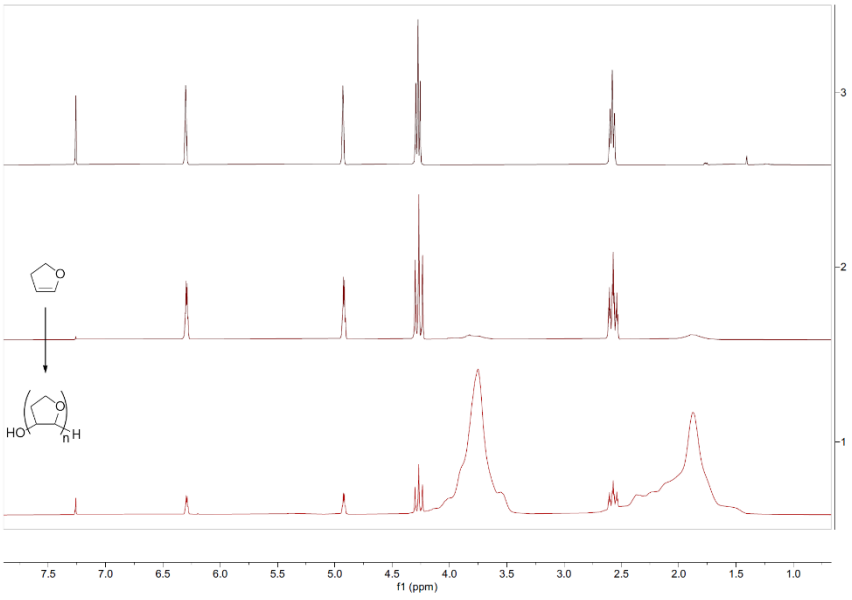


Figure 80: ^{1}H NMR spectra of fresh and aged samples of commercial 2,3-dihydrofuran. 1: dihydrofuran after > 50 h in $CDCl_3$, 2: neat, refrigerated dihydrofuran, after 1.5 years of storage, 3: fresh sample of a new batch.

The polymerization of dihydrofuran in the NMR sample of the fresh batch likely was acid catalysed in acidic $CDCl_3$ with about 97% of the substance polymerized. A

plausible product, fitting to the NMR spectrum is displayed in Figure 80. The mechanism of polymerization in the neat refrigerated substance over time is unknown. In general, polymerization proceeds catalytically with an acid, a base, a radical, or a nucleophile, which is plausible to occur under the reaction conditions of our dihydroxylation. After > 50 h in $CDCl_3$, approximately 20% of **20** had polymerized (Figure 80 - spectrum 2) according to NMR integration. Thus, a similar polymerization, or light driven cycloaddition might also take place during the overnight reaction with OsO_4 and hydrogen peroxide, lowering the yield and complicating workup. Also, since these side reactions take place over time, the reaction benefits from larger catalyst loadings and thus, shorter reaction durations.

Even though the yield still seems low (35%), it should be noted that the substrate 2,3-dihydrofuran is very cheap and OsO_4 is very toxic, hazardous and expensive. Also, hydrogen peroxide is even less expensive. The initially sought 2 mol% could not be employed, because one of the two commercial ampules (4% OsO_4, aq.) broke at the wrong position when opened due to an error in manufacturing (it was replaced by the manufacturer). When calculating the yield with respect to the catalyst, the yield is even higher (> 1.5-fold) when assuming the literature protocols employed 4 – 5 mol% of catalyst. The initial product is the racemic *syn*-furanose ***syn*-11b** (3-deoxy-D/L-*glycero*-tetrose), with the anomeric hydroxide slowly racemizing, due to the hemiacetal being in slow equilibrium with its open aldehyde form. It should be noted that the 2-hydroxy-group is also being racemized, but even slower, due to the tautomerism of the open aldehyde (*vide supra*).

When assessing route **(I)** and comparing it to the other routes that were explored in this work, it should initially be noted that it is the shortest route with the highest atom economy (perfect on paper when ignoring the catalyst), with respect to the racemic product. *Via* this route, a fridge stable derivative of 2,4-dihydroxybutanal can be obtained within 2 – 3 days. However, it should be noted that the route has zero enantioselectivity, which decreases the atom economy with respect to the required enantiomer 2(*R*),4-dihydroxybutanal substrate of transketolase. Also, OsO_4 should be considered a serious hazard, even when employed as a 4% aqueous solution, as it can easily pass through the barrier of the skin from the aqueous phase. [303] For this reason, nowadays OsO_4 is rarely employed as standard reagent in synthetic

laboratories and certainly not considered green chemistry, even with respect to the high atom economy and benign conditions and solvents.

4.3.2.2 Routes **(II)**: *anti*-Dihydroxylation

The second route **(II)** that was considered also started from 2,3-dihydrofuran and employed the standard epoxide forming reagent *m*-chloroperbenzoic acid for dihydroxylation (Figure 81).

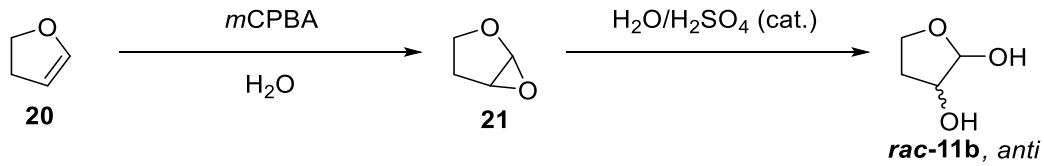


Figure 81: Route IV: Dihydroxylation of **20** (2,3-dihydrofuran) *via* intermediate epoxide (**21**) generated by the addition of stoichiometric *m*-chloroperbenzoic acid (*m*CPBA). Opening **21** (epoxide) under acid catalysis leads to formation of ***rac*-11b.**

A literature procedure for conversion of cyclopentene in water was followed. [304] The initial epoxide was opened in an acidic aqueous solution, yielding the *anti*-furanoses of ***rac*-11b**. The crude product was subjected to column chromatography without acetalization and yielded a colourless waxy solid. Even though NMR analysis proved the product to be present, additional aromatic peaks were visible and the liquid product presumably was not fully separated and adsorbed to voluminous, colourless, waxy *m*-chlorobenzoic acid salt, concluding from appearance, smell, and NMR. Since purification of the *syn*-diol from route III was achieved, the epoxide route was not further explored, due to its poor atom economy. However, it should be noted that separation probably is feasible after acetalization. Also, dihydroxylation *via* epoxidation with *m*-chloroperbenzoic acid represents a standard protocol in organic laboratories and thus such a procedure would be feasible for student practice.

4.3.2.3 Routes **(IIIa)**: Routes Involving Reduction of a Lactone Precursor

All other routes are based on the selective reduction of α-hydroxy-γ-butyrolactone (**16**) derivatives to the corresponding lactols. The racemic lactone ***rac*-16** is commercially available at a reasonable price, however around 5 to 10 times more expensive than **20** (2,3-dihydrofuran). The route is especially attractive, because the α-hydroxyl stereocentre is already defined in the lactone precursor (**16**), contrary to the olefine (**20**) and preserved through the reduction, giving an opportunity to produce the specific 2(*R*),4-dihydroxybutanal (***R*-11**) substrate for the transketolase reaction.

In general, reduction of acids and acid derivatives is difficult to terminate at the stage of the aldehyde, because aldehydes are more prone to being reduced than acid derivatives due to their higher electrophilicity. Standard organic chemistry textbooks provide methods for the reduction of esters, cyclic ester lactones, amides, and acid chlorides, mostly relying on the formation of a stable tetrahedral intermediate that cannot be attacked by the reducing agent. For example, a standard methodology for the reduction of lactones in synthetic laboratories is the reduction with diisobutylaluminium hydride (DIBAL-H), where the initially formed tetrahedral intermediate is stable at - 70 °C in nonpolar solvents (Figure 82).

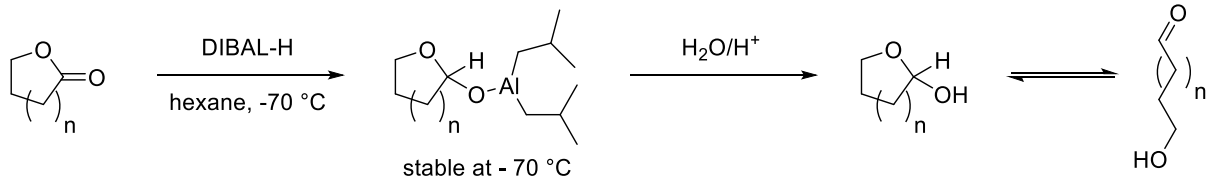


Figure 82:Reduction of lactones with diisobuylaluminium hydride (DIBAL-H), n = 1 - 3. [305]

Due to the stability of the tetrahedral intermediate, a reactive aldehyde is not formed until after hydrolysis, which also quenches excess DIBAL-H. However, the often-required excess of hydride lowers the atom economy, and alanes and aluminates hydrolyze to mixed aluminium hydroxides that form gels and emulsions, which are difficult to workup, especially when polar products are to be isolated. In this work, DIBAL-H reduction of the *tert*-butyldimethylsilyl (TBDMS)-protected lactone was attempted. Even though TLCs during the reaction looked promising, during workup, difficulties were encountered and the ^{1}H NMR of the product after column chromatography showed only traces of the hemiacetal. Nonetheless, a more successful protocol including *in situ* protection of the hemiacetal and an optimized extraction protocol may be developed in future.

Another method that has been shown to reduce lactones to lactols is reduction with diisamylborane. [306]

2016, Gonzalez *et al.* published the reduction of sugar lactones with triethylborohydride. [307] They claimed to be able to reduce several protected sugar γ- and δ-lactones with protected hydroxyl groups (*e.g.* **24**) to the corresponding lactols (*e.g.* **25**) in high yields with minimum diol formation, while the unsubstituted γ-butyrolactone (**22**) was fully reduced to the diol (**23**) with their protocol (Figure 83).

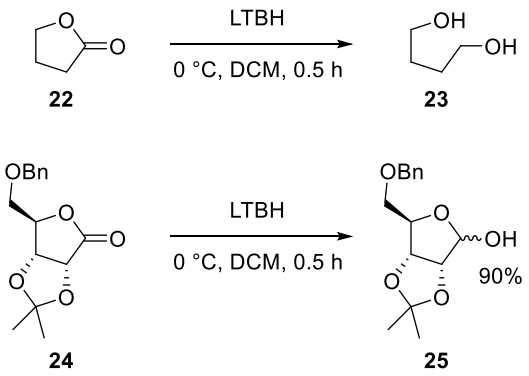


Figure 83: Reduction of lactones with lithium triethylborohydride (LTBH), Bn: benzyl. [307]

Reduction of **22** to **23** (diol) in THF is known since the 1980s. [308] Because, contrary to the substituted substrates, the unsubstituted γ-lactone did not yield the lactol, also in DCM, the authors reasoned that the exocyclic oxygen atoms chelate the borane reagent and that this is critical to the reduction.

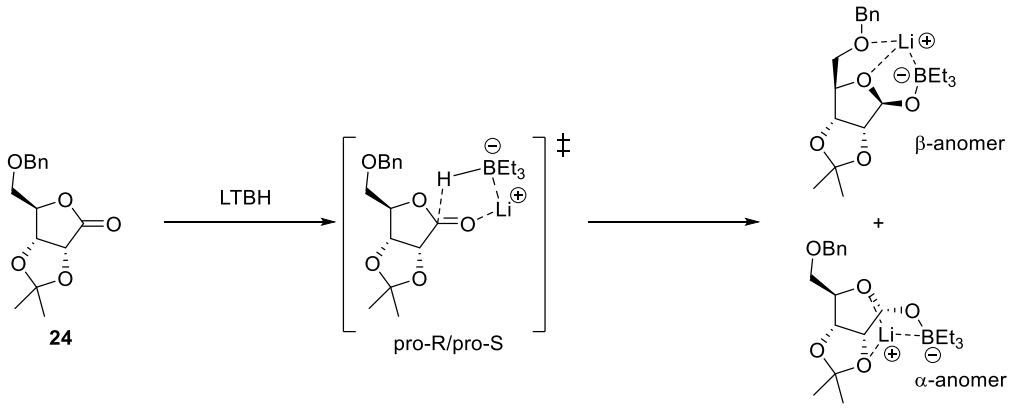
Comparing lithium triethylborohydride with the standard borohydride sodium borohydride, there are three main differences. Firstly, the reagent is far more reactive, owing to the inducting effect of the three alkyl substituents rendering it 10 000 times more nucleophilic than lithium borohydride. [308,309] The lithium counterion also has a profound effect on the reactivity, as it is a strong and hard Lewis acid, polarizing carbon-oxygen double bonds, activating them for the nucleophilic attack. For example, while most esters are not reduced with sodium borohydride (except when electron poor, e. g. nitrophenyl esters), this reaction is possible with lithium borohydride. These two activating effects render lithium triethylborohydride a very strong reducing agent, able to reduce carboxylic acids and acid derivatives with ease. Secondly, the three alkyl groups are sterically demanding, however the increased reactivity outweighs the steric hinderance during the initial attack of the reagent on an electrophile. For stereo- and regioselective reductions, also Selectride® is available, in which the ethyl groups are replaced by *sec*-butyl groups for an even more increased steric demand. The third difference is its increased solubility in apolar solvents. While sodium borohydride is

already hardly soluble in diethyl ether, [310] lithium triethylborohydride is well soluble in DCM.

The question of why the lactol can be isolated boils down to why the tetrahedral intermediate after the initial nucleophilic attack of the reagent is stabilized, or in other words how the free aldehyde is prevented to be formed. This circumstance is even more interesting due to the fact that the unsubstituted lactol cannot be isolated, thus its intermediate being much less stable, compared to the substituted one. Moreover, the authors claim that the reduction can be carried out at - 78, 0, and 20 °C ruling out the idea of a relatively small stability difference between the tetrahedral intermediates of substituted and unsubstituted substrates. The possibility to carry out the reaction in apolar solvent clearly plays a critical role, since only 30% yield was isolated when the reduction of 5-*O*-benzyl-2,3-*O*-isopropylidene-D-ribono-1,4-lactone was carried out in THF, compared to 99%, when carried out in DCM.

Molecular dynamics simulations of the reduction of acetone with sodium borohydride in methanol suggest that the sodium ion is never separated from the negative charge, located on the boron atom throughout the entire reaction. [311] Similar results have been obtained when studying $LiAlH_4$ in toluene and THF, however the addition of crown ether led to the separation of Li^+ and AlH_4^-. [312]

These observations give enough reason to believe that the complexation of the lithium ion by the oxygen atoms of the tetrameric intermediate leads to a large stabilization of the cyclic state, preventing the boron species to dissociate and the ring to open.



24

Figure 84: Reaction mechanism of the reduction of **24** with lithium triethylborohydride, leading to a tetraedric intermediate with proposed geometries of lithium ion-coordination.

The reaction leads to a proportion of α/β-anomers of 1:4. In absence of the 2,3-*O*-isopropylidene substituent, the lactol is still formed, but only at 30% yield.

However, it cannot be completely ruled out that the reasons are rather to be found in the difference of the equilibrium position between the lactol and the open aldehyde. In apolar solvents, the equilibrium is on the side of the lactol ring and substituents further influence this equilibrium position. This possibility is supported by the fact that longer reaction times lead to increased diol formation.

Overall, the system is an interesting model for such lactone reductions that could further be elucidated by testing different lactone substituents, borohydride substituents, counterions, solvents and even apply *in situ* NMR to elucidate the nature of reactive intermediates.

We reasoned that such reductions should also be possible with 2-hydroxy-γ-butyrolactone and that different oxygen containing protecting groups should have different effects on the outcome of these reductions.

To test this hypothesis, three different hydroxide protecting groups were introduced: *tert*-butyldimethylsilyl (TBDMS), methoxymethyl (MOM), and acetyl (Ac). Performed reactions for hydroxide protection are displayed in Figure 85.

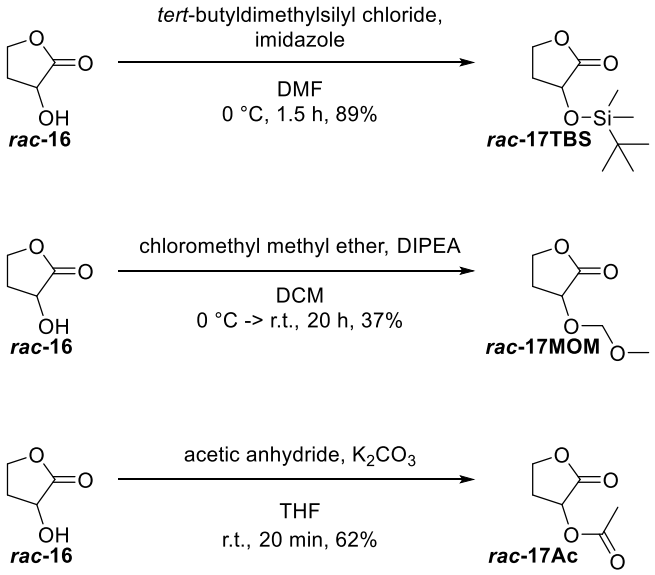


Figure 85: Introduction of TBDSM, MOM, and Ac protecting groups.

The silyl group is easy to introduce and a 2-TBDMSO derivative of γ-butyrolactone was not tested in the paper, [307] however only has one-potentially coordinating oxygen atom with a bulky protecting group and the worst atom economy.

The MOM-group is considered especially interesting, because it mimics part of a crown ether, thus providing two possible coordination centres and containing an additional

acetal moiety, reasoned to be stable under the basic reaction conditions. The atom economy is considered acceptable, but the reagent for the introduction of this protecting group (chloromethyl methyl ether) is a volatile cancerogenic substance.

The acetyl group also contains two potential coordination centres, has a good atom economy and reagents (acetic anhydride, acetyl chloride, vinyl acetate) for the introduction of acetyl groups are cheap. However, the acetic acid ester is prone to reduction itself. Additionally, it holds great potential for kinetic resolution of the racemic γ-butyrolactone employing lipase prior to reduction.

All three derivatives were successfully reduced to the corresponding lactol at - 80 °C in DCM. The lithium triethylborohydride was commercially available only as a solution (1 M) in THF, which is unfortunate, since THF is expected to have a negative effect on the selectivity of the reduction and can promote overreduction to the diol as described in the aforementioned paper. [307] Overreduction and the formation of different side products has been found during our experiments as well.

The reduction of the TBDMS derivative yielded two fractions after column chromatography. Both fractions were analysed *via* NMR. The peaks from fraction A fit to the corresponding lactol and the fully reduced diol in a ratio of 1/0.9, according to the ^{1}H NMR integrals. The peaks from fraction B do not contain the same lactol as in fraction A, but two other lactols, possibly due to two different conformers. Two different lactol diastereomers (pro (*R*)/pro (*S*)-reduction) with inverted configuration (*syn*/*anti*) of the substituents at C-1 and C-2, are expected to be formed during the reduction. Fraction B also still contained some diol. The three different products in fraction B (lactol with duplet of duplet hemiacetal peak, lactol with singlet hemiacetal peak, diol) in a ratio of 1/0.7/0.2, according to the ^{1}H NMR peak integrals. Nonetheless, clearly it was possible to reduce part of the lactone to the lactol, however not in high yield (< 30%) and accompanied by overreduction.

The MOM-protected lactone was reduced to the corresponding lactol, yielding a single spot on the TLC plate. The NMR analysis indicated two lactol isomers, either corresponding to the two diastereomers, or to two conformers and one unidentified side product. In hindsight, it seems likely that the second diastereomer was not eluted from the column. The yield (~35%) is better than that from the TBDMS derivative.

The acetyl-protected lactol was obtained from the reduction as well, however isolation proved difficult, because of strong interactions with the silica gel. Additionally, the product fraction contained a lot of the diol (~1/1). The yield was estimated to be 30% as well.

All ^{1}H NMRs also contained a number of aldehyde peaks (9.6 – 10.3 ppm), corresponding to the number of identified lactols. Other possible side products, like reduction of the protected alcohol to the corresponding alkane, elimination of the protected 2-hydroxide group, and deprotection were not identified in the NMR spectra. [313]

Concluding, the reduction of all three protected lactones with the lithium triethylborohydride reagent is possible, surpassing the expectations. The MOM-protected lactone seems to deliver the best results in terms of yield, prevention of overreduction and enantioselectivity. The reactions still hold a lot of potential for optimization and analysis.

The protocol, including a mild quench with a mixture of methanol and water should be reviewed considering quenching with water only. This might prevent side product formation. Due to parallelization of the three test reactions, the two-phasic product was stored overnight in the fridge. Instead, the phases should be quickly separated and the organic phase dried, to prevent open aldehyde formation during storage and isolation.

A reduction with lithium triethylborohydride, dissolved in DCM instead of THF should be considered. Also, comparing lithium triethylborohydride with commercially available lithium and potassium tris(*sec*-butyl)borohydride will deliver further information about the influence of the Lewis acidity of the counterion and the influence of the steric bulk of the borohydride. If assumptions are correct, lithium tris(*sec*-butyl)borohydride should deliver a similar result with increased enantioselectivity, while the potassium tris(*sec*-butyl)borohydride should lead to increased overreduction and decreased anomer selectivity, compared to the lithium tris(*sec*-butyl)borohydride. The addition of different Lewis-acids, like Cs^{+}, Mg^{2+} and Al^{3+} should also be considered. Produced conformers and diastereomers should be identified.

In the meantime, kinetic resolution of 2-hydroxy-γ-butyrolactone proved to be feasible with CalB and vinylchloride (*vide infra*). It was shown that reduction of the acetyl-protected lactone is feasible and since the focus of this project is to produce ***R*-11**

(2(*R*),4-dihydroxybutanal) as acceptor substrate for the transketolase reaction, optimization of the reduction of the other protected lactol derivatives was not continued. Instead, the focus was on production of the 2(*R*)-acetoxy-γ-butyrolatone and optimization of the reduction.

Main issues during the production of the acetylated lactol were identified to be purification and overreduction. Aside from that, due to our experience during the production of (***rac*-11**) *rac*-2,4-dihydroxybutanal *via* OsO_4 catalysed dihydroxylation, the long-term stability of these aldehydes is considered to be low. Thus, an *in situ* hemiacetal acetylation strategy was developed, tackling all three challenges (overreduction, purification, aldehyde stability) at once.

It has previously been shown that acetylation with acetic anhydride is compatible with DIBAL-H reduction. [314,315] In the protocol, the authors enantioselectively reduced **26** (ethyl *O*-benzylglycolate) to the corresponding acetals (**27**) with $NaAlH_4$ and a chiral diol (not shown) with *in situ* acetylation of the developing tetrahedral intermediate with acetic anhydride and DMAP (Figure 86).

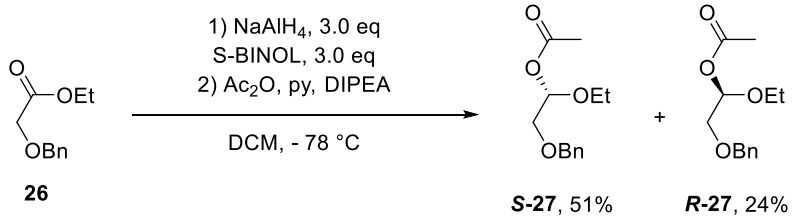


Figure 86: Reduction of benzyl (Bn) protected ethyl glycolate (**26**) with sodium aluminium hydride, and (*S*)-1,1'-bi-2-naphthol (*S*-BINOL) and *in situ* protection of the hemiacetal with acetic anhydride (Ac_2O), pyridine (py) and 4-(*N,N*-dimethylamino)pyridine (DMAP).

DMAP is a very toxic substance, with an LD_{50}, oral apprx. 6 times lower than pyridine and a high potential for cutaneous absorption. However, it is known that DMAP increases the rate of acetylation by a factor 10^4, which exceeds the performance of pyridine due to its increased nucleophilicity. Both reagents are known to activate acetic anhydride through the formation of a pyridinium salt, and fast acetylation of the tetrahedral intermediate, formed during the reduction, is expected to be beneficial to the reaction. For this reason, acetic anhydride, pyridine, and DMAP were added to the cooled reaction solution in close succession after completion of the reduction. Excess borohydride reagent was expected to react either with the acetic anhydride, or with one of the pyridines, forming tetrahydropyridine. Characteristic yellow-green colouring of

the reaction mixture, reported for the reduction of pyridine, has been observed upon the addition of DMAP. [308]

In the meantime, production of the 2(*R*)-acetoxy-γ-butyrolactone has been successful (*vide infra*) and was employed as substrate for the improved reduction protocol (Figure 87).

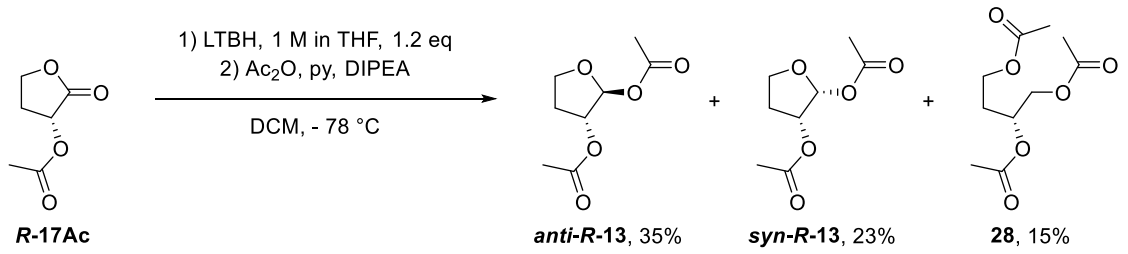


Figure 87: Reduction of ***R*-17Ac** (2(*R*)-acetoxy-γ-butyrolactone) with lithium triethylborohydride (LTBH) and *in situ* protection of the formed hemiacetal with acetic anhydride (Ac_2O), pyridine (py) and 4-(*N,N*-dimethylamino)pyridine (DMAP).

The desired product ***R*-13** (*syn*, *anti*-2,3(*R*)-diacetoxytetrahydrofurane) was isolated in 58% yield with high purity, drastically improving the purification and overall yield, relative to the reduction without *in situ* protection. Overreduction to the triacetylated triol (**28**) was minimized and is expected to be further decreased, when lithium triethylborohydride is employed as a solution in DCM, instead of the reagent in THF.

The two diastereomers were isolated in a ratio of 3/2 (*anti*/*syn*), exhibiting the potential of the bulky borohydrides to be stereoselective. Since desired deprotection leads to fast isomerization, the two diastereomers were not fully separated, even though this is possible *via* column chromatography.

4.3.2.4 Routes (**IIIb**): Kinetic Resolution of a Lactone Precursor with CalB

To produce the enantioenriched substrate ***R*-17Ac** (2(*R*)-acetoxy-γ-butyrolactone) for the reduction, two different methods were considered. One of them starts from the commercially available ***rac*-16** (2-hydroxy-γ-butyrolactone), separating the two enantiomers *via* kinetic resolution with lipase. The other route begins with an enzymatic cascade, converting ***R*-14** (L-homoserine) to ***R*-15** (2(*R*),4-dihydroxybutyric acid), and subsequently cyclising the linear hydroxy acid into the corresponding lactone.

The first tested enzyme for the kinetic resolution was the in-house available immobilized *Candida antarctica* lipase B (CalB). The immobilized enzyme is an industrial power-house with diverse applications in kinetic resolutions. The very stable enzyme preparation can be employed in organic solvent, enabling the conversion of

acids into esters and amides. The standard in-house procedure employs the immobilized enzyme granules and molecular sieves in dry cyclohexane with a flying stirring bar, to avoid milling of the granules. The chiral hydroxide is then added together with vinyl acetate. Transesterification, induced by the lipase, leading to transfer of the preferred enantiomer of the alcohol to the acetate, and the liberation acetaldehyde. Thus, vinyl acetate is an irreversible substrate, driving the reaction.

Prior to conducting the kinetic resolution, the preferred stereochemistry was predicted. Dividing the two substituents of a secondary alcohol into a large and a medium-sized group, according to Kazlauskas' rule, the (*R*)-ester is formed preferably, assuming the large group has the higher CIP-priority (Figure 88). [316]

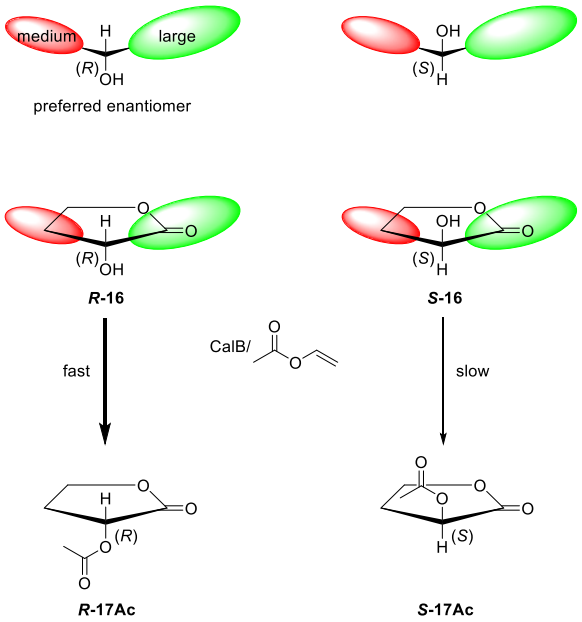


Figure 88: Kazlauskas' rule, applied to the transesterification of vinyl acetate with **16** (2-hydroxy-γ-butyrolactone) and *Candida antarctica* lipase B.

In other words, the enzyme binds the (*R*)-alcohol stronger than the corresponding (*S*)-enantiomer. In the case of **16** (2-hydroxy-γ-butyrolactone), the ideal ***R*-17Ac** (2(*R*)-acetoxy-γ-butyrolactone) was predicted to be formed preferentially during the reaction. This outcome would alleviate the need for additional alcohol protection of the ***R*-16** (2(*R*)-hydroxy-γ-butyrolactone) prior to the subsequent reduction, saving one reaction step.

However, due to a small size-difference of the two groups in the cyclic ester (Figure 88), it was not clear if the *E*-factor (enantioselectivity) would be sufficient to isolate ***R*-17Ac** (2(*R*)-acetoxy-γ-butyrolactone) with good yield and high enantiomeric excess

(*ee*) and if there would be a need to repeat the reaction with intermediate separation steps to increase the product *ee*.

Additionally, the alcohol substrate is an ester itself, bearing the potential for side reactions even without vinyl acetate (*Figure 89*).

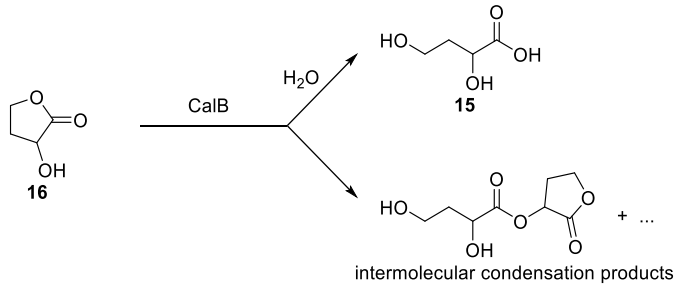


Figure 89: Possible side reactions in the reaction of **16** (2-hydroxy-γ-butyrolactone) with *Candida antarctica* lipase B.

A first test reaction conducted in cyclohexane illuminated the stereochemical outcome. The product ester was separated from the alcohol *via* column chromatography and analysed *via* chiral GC/MS. The chromatogram was compared with two authentic reference materials, produced from a commercial reference material of ***R*-16** (2(*R*)-hydroxy-γ-butyrolactone) and ***rac*-16** (*rac*-2-hydroxy-γ-butyrolactone *via* chemical acetylation. As predicted the preferred enantiomer was ***R*-17Ac** (2(*R*)-acetoxy-γ-butyrolactone).

The CalB test reaction was not yet optimized and two issues were found. The hydroxy lactone (***rac*-16**) substrate was not well soluble in cyclohexane and a considerable amount of a more polar side product was formed, identified to be the open chain **15** 2,4-dihydroxybutyric acid. Apparently, addition of molecular sieves to pre-dried cyclohexane and employing a drying tube is insufficient to fully exclude water from the reaction mixture in the overnight reaction, since saponification is very fast in wild-type CalB even with very low concentrations of water. [317] Luckily, large amounts of other transesterification-products were not found, presumably due to the large driving force with vinyl acetate.

To increase the solubility of **16** (2-hydroxy-γ-butyrolactone), subsequent reactions were conducted in a cyclohexane/THF (3 + 2, V + V) solvent mixture.

The enantioselectivity of an enzymatic reaction is usually expressed as the *E*-factor, representing a measure of how many molecules of the preferred enantiomer of a

racemate are converted until one molecule of the other enantiomer is converted. It depends on the solvent, the reaction temperature, and the substrate. The *E*-Factor of CalB against **16** (2-hydroxy-γ-butyrolactone) in THF, at 50 °C was determined from a kinetic resolution experiment, terminated after 16.5 h.

To improve the reproducibility of the method, an internal standard (2-butanol) was initially added to each sample. Alcohols in the analytical samples were first derivatized with *tert*-butyldimethylsilyl chloride to the corresponding silyl ethers to lower the boiling point for the subsequent gas-chromatography (GC) analysis.

GC/MS analysis yielded the concentration of **16** (2-hydroxy-γ-butyrolactone) to be 57.1 mM, and the conversion was calculated to be 40.5% from the starting concentration of the kinetic resolution (100 mM). Another sample of the reaction mixture was analysed by GC/MS without derivatization, determining the ratio of the two enantiomers 2(*R*)-acetoxy-γ-butyrolactone and 2(*S*)-acetoxy-γ-butyrolactone to be 90.37 and 9.63%, respectively. Thus, the *ee* of 2(*R*)-acetoxy-γ-butyrolactone product was 80.74%.

From calculated conversion and *ee*, the *E*-factor was calculated to be 16, after:

$$E = \frac{\ln\left[1 - c \cdot \left(1 + ee\left((R) - product\right)\right)\right]}{\ln\left[1 - c \cdot \left(1 + ee\left((R) - product\right)\right)\right]} = \frac{\ln\left[1 - 0.405 \cdot (1 + 0.8074)\right]}{\ln\left[1 - 0.405 \cdot (1 - 0.8074)\right]} = 16$$

where E is the *E*-factor, *ee* the enantiomeric excess, and c the conversion.

The *E*-factor is mediocre. To reach a perfect product *ee*, *E*-factors should usually be above 100. However, it is still considered to be suitable for the purpose of producing the acceptor substrate for the transketolase for the following reasons.

Transketolase reactions represent another kinetic resolution step with very high enantioselectivity, generally not requiring an enantiopure substrate to deliver the desired sugar. The main reason to employ the acceptor as an enantioenriched substrate is not to allow the production of the desired sugar, but to halve the substrate concentration. This is considered to be beneficial to the reaction and purification, enabling higher oxygen solubility, higher substrate concentrations and better performance of the subsequent column chromatography. Also, the kinetic resolution leads to a better atom economy in the subsequent reduction step. Therefore, a ratio of

10/1 in the acceptor enantiomers is considered to be sufficient to reach the desired improvements.

The recorded ^{1}H NMR spectrum of **17Ac** (Figure 90) has insufficient resolution, so that most H-H-coupling constants could not be identified.

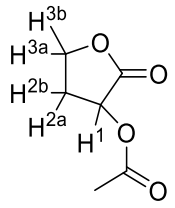


Figure 90: 2-acetoxy-γ-butyrolactone (**17Ac**) with numbered hydrogen atoms for ^{1}H NMR peak identification.

Additional coupling constants were identified in another well-resolved spectrum (Figure 91) of a mixed fraction from previous reactions:

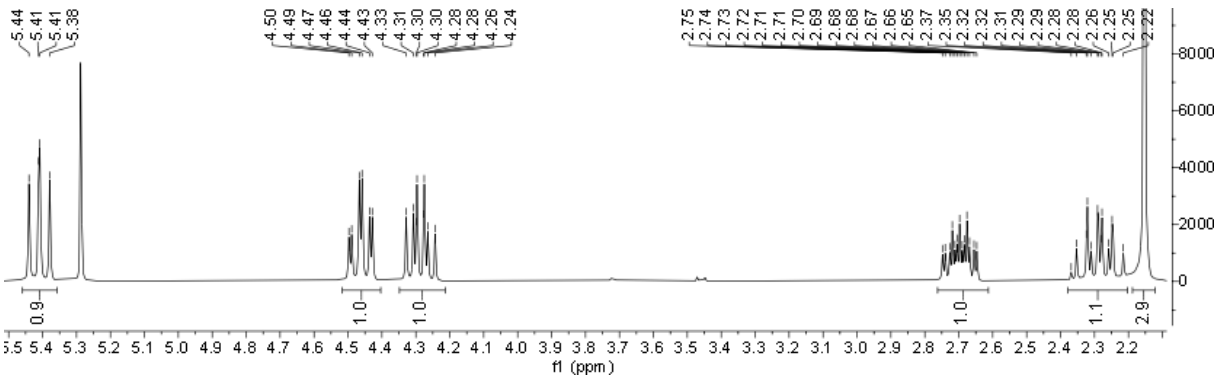


Figure 91: ^{1}H NMR spectrum (5.5 - 2.1 ppm) from another synthesis (April 2022) with better resolution for additional identification of coupling constants.

Additionally gained information is summarized in Table 11.

Table 11: Additionally identified coupling constants of hydrogen atoms in the ^{1}H NMR spectrum of **17Ac** (2-acetoxy-γ-butyrolactone).

Atom (Figure 90)	^{1}H NMR peaks (500 MHz, $CDCl_2$), given spectrum [ppm]	^{1}H NMR peaks (300 MHz, $CDCl_3$), additional spectrum [ppm]
H-1	5.41 (t, $^{3}J_H$ = 9.1 Hz)	5.41 (dd, $^{3}J_{2b}$ = 9.7, $^{3}J_{2a}$ = 8.6 Hz)
H-3a	4.45 (t, $^{3}J_H$ = 9.1 Hz)	5.47 (~td, $^{2}J_H$ = 9.1, $^{3}J_H$ = 2.5 Hz)
H-2a	2.74 – 2.64 (m)	2.70 (dddd, $^{2}J_H$ = 12.9, $^{3}J_H$ = 8.7, $^{3}J_H$ = 6.4, $^{3}J_H$ = 2.4 Hz)
H-2b	2.34- 2.21 (m)	2.27 (ddd, $^{2}J_H$ = 12.9, $^{3}J_H$ = 9.8, $^{3}J_H$ = 0.9 Hu)

4.3.2.5 Routes **(IV)**: Production of a Lactone Precursor (***R*-15**) from L-Homoserine

Another method, considered to yield perfect enantioselectivity and high yields, instead of the maximum possible 50% in the previously described kinetic resolution, is the enantiospecific reduction of a corresponding α-keto acid with D-hydroxy acid dehydrogenase and subsequent cyclisation. The needed keto acid can be produced from the corresponding L-amino acid, employing amino acid racemase and DAAO, in combination with catalase (Figure 93).

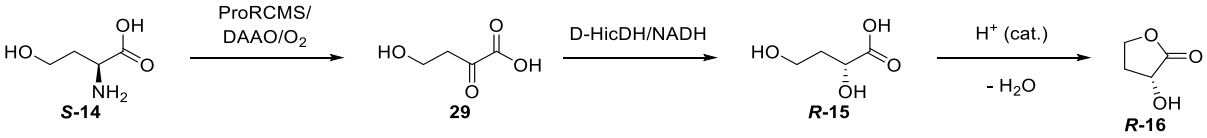


Figure 92: Reaction sequence for the production of ***R*-16** (2(*R*)-hydroxy-γ-butyrolactone) from ***S*-14** (L-homoserine), employing Prozomix amino acid racemase (ProRCMS), DAAO, and D-hydroxyisocaproic acid dehydrogenase (D-HicDH).

The D-2-hydroxybutyric acid was successfully produced in one-pot, as described in Chapter 4.5 about the production of hydroxy acids from L-amino acids (Figure 93).

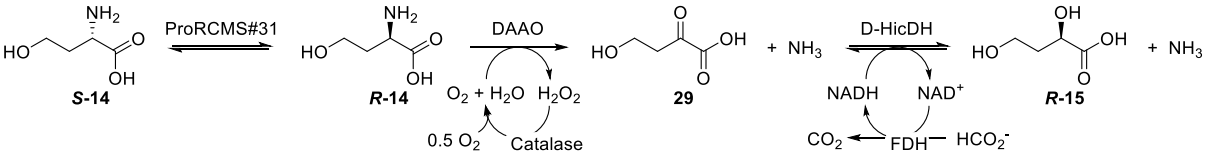


Figure 93: Biocatalytic one-pot cascade reaction for the production of ***R*-15** (D-α-hydroxybutyric acid) from ***S*-14** (L-homoserine), employing a panel racemase (ProRCMS#31), the DAAO from *Rhodotorula gracilis*, commercial catalase, D-hydroxyisocaproic acid dehydrogenase from *Lactobacillus paracasei* (D-HicDH) and commercial formate dehydrogenase (FDH) for cosubstrate recycling.

ProRCMS#31 has been found to be suitable in a comparative test reaction, conducted in the same way as the panel screening was conducted (*vide supra*). In this test reaction, ProRCMS#9 and #31 have been applied. Both racemases are suitable for the purpose and ProRCMS#31 has been found to have the higher activity on ***S*-14** (L-homoserine) under the screening conditions (*vide supra*), and therefore was chosen for further testing, also due to the high stability.

In the next step, reaction conditions for the cascade reaction from ***S*-14** (L-homoserine) to ***R*-15** (D-2-hydroxybutyric acid) were screened at small scale (0.5 mL), at 30 °C, 600 rpm in Eppendorf vials with punctured caps in an overnight (14 h) reaction.

The varied enzyme and substrate concentrations are displayed in Table 12.

Table 12: Parameter variation during the screening for suitable conditions of the biocatalytic one-pot cascade from L-homoserine (***S*-14**, L-hser) to D-2-hydroxybutyric acid (***R*-15**).

	L-hser [mM]	Formate [eq.]	NAD^+ [mol% eq.]	RCMS#31 [mg/mL]	DAAO [mg/mL]	D-HicDH [mg/mL]	FDH [U/mL]
1	120	1.7	4	5	2.5	0.5	0.6
2	120	1.7	4	5	2.5	0.5	0.3
3	120	1.7	4	2	5	0.5	0.6
4	120	1.7	4	2	5	0.5	0.3
5	100	2	4	4	4	4	0.6
6	100	2	4	4	4	4	0.3
7	100	2	4	4	4	2	0.6
8	100	2	4	4	4	2	0.3
9	100	2	2	2	2	2	0.6
10	100	2	2	2	2	2	0.3
11	100	2	4	2	2	2	0.6
12	100	2	4	2	2	2	0.3
13	100	2	8	2	2	2	0.6
14	100	2	8	2	2	2	0.3
15	50	2	8	0	2	2	0.6
16	50	2.5	16	0	2	2	0.3

The results were analysed *via* HPLC/MS with a hydrophilic interaction liquid chromatography (HILIC) method.

HILIC chromatography is a powerful tool for the separation of highly polar compounds, employing polar solvents with the ability to dissolve polar compounds, but separating in quasi-normal mode.

A simplified model of the underlying principle is described hereafter: A water miscible solvent (*e.g.* ^{i}propanol, acetonitrile) is used, containing a low fraction of water. During chromatography, the surface of a relatively polar or ionic column (*e.g.* Waters BEH Amide) is wetted with an aqueous layer, forming the stationary phase, while the other solvent functions as the mobile phase. Polar compounds are partitioned between both

liquid phases, depending on their hydrophilicity. Saturating the aqueous fraction of the eluent with buffer enables control of pH and even stronger film formation. It also leads to stronger repulsion of less hydrophilic compounds from the stationary aqueous phase. In our case, a gradient was employed, slowly increasing the water content of the eluent from 13 to about 50% during the chromatography. The limits of the method were tested, successfully separating two different challenging mixtures of polar compounds. The first mixture contained L-serine and lithium hydroxypyruvate and the second mixture contained L-isoleucine and L-leucine.

Due to the limits of the HPLC's MS detector, the average fault in the measured concentration was estimated to be relatively high (≥ 10%). L-Alanine was added to the samples as an internal standard. Also, no individual calibrations were prepared. For these reasons, the results were normed on the sum of ***S*-14** (L-homoserine), **29** (4-hydroxy-2-ketobutyric acid), and **15** (2,4-dihydroxybutyric acid) peak integrals as 100%. Side reactions were not considered. However, since the goal was to reach full conversion and the employed single mass search of the MS is very sensitive in detecting the searched ion masses (***S*-14**: M^+ = 120 u/z, **29**: M^- = 117 u/z, ***R*-15**: M^- = 119 u/z), the method is considered to be fit for the purpose of finding suitable conditions, looking for the absence of peaks for ***S*-14**-, and **29** and a large integral of the ***R*-15**-peak.

The results are displayed in Figure 94 as fractions of the corresponding HPLC/MS peak integrals relative to the sum of these integrals.

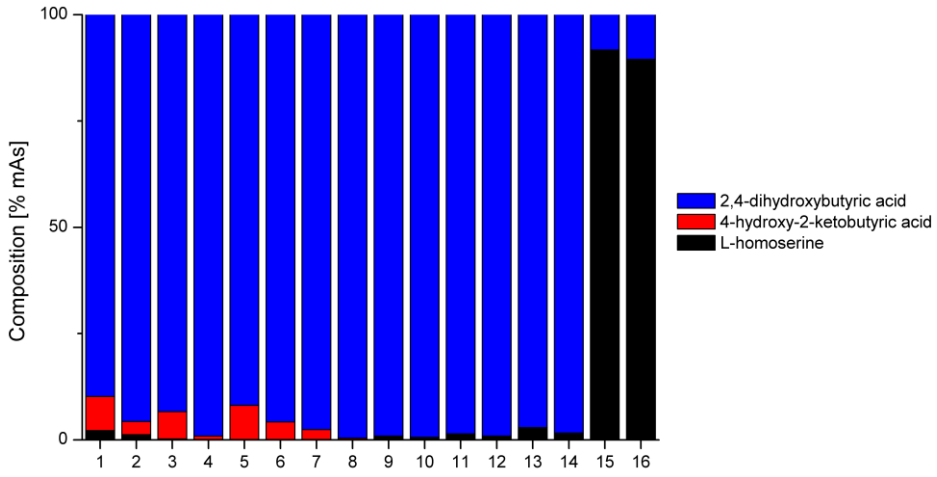


Figure 94: Results of a screening for conditions in the L-homoserine (***S*-14**) to 2,4-dihydroxybutyric acid (***R*-15**) one-pot cascade reaction. 16 reactions (500 µL, 25 °C, 600 rpm, aeration) were run under varied conditions, as indicated in *Table 12*.

Most conditions were suited to reach high conversion, which points out that there is still a lot of room for improvements. De-colouration of the yellow reaction mixtures (FAD -> $FADH_2$) in deeper layers of the reaction mixture was observed, which is indicating insufficient aeration and highlighting the value of micro-scale aeration as described above. The small conversions in the ProRCMS-blanks (reaction 15 and 16) can be interpreted either as a limited enantiopurity of the employed ***S*-14** (L-homoserine) reagent, or as an ability of DAAO from *Rhodotorula gracilis* to convert ***S*-14** (L-homoserine) with very slow turnover.

The best results were achieved in samples 4, 8, and 10. The screening could have been repeated with shorter reaction times, a new set of conditions, and better aeration for further optimization. Since almost all reactions indicated high conversion, rather another single small-scale test reaction was run with increased substrate concentration (250 mM, 2.0 eq. formate, 1 mg/mL D-HicDH, 2 mg/mL ProRCMS#31, 0.12 U/mL FDH) and better aeration, to assess preparative conditions. The new conditions were chosen on the basis of the following results in the initial screening:

- 250 mM amino acid substrate (***S*-14**), instead of 100 mM: no effect observed with different substrate concentrations, goal is to demonstrate the capabilities of the cascade system to reach high product-titers and make oxygen provision the limiting factor.
- 2 eq. ammonium formate (500 mM): no negative effect, depending on the total concentration of salts detected; cheap buffering substance.
- Lower NADH concentration (1.5 mol% eq.): no limitation observed
- Lower FDH concentration: no concentration dependent effect observed.
- D-HicDH concentration between 0.5 and 2 mg/mL (1 mg/mL): 0.5 mg/mL seemed to be almost sufficient to convert all 2-keto acid (**29**), **29** was not detected with 2 mg/mL, concentrations above 10 mg/mL cell-free extract, or of more than 4 mg/mL D-HicDH cell-free extract seem to inhibit the hydrogenation.
- DAAO and ProRCMS#31 were employed in 2 mg/mL concentration, as in reaction 10.

After 24 h, small quantities of ***S*-14** were still detectable on TLC with sensitive ninhydrin dyeing. Since DAAO is prone to deactivation, another portion (2 mg/mL) of the enzyme was added and the mixture incubated for another 24 h, leading to full conversion.

Part of the product (***R*-15**) was further purified *via* column chromatography to around 80% purity, still containing some formic acid from column chromatography, but demonstrating the capabilities of the cascade system with a volumetric productivity of 30.5 g/L and quantitative yield from > 24 to < 48 h reaction times.

Two methods to cyclise the hydroxy acid ***R*-15** to the corresponding enantiopure ***R*-16** (2(*R*)-hydroxy-γ-butyrolactone), both based on an acid catalysed intramolecular ester condensation with removal of water were tested. [318,319]

In the first experiment, a small portion (~ 60 mg) of the crude product ***R*-15** (~ 65% purity) was acidified in a mixture of acetic and formic acid with molecular sieves and refluxed, yielding around 50% of the corresponding lactone ***R*-16** after 18 h.

A second experiment, intended to further drive the reaction equilibrium was conducted in a microdistillation apparatus, replacing the acetic acid solvent with non-volatile sulfuric acid. However, heating a mixture of neat impure product with concentrated sulfuric acid led to strong foaming, indicating decarboxylation. Nonetheless, a small product ***R*-16** quantity (< 2%) could be isolated from the condenser and was analysed *via* NMR.

Since the kinetic resolution that was conducted in parallel proved to deliver the desired acceptor, the cascade development experiments were discontinued in favour of the simple resolution protocol.

It should be clear that the conversion of ***S*-14** (L-homoserine) to ***R*-15** (2(*R*),4-dihydroxybutyric acid) is highly effective as a one-pot cascade reaction system with good productivity and easy purification, which delivers a highly enantiopure product based on the selectivity of D-HicDH. The atom efficiency is very high compared to the kinetic resolution.

The cyclisation is expected to deliver high yield upon further optimization. The lactonization under reduced pressure could be repeated in a high boiling solvent, successively removing water and the product lactone ***R*-16** from more dilute acid, maybe employing trichloroacetic acid instead of sulfuric acid. Alternatively, cyclisation may be conducted in refluxing benzene with *p*-toluenesulfonic acid catalysis, as reported by Caro et al.. [319]

In principle, these results indicate the cascade process to be more efficient than the kinetic resolution. However, ***S*-14** (L-homoserine) is about 5 times more expensive than ***rac*-16** (*rac*-2-hydroxy-γ-butyrolactone), when comparing commercially available 5 g portions from Sigma-Aldrich (Merck), a price difference still large enough to justify the kinetic resolution process.

However, biocatalytic cascades for the production of ***S*-14** (L-homoserine) from cheap starting materials become increasingly available and an increasing demand for bio-based building blocks is expected, thus favouring the L-homoserine-based process in the foreseeable future. [320,321,322]

4.3.3 Donor Substrates for Transketolases

It should be clear by now that transketolases convert diverse, so-called acceptor aldehydes to the corresponding ketose products, especially when exploiting the inherent driving force of the hydroxypyruvate donor, as discussed above.

However, not only is it now possible to employ a vast number of different aldehyde acceptors, but there are now also engineered enzyme variants available for the conversion of pyruvate as acyl-donor. Thus, the enzyme becomes a transacyl-, or 2-oxoalkylase in the case of even higher homologues. An excellent example for this purpose is the double variant H102L/H474S of the thermostable transketolase from *Geobacillus stearothermophilus* for pyruvate conversion, engineered by Saravanan *et al.*, [121] using an *in vitro* evolution approach. The same publication also presents enzyme variants for 2-oxobutyrate and 2-oxo-valerate conversion.

As donors for our racemisation experiments, hydroxypyruvate and pyruvate were selected, generated from the corresponding amino acids. As already outlined, hydroxypyruvate is the classic donor substrate in synthetic transketolase reactions, because it is well accepted by the wild-type enzymes and because the irreversible formation of CO_2 drives the transketolase equilibrium to the product side, enabling even less reactive aldehyde acceptor substrates to be converted.

For preparative biotransformations, we selected the wild-type transketolase from *Geobacillus stearothermophilus* for hydroxypyruvate conversion, due to its high operational stability and the simple purification procedure (*vide infra*). Also, it has been proven to convert the selected acceptor 2,4-dihydroxybutanal, [299] and was available in-house.

The same driving force is also intrinsic to pyruvate as transketolase donor, but it is not a suitable substrate for wild-type transketolases. Thus, for pyruvate conversion, the in-house-available variant TK(H102L/H474S) was selected. [121]

For the *in situ* generation of hydroxypyruvate from serine, the two enzymatic routes (transamination and oxidation) were compared and discussed for example by L'Enfant *et al.*. [233] For transamination, we selected the Prozomix transaminase ProTA#17, because out of the transaminases available in-house it had one of the highest activities on serine at the selected temperature (25 °C). The enzyme has been recently characterized to be (*R*)-selective, which is unusual, since most known transaminases are (*S*)-selective. For oxidative deamination we selected the well-known DAAO from *Rhodotorula gracilis* which excels in terms of catalytic activity and stability, when compared to other common DAAOs. [234]

Discussing the price of bulk chemicals on the basis of small-scale purchases of research chemicals is not easy. In general, common amino acids like alanine and serine are relatively cheap in all isomer forms. For example, BLDpharm sells L-, D-, and D/L-forms of serine for around 70 € per kg (63 (95%), 71 (97%), 67 € (98%), respectively) [323,324,325] However, one should keep in mind that racemic amino acids are usually the cheapest, produced *via* chemical synthesis and sold at far lower prices in bulk than available from the common small-scale chemical vendors. Unfortunately, it is difficult to acquire accurate prices for these products from the manufacturers without a big company background and all inquiries from my side remained unanswered.

Concluding, even though simple amino acids are in general not a limiting factor for the costs of the presented processes, the application of amino acid racemases from the ProRCMS panel enabled us to always use the least expensive commercially available form.

For the generation of pyruvate, a newly discovered DAAO became available (ProDAAO#4536), [326] which is the first example of a DAAO with a high thermostability, giving us the opportunity to execute a cascade reaction at elevated temperature and test the thermostable ProRCMS#31. Unfortunately, this thermostable DAAO was found to be not active on D-serine (data not shown).

For alanine prices, the same trends as for serine can be observed.

Even though racemic alanine and serine are the least expensive, the L-enantiomers were selected as starting materials, because this provides us a valuable negative control, requiring a mandatory racemization for conversion. The reaction setup is also fit for the application of the cheaper racemate as substrate, providing for higher initial reaction rates. A possible accumulation of the corresponding oxo-acids due to higher initial rates can be counteracted, for example through lowering of the initial oxidase load, or amino acid concentration or through increasing the transketolase concentration.

4.3.4 Production of 5-Deoxy and 1,5-Dideoxyfructose

The complete cascade reactions are displayed in Figure 95.

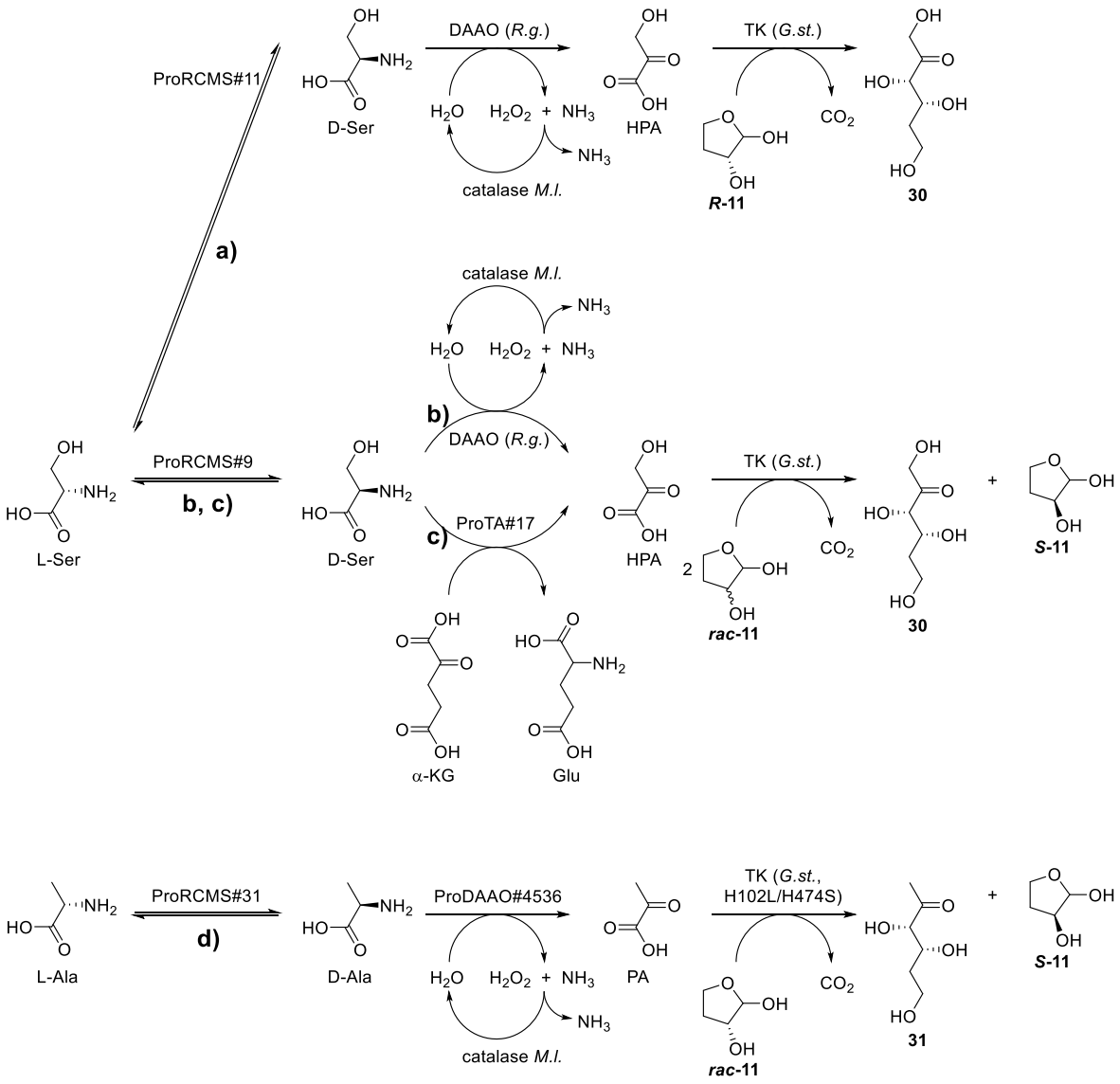


Figure 95: Reaction schemes of amino acid racemase (ProRCMS#)/transketolase (TK) cascade reactions: a) L-serine (L-Ser)/ProRCMS#31/DAAO/TK pathway with acceptor ***R*-11** (2R,4-dihydroxybutanal, depicted as hemiacetal) for the production of **30** (5-deoxyfructose) , b) L-serine (L-Ser)/ProRCMS#11/DAAO/TK pathway with acceptor ***rac*-11** (*rac*-2,4-dihydroxybutanal, depicted as hemiacetal) for the production of **30** (5-deoxyfructose), c) L-serine (L-Ser)/ProRCMS#11/transaminase (ProTA#17)/TK pathway with acceptor ***rac*-11** (*rac*-2,4-dihydroxybutanal, depicted as hemiacetal) for the production of **30** (5-deoxyfructose), d) L-alanine (L-Ala)/ProRCMS#31/DAAO/TK pathway with acceptor ***rac*-11** (*rac*-2,4-dihydroxybutanal, depicted as hemiacetal) for the production of **31** (1,5-dideoxyfructose). ProDAAO#31: Prozomix D-amino acid oxidase panel, enzyme #4536, ProRCMS#: Prozomix amino acid racemase panel, enzyme number given, *R.g.*: *Rhodotorula gracilis*, *G.st.*: *Geobacillus stearothermophilus*, *M.l.*: *Micrococcus lysodeicticus*, α-KG: α-ketoglutaric acid, Glu: glutamic acid, HPA: hydroxypyruvic acid, PA: pyruvic acid.

Reaction conditions were selected according to the following considerations, or due to experimental results, gathered as practical guidelines.

4.3.4.1 Selection of Reaction Conditions

The maximum applicable **temperature** is limited by the oxidase. As described before, the DAAO from *Rhodotorula gracilis* for serine oxidation represents a bottleneck with all other reaction components available as more thermoresistant variants. Since this yeast oxidase is quickly deactivated in solution at temperatures above 25 °C, [202] this temperature was selected for reaction a, b, and c. Compatible to that, ProTA#17 was one of the in-house available transaminases with highest activities at this relatively low temperature. Alternatively, in the D-amino oxidase ProDAAO#4536 this temperature limitation is lifted. The enzyme is considered thermostable, making it a useful tool for research applications. [326] In fact, it allowed us to test the thermostable ProRCMS#31 at a temperature of 45 °C, as a compromise between enzyme activity, oxygen solubility, and long-term stability of the system.

Several **buffers** were selected for testing, according to the considerations of Good *et al.* [327] and the individual requirements of the envisioned cascade reactions. Prerequisites for the buffer were a suitable buffer-range, ideally with excellent buffer capacity at the starting pH (7.5) of the cascade with greater capacity required against the basic range, to compensate for the formed hydroxide (transketolase pH shift, *vide supra*). Also, the buffer was sought to be inert against the applied catalysts and all other components of the reaction mixture. In particular, amino acids and other easily oxidized components were not used, since they might be oxidized by the oxidase or by the formed H_2O_2. Primary and secondary amino groups were avoided due to their ability to form imines with aldehydes and ketones involved. Additionally, the buffer should not have a negative effect on the stability of the oxidase, which was reported for example for HEPES. [328] Among the tested buffers were potassium phosphate, MOPS, triethanolamine (TEA), and HEPES.

To test the influence of the buffers on the stability of resting DAAO from *Rhodotorula gracilis*, the enzyme (10 mg/mL, cell-free extract) was incubated with the selected buffers (potassium phosphate, HEPES, MOPS, TEA) at 100 mM, pH 7.0 with 0.1 mM FAD for 20 h at 25 °C in closed Eppendorf vials with perforated caps. These conditions are similar to those in the cascade reactions, however void of substrates and other enzymes.

The activity of aliquots of the incubated enzymes was measured with D-serine and the colorimetric horseradish peroxidase assay, described in the screening of the ProRCMS panel (Figure 96).

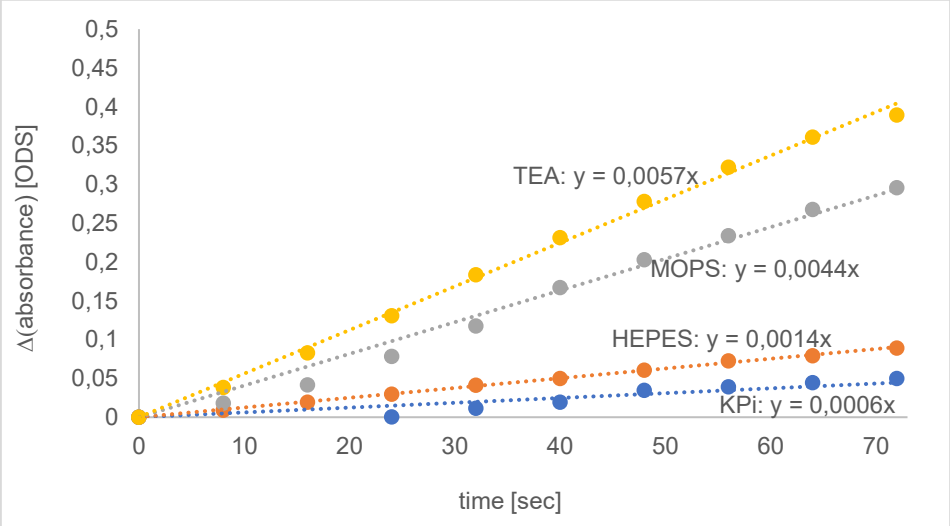


Figure 96: Results of the spectrophotometric measurement, displaying the initial increase in absorbance at 498 nm of reaction wells of a microtiter plate, each containing a solution (100 µL) of D-serine (50 mM), 4-aminoantipyrine (2.2 mM), vanillic acid (2.2 mM), horseradish peroxidase (8.8 U), and the preincubated (10 mg/mL in 100 mM buffer, pH 7, 20 h, 25 °C, under air) oxidase (2 mg/mL in well) in the respective buffer. Blue: potassium phosphate (KPi), orange: 2-[4-(2-hydroxyethyl)piperazin-1-yl]ethane-1-sulfonic acid (HEPES), gray: 3-(morpholin-4-yl)propane-1-sulfonic acid (MOPS), yellow: triethanolamine (TEA).

It became clear that the stability of the oxidase preparation differed a lot depending on the buffer. In potassium phosphate buffer, there was almost no residual activity. The deactivation of the enzyme was even visible to the naked eye, where the preincubation vials exhibited different degrees of turbidity. The oxidase incubated in TEA was most active and therefore, TEA was selected as buffer for cascade reactions.

In a follow-up experiment, the conversions of complete cascade reactions in TEA and HEPES were determined in different buffers, using different racemases and the commercially available D/L-glyceraldehyde.

For analysis of the D/L-glyceraldehyde proportion, samples were derivatized with PMP (1-phenyl-3-methyl-5-pyrazolone), following a procedure by Zhang *et al.*.[329] The derivatization reaction leads to a bis-PMP adduct with reducing sugars, as depicted in Figure 97.

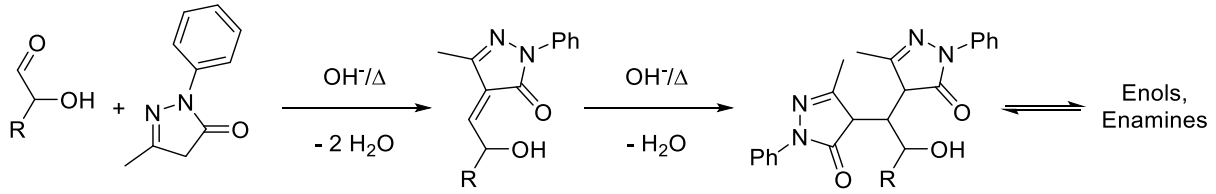


Figure 97: Reaction sequence, leading to the bis-PMP adduct, as described in Bailly *et al.* 2020. [330]. RCHOCHO: reducing sugar.

For detailed information about the applied conditions, see the experimental part. At times of 1, 2, 3 and 22 h after the start of the reaction, aliquots were derivatized with PMP and analysed *via* HPLC/UV. As a measure of the conversion, the ratio between the peaks of the two enantiomers of the glyceraldehyde acceptor substrate were used.

No significant differences were found during these reactions, comparing the proportions of L and D-glyceraldehyde during the conversion in TEA and HEPES. This is attributed to fast conversion of D-serine to HPA during the first 8 hours of the reaction (compare 4.3.4.4).

To check whether under the conditions of a cascade reaction aldehyde substrates are inhibiting anyone of the employed enzymes, the conversions of complete cascade reactions (Figure 95*)* were determined, with ProRCMS#11 and the commercially available D/L-glyceraldehyde as acceptor at different **concentrations** (80, 160, 240 mM). For detailed information about the applied conditions, see the experimental part.

After 22 h, in all three reactions around 30 – 40 mM of the D-glyceraldehyde enantiomer were converted, as calculated from the ratio of the bis-PMP derivatized L- and D-glyceraldehyde educt (Table 13).

Table 13: Calculated conversions of D-glyceraldehyde, based on the integral ratio of bis-PMP derivatized L- and D-glyceraldehyde HPLC (Method chiral reverse isocratic I, 254 nm) peaks. All three reactions (500 µL, 25 C, 22 h, under air) contained *rac*-glyceraldehyde concentration given), L-serine (1.0 eq relative to D-glyceraldehyde), ProRCMS#11 (10 mg/mL, cell-free extract), DAAO (10 mg/mL, *Rhodotorula gracilis*, cell-free extract), transketolase (20 mg/mL*, Geobacillus stearothermophilus*, heat shocked lyophilizate, PLP (0.5 mM), FAD (0.5 mM) in HEPES (100 mM, pH 7.0).

Initial *rac*-aldehyde concentration [mM]	Calculated conversion [mM]
80	31
160	33
240	38

The error in the calculation of conversions, based on the relative integrals of incompletely resolved enantiomer peaks of bis-PLP derivatized samples is estimated to be relatively high (~ 10%). Concluding, no inhibition of the cascade, in response to aldehyde concentrations of up to 240 mM, was found in the case of *rac*-glyceraldehyde.

The optimization of **enzyme proportions** was targeted to reach high conversion of L-serine without accumulation of hydroxypyruvate. Accumulation of unstable hydroxypyruvate was sought to be avoided, reducing its residence time and thus the level of decomposition. This way, the yield of product compared to serine is increased and the excess of serine in the reaction mixture is minimized. The second possibility to increase the serine to product yield was to minimize the reaction duration, also decreasing the hydroxypyruvate residence time. The reaction duration can be decreased by increasing the enzyme loading. However, at some point, the enzyme concentration cannot be further increased due to limited protein solubility. Additionally, at high concentrations of cell-free extract, dehydration of serine can increasingly occur due to *E. coli* enzymes. Also, high viscosity and concentrations are a trade-off with the oxygen concentration.

To optimize the proportions of oxidase and racemase, a straightforward approach was followed, and conversions of D-serine with DAAO were compared with conversions of L-serine with higher concentrations of ProRCMS#9 and four times the amount of DAAO. For fast and reliable determination of hydroxypyruvate, using purified enzymes at low concentrations, the WST-1 assay was employed.

In a first experiment, the conversion of 50 mM D-serine with 0.2 mg/mL DAAO from *Rhodotorula gracilis* after 1 h of incubation was compared with the conversion of 50 mM L-serine, four times the concentration of oxidase and varying amounts of ProRCMS#9 (Figure 98).

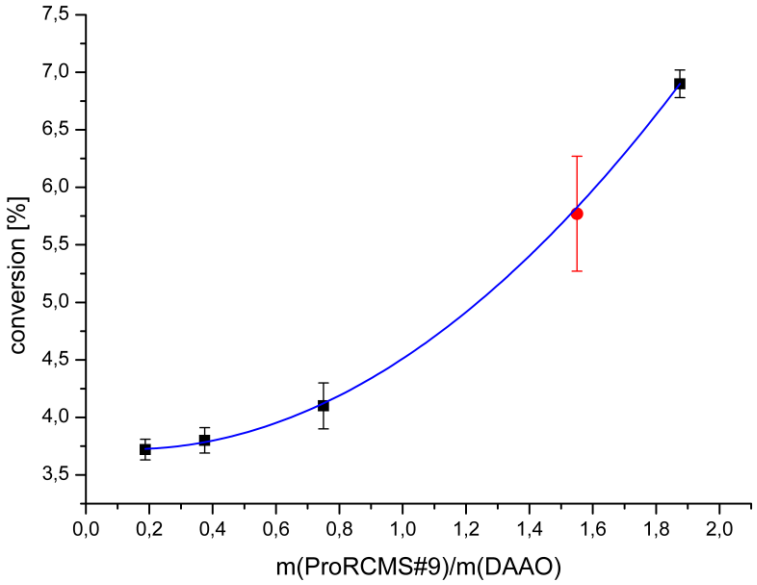


Figure 98: Conversion of 50 mM L-serine (■) with 0.8 mg/mL purified DAAO from *Rhodotorula gracilis* and varying amounts of purified ProRCMS#9 in potassium phosphate buffer (50 mM, pH 7.0) at 25 °C and 1 atm air. The results are interpolated with a 2nd degree polynomial fit (blue) and compared with the conversion of 50 mM D-serine (●) with 0.2 mg/mL DAAO under the same conditions.

The results suggest that approximately the same amount of hydroxypyruvate forms in one hour at 25 °C, when applying 0.8 mg DAAO and 1.25 mg/mL ProRCMS#9 on 50 mM L-serine, compared to when applying 0.2 mg/mL DAAO on 50 mM D-serine under the same conditions.

In a second experiment, the DAAO concentration was varied as well (Figure 99).

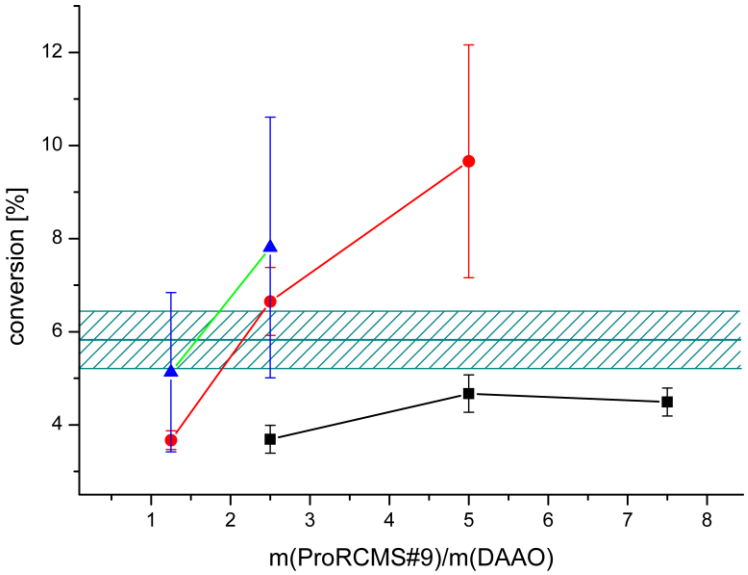


Figure 99: Conversion of 50 mM L-serine with 0.4 (■), 0.6 (●), and 0.8 mg/mL (▲) purified DAAO from *Rhodotorula gracilis* and varying amounts of purified ProRCMS#9 in potassium phosphate buffer (50 mM, pH 7.0) at 25 °C and 1 atm air. The results are compared with the conversion of 50 mM D-serine (50 mM) with 0.2 mg/mL DAAO under the same conditions, displayed as hatched area (cyan).

As shown in the previous experiment, the conversion of 50 mM D-serine with 0.2 mg/mL DAAO reaches apprx. 6%. As expected, when doubling the amount of DAAO, the conversion of L-serine cannot reach that level with any amount of added racemase. Considering the oxidase's Michaelis constant for D-serine (K_M = 13.75), [202] it would have to be accumulated to a concentration of above 13.75 mM within the first minutes of the reaction, assuming 50 mM is enough to reach v_{max}.

With three times the amount of oxidase, similar conversions were reached with m(ProRCMS#9)/m(DAAO) proportions between 1.8 and 2.5. With four times the amount of oxidase, it was reached with proportions between 1.25 and 1.8, or 1.58, assuming the error width to be zero, fitting well to the previous result of 1.55.

A similar optimization was also conducted with the immobilized DAAO preparation (data not shown).

4.3.4.2 Continuous Monitoring

Continuous monitoring of **acceptor aldehyde 11** (2,4-dihydroxybutanal) and the **products 30** (5-deoxyfructose) and **31** (1,5-dideoxyfructose) was based on ^{1}H NMR analysis (D_2O, 500 MHz, 256 scans, relaxation delay 1.0000 sec) using the reference signal of one of the two methylene protons at C3 of **11** (2,4-dihydroxybutanal, 2.22 ppm) and C5 of **30** and **31**(5-deoxy- and 1,5-dideoxyfructose, respectively, 1.64 ppm). The samples were prepared by lyophilization of an aliquot (200 µL) of the reaction mixture and dissolution of the lyophilized sample in D_2O (600 µL). The sample was spiked with an authentic reference material (maleic acid, 70.8 mM) as internal standard prior to the sample dissolution.

Figure 100 shows the time course of a cascade reaction of L-serine and ***rac*-11** (*rac*-2,4-dihydroxybutanal), displaying the corresponding NMR spectra in a time-resolved manner.

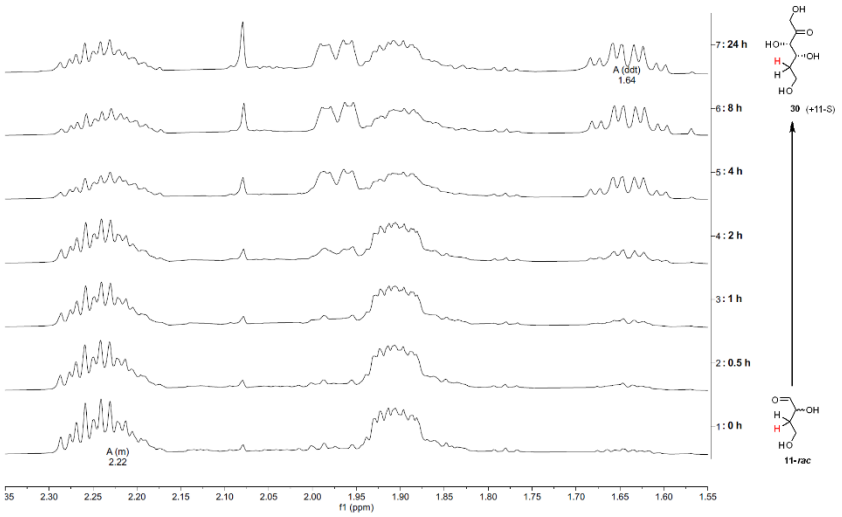


Figure 100: Stacked ^{1}H NMR spectra of reaction samples (0, 0.5, 1, 2, 4, 8, 24 h) of a cascade reaction of L-serine and ***rac*-11** (*rac*-2,4-dihydroxybutanal) to **30** (5-deoxy-D-fructose), zoomed in on the methylene peaks of educt acceptor and product hexose.

Baselines of the original ^{1}H NMR spectra were corrected with MestReNova's Bernstein polynomial fit (3rd order), or manually if the result of the polynomial fit were unsatisfactory (spectrum of cascade c, 480 min). Peaks were normalized on the peak integral of the internal standard (maleic acid 70 mM). Following the reaction over 24 h reveals the concentrations of educt acceptor **11** (2,4-dihydroxybutanal) and product hexose (**30**, **31**), according to the integrals of the peaks of one of their methylene protons (2.22 and 1.64 ppm, respectively), relative to the integral of the maleic acid peak (6,36 ppm). The deviation of the two selected peaks' (product and educt) integral sum is small during the time course of cascade a, b, c, and in the first 3 samples of cascade d (6, 5, 10, and 4% standard deviation, respectively), but large in the later samples of cascade d (15%). Also, converting the time course' mean of this integral sum into concentrations *via* the internal standard's integral is in good accordance with the selected starting concentration cascades a - c, and the first samples of d (135, 213, 214 mM, 226 and 103 mM). The small excess with respect in the calculated concentration is attributed to fluctuations in the NMR baseline and peak overlap with other minor reaction components. Therefore, the integral sum of individual spectra was standardized on the acceptor aldehyde's starting concentration. In cascade a (with 2R,4-dihydroxybutanal), the sum is ~ 1.2-times higher as compared to the starting concentration, delivering a possible explanation to the relatively large amount of unconverted educt.

This method, based on ^{1}H NMR analysis, was selected due to a lack of satisfying alternatives and because it enabled parallel specific analysis of both acceptor and product, with little interference from other reaction components.

The alternative PMP derivatization did not lead to reproducible results when attempting absolute quantification, which is mainly due for two reasons. The first reason is that in spite of extensive method development, full resolution of derivatized 2(*R*),4- and 2(S),4-dihydroxybutanal could not be achieved on our chiral HPLC columns. Secondly, the total intensity of the two peaks was found to be very sensitive towards the handling during the derivatization protocol (chloroform extraction and heating), which is difficult to exactly reproduce, and the timespan and storage temperature before measurement. Therefore, only relative peak integrals could be used to determine the ratio between the two enantiomers.

Derivatization with dinitrophenylhydrazine and benzylhydroxylamine was also attempted, but with unsatisfactory resolution using HPLC analysis.

Figure 101 shows the decreasing intensity of the ***R*-11** 2(*R*),4-dihydroxybutanal peak relative to the peak of the (S)-enantiomer during a cascade reaction between L-serine and ***rac*-11** (*rac*-2,4-dihydroxybutanal).

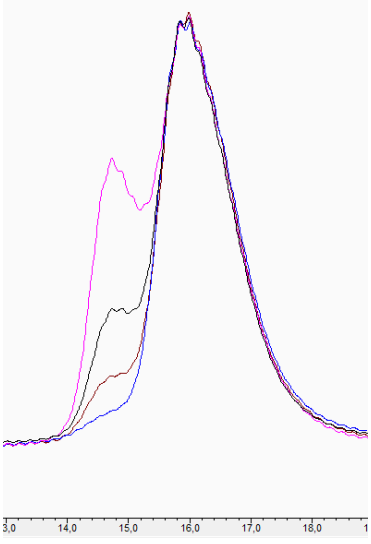


Figure 101: Overlay of the HPLC (method chiral reverse isocratic I, 270 nm) chromatograms of PMP-derivatized samples of a cascade reaction, taken at 120, 240, 480, 1440 min after the start of the reaction. With the bis-PMP (*R*)-derivative at 14.6 and the (S)-derivative at 16 min.

To monitor the consumption of **L-serine**, aliquots of the reaction mixture were derivatized with excess FmocCl (Figure 102) and analysed with chiral HPLC/UV.

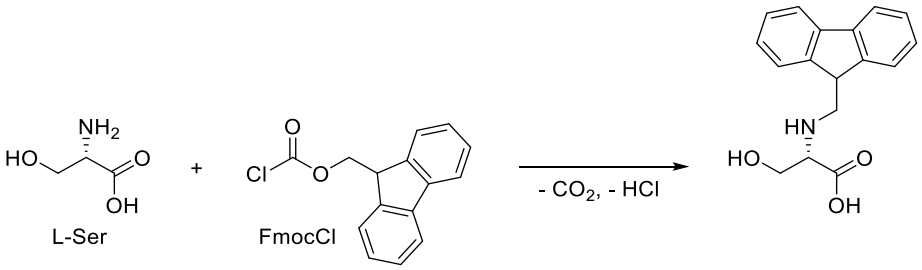


Figure 102: Fmoc derivatization of the L-serine amine with FmocCl.

Fmoc derivatization and HPLC/UV is a standard protocol for amino acid analysis, and all needed reactants and protocols were available in-house. For total serine, UV/Vis-detection at 254 nm was used. Conducting HPLC on a chiral column enables differentiation between D- and L-enantiomers of serine. Chiral separation was applied to the samples of reaction cascade b (formation of 5-deoxyfructose from ***rac*-11** and L-serine, employing DAAO) with mass spectrometric single ion search detection of enantiomers. Mass spectrometric quantification usually shows a high peak integral drift over time, but this method is quite stable for short measurement intervals (serine enantiomer peak distance = 1.0 min) and is a sensitive probe for the existence of an analyte (low limit of detection).

In each series of HPLC samples from the individual cascades, the sample with the longest reaction duration was used as blank. All blanks ranged ~ 5% around the same value, which was confirmed as blank in the mass spectrometric single ion search of cascade b (see 4.3.4.4).

Other possible methods for quantification of alanine and serine are enzymatic assays (*vide infra*), derivatization with dansylchloride, GC separation after peracetylation, *tert*-butyltrimethylsilyl chloride derivatization, or separation of the neat amino acid *via* hydrophilic interaction liquid chromatography (HILIC). Employing Marfey's reagent even enables resolution of diastereomers on achiral HPLC columns but this is relatively expensive.

Hydroxypyruvate was spectrophotometrically monitored during early cascade reactions, tested by several methods. For example, the WST-1 method discussed above, or monitoring of NADH consumption during reduction of HPA (hydroxypyruvic acid) with lactate dehydrogenase (Figure 103).

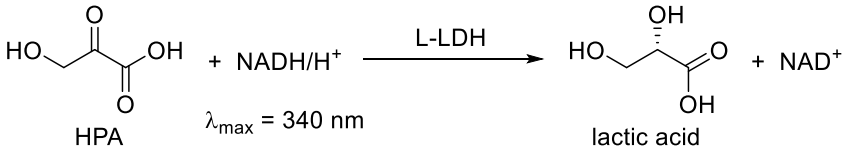


Figure 103: Reaction equation of the colorimetric L-lactate dehydrogenase (L-LDH) assay for hydroxypyruvate quantification with UV-detection of the consumed nicotinamide adenine dinucleotide hydride (NADH) at 340 nm.

Both methods were feasible but are relatively sensitive towards components within the cell-free extract and thus had relatively high limits of quantification, estimated to lie around 1 – 2 mM of analyte in the reaction mixture. The L-lactate dehydrogenase (L-LDH) assay could be applied to measure concentrations of hydroxypyruvate, D-serine, and total serine. To measure the sum of D-serine and hydroxypyruvate concentrations, DAAO was added to the L-LDH assay mixture and to measure hydroxypyruvate and total serine, ProRCMS#9 is added as well.

The addition of **coenzymes** is in general not necessary, except for transketolase. Small concentrations (0.1 mM) of pyridoxal-5'-phosphate (PLP) and flavin adenine dinucleotide (FAD) were added anyways, because this allowed replacement of decomposed cofactors and therefore increased enzyme long term activity. Specifically, it has been shown that PLP can be degraded in enzymes upon exposure to light. [331] Therefore, reaction vials were painted black. Addition of FAD has been shown to increase DAAO stability. [328]

4.3.4.3 Test Reaction Cascades

Conditions of the initial cascade reactions were designed to maximize comparability with cascades conducted with D-serine and without racemase, conducted by L'Enfant *et al.*. [233] The publication states that she employed 0.3 mg/mL purified transketolase from *Geobacillus stearothermophilus* and 0.2 mg/mL purified DAAO from *Rhodotorula gracilis* under very similar conditions to what we initially employed, including bubbling of oxygen. Also, they claim the system to be stable for 32 h and the acceptor to be used up at a molar ratio of 1/1 with regard to serine. We found this claim to be incorrect and early progress was hindered following the literature.

Later cascade reactions following our own protocol were designed to firstly be quick and easy to establish and secondly, to exhibit maximized product production rates with high product titers and minimum loss of hydroxypyruvate. For this purpose, enzyme loading was maximized and enzyme proportions, including transketolase, were optimized not to accumulate hydroxypyruvate.

It should be noted that also early cascades, where enzyme proportions were not yet optimized, exhibited only small amounts of hydroxypyruvate (< 5 mM). This fact is shown in Figure 104, which displays the time course of one of our first continuously monitored large-scale (20 mL) cascade reactions. In these experiments, L-serine, D-serine, and hydroxypyruvate were monitored with the L-LDH assays, while the reaction conditions mimicked those of L'Enfant *et al.* (*vide supra*) with the literature loading of 0.2 mg/mL of oxidase and a 1.5/1 racemase/oxidase proportion.

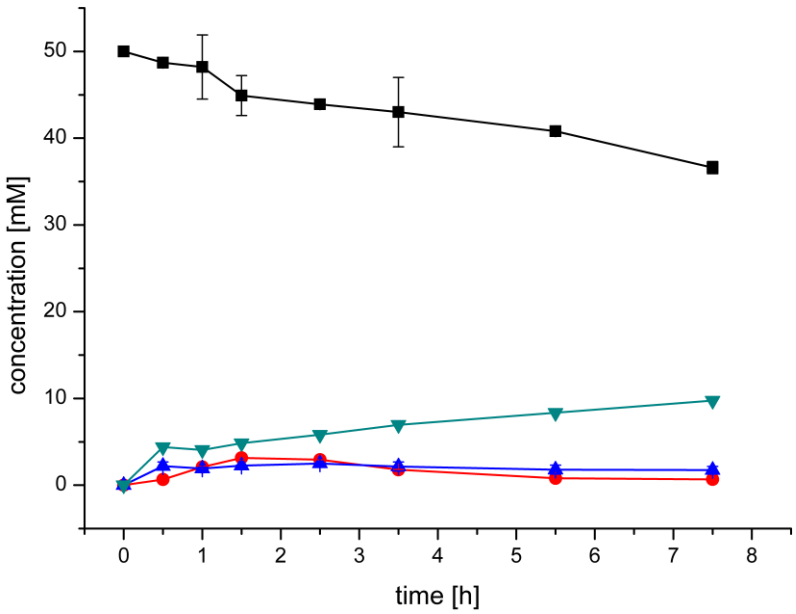


Figure 104: Time course of concentrations during the cascade reaction between 50 mM L-serine (■) and 100 mM ***rac*-11** (*rac*-2,4-dihydroxybutanal), yielding **30** (5-deoxyfructose). Red circles (●) display D-serine concentrations, blue triangles (▲) display hydroxypyruvate concentrations, cyan diamonds (▼) display converted ***R*-11** (2(*R*),4-dihydroxybutanal) as determined by HPLC/UV of PLP-derivatized ***R*-11/*S*-11** (2R, 2S,4-dihydroxybutanal, respectively) peak ratios. Reaction conditions: 0.23 mg/mL purified ProRCMS#9, 0.21 mg/mL purified DAAO from *Rhodotorula gracilis*, 0.32 mg/mL purified transketolase from *Geobacillus stearothermophilus* in H_2O, pH 7.0, maintained by titration of 0.1 M HCl, 0.1 mM FAD, 0.1 mM PLP, 1.2 mM $MgCl_2$, 0.1 mM ThDP, O_2 bubbled at 15 mL/min, 25 °C, 2297 U of commercial catalase from bovine liver added every 30 min.

The diagram shows that both, D-serine and hydroxypyruvate are maintained at a low concentration below 5 mM throughout the cascade reaction at an enzyme ratio of 1.1/1/1.5 (ProRCMS#9/DAAO/TK) of preparations with 0,32 mg/mL transketolase and 50 mM as starting concentration of L-serine. After overnight incubation, the conversion had come to a halt, strong foaming was observed and addition of hydroxypyruvate lead

to a burst of conversion. This led to the conclusion that the DAAO had been completely deactivated.

To overcome the oxidase deactivation, a new method of oxygen introduction was developed by introducing oxygen *via* a semipermeable silicone membrane, without bubbling, as described above. This led to an improvement of the long-term stability of the reaction system with slow ongoing conversion up to 22 h (Figure 105).

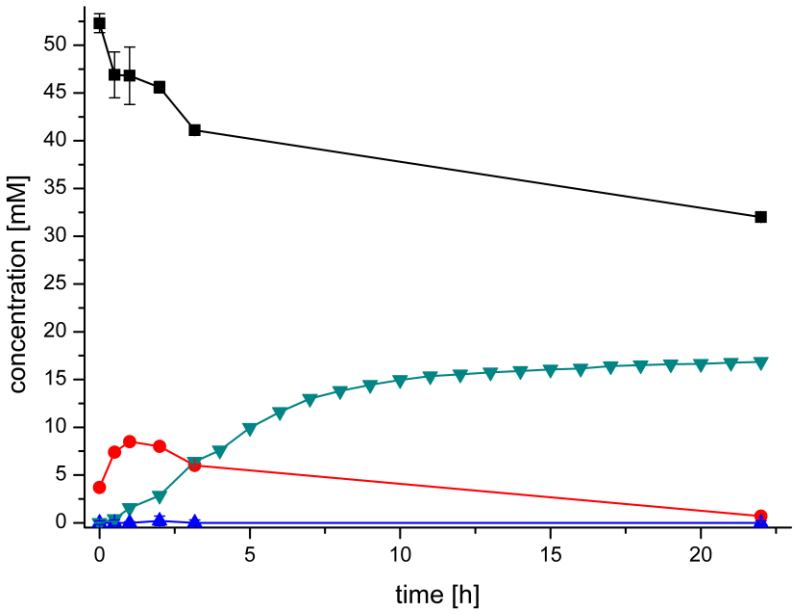


Figure 105: Time course of concentrations during the cascade reaction between 50 mM L-serine (■) and 100 mM ***rac*-11** (*rac*-2,4-dihydroxybutanal), yielding **30** (5-deoxyfructose). Red circles (●) display D-serine concentrations, blue triangles (▲) display hydroxypyruvate concentrations, cyan diamonds (▼) display converted ***R*-11** (2(*R*),4-dihydroxybutanal) as deduced from the titrated volume of HCl. Reaction conditions: 0.53 mg/mL purified ProRCMS#9, 0.12 mg/mL purified DAAO from *Rhodotorula gracilis*, 0.35 mg/mL purified transketolase from *Geobacillus stearothermophilus* in H_2O, pH 7.0, maintained by titration of 0.1 M HCl, 0.1 mM FAD, 0.1 mM PLP, 9.0 mM $MgCl_2$, 2.4 mM ThDP, O_2 *via* a silicone tubular membrane with 2 bar internal pressure of oxygen, 25 °C, 1149 U of commercial catalase from bovine liver added every 30 min.

During the experiment, hydroxypyruvate remained at a concentration below 0.5 mM, while D-serine concentration peaked at 1 – 2 h after the start of the cascade around 8 mM. In this cascade reaction, enzymes were employed in a ratio of 4.4/1/2.9 (ProRCMS#9/DAAO/TK) with 0.35 mg/mL transketolase and 50 mM as starting concentration of L-serine.

The results indicated unclear inhibition of the racemase, which could not be confirmed in later experiments.

A third experiment, with a higher oxidase loading of 0.65 mg/mL (3.25-times the original loading) and a racemase/oxidase proportion of 3.8/1 was conducted. According to our racemase/oxidase proportion experiments, this should correspond to

a hydroxypyruvate production of apprx. 1.3-times of that of the original oxidase loading of 0.2 mg/mL. This time, only the consumption of HCl was monitored, with results corresponding to a significantly higher conversion to **30** (5-deoxyfructose, Figure 106).

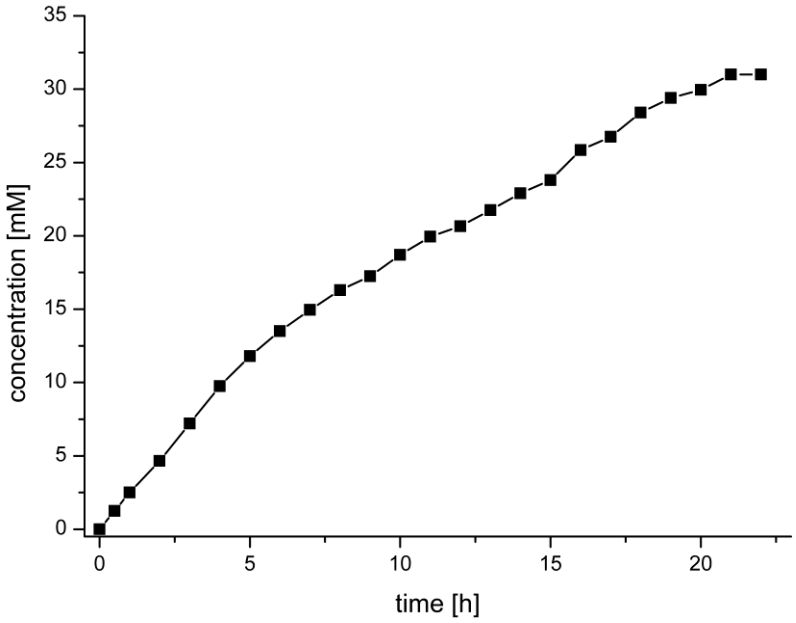


Figure 106: Time course of concentrations during the cascade reaction between 50 mM L-serine and 100 mM ***rac*-11** (*rac*-2,4-dihydroxybutanal), yielding **30** (5-deoxyfructose). Black squares (■) display converted ***R*-11** (2(*R*),4-dihydroxybutanal) as deduced from the titrated volume of HCl. Reaction conditions: 2.49 mg/mL purified ProRCMS#9, 0.65 mg/mL purified DAAO from *Rhodotorula gracilis*, 1.49 mg/mL purified transketolase from *Geobacillus stearothermophilus* in H_2O, pH 7.0, maintained by titration of 0.1 M HCl, 0.1 mM FAD, 0.1 mM PLP, 9.0 mM $MgCl_2$, 2.4 mM ThDP, O_2 *via* a silicone tubular membrane with 2 bar internal pressure of oxygen, 25 °C, 1149 U of commercial catalase from bovine liver added every 30 min.

In this cascade reaction, about 64% conversion was reached after 22 h with an enzyme ratio of 3.8/1/2.3 (ProRCMS#9/DAAO/TK), 1.49 mg/mL of transketolase and a serine starting concentration of 50 mM. The oxygen provision was shut off during the night by accident as described above.

Table 14: Comparison of initial results from the conversion of L-serine and acceptor (***rac*-11**) to **30** (5-doxyfructose) under different one-pot cascade conditions.

	Enzyme proportions	Transketolase [mg/mL]	Conv. 22 h	Oxygen	Conversion ongoing
1	1.1/1/1.5	0.32	20%	bubbling	no
2	3.8/1/2.3	0.35	33%	membrane	yes
3	4.4/1/2.9	1.49	64%	membrane	yes

The results clearly indicate that the system profits from higher enzyme loadings.

A cascade reaction at larger scale (20 mL) was performed with 210 mM ***rac*-11** (*rac*-2,4-dihydroxybutanal, 1.05 eq.) and 100 mM L-serine with the optimized enzyme concentrations and continuously monitored as described above (Figure 107).

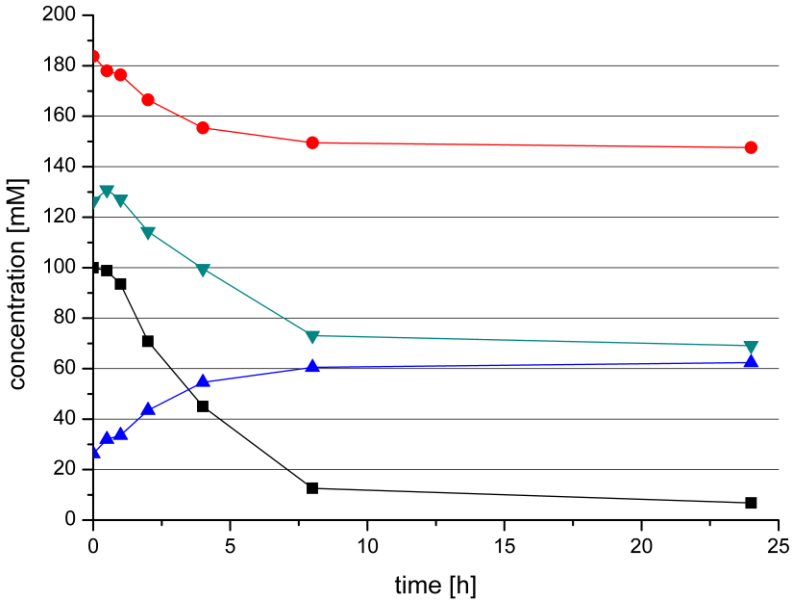


Figure 107: Time course of concentrations during the cascade reaction between 100 mM L-serine (■) and 210 mM *rac*-2,4-dihydroxybutanal (●), yielding 5-deoxyfructose (**30**, ▲). Cyan diamonds (▼) show the sum of product **30** (5-deoxyfructose) and L-serine, indicating a loss of donor. Reaction conditions: 4 mg/mL ProRCMS#11, 2 mg/mL DAAO from *Rhodotorula gracilis*, 6 mg/mL transketolase from *Geobacillus stearothermophilus* in TEA, 100 mM, pH 7.0, 0.1 mM FAD, 0.1 mM PLP, 9.0 mM $MgCl_2$, 2.4 mM ThDP, 1 atm O_2, 25 °C.

It should be kept in mind that contained errors are estimated to be above 10%. Nonetheless, the diagram clearly indicates that a part of the L-serine is lost during the reaction (~ 50 mM), without leading to product formation. Reasons for the loss of L-serine and hydroxypyruvate were already discussed (*vide supra*). The hydroxypyruvate CH_2 peak (4.70 ppm) has not been found in any of the NMRs and a small D-serine HPLC-peak has only been found in samples at an early stage.

The apparent loss of donor and the incomplete conversion was initially surprising, because early cascade test reactions were performed with an old batch of ***rac*-11** (*rac*-2,4-dihydroxybutanal). These test reactions always indicated full conversion of ***R*-11** (2(R),4-dihydroxybutanal) from HPLC/UV with PMP-derivatized samples. However, more detailed analysis of this batch indicated its concentration to be only around 50% of the concentration indicated on the vial, thus corresponding to a large excess of L-serine in the reactions.

With equimolar amounts of serine and aldehyde acceptor, in 24 h, only about 70 to 80% of aldehyde acceptor could be converted. Since the acceptor is expensive and difficult

to separate from the product hexose, it was decided to add serine in excess to reach 100% conversion.

To compensate for the loss of serine and reach 100% conversion of the expensive acceptor, following cascade reactions were performed with an **excess of L-serine**. The excess was optimized in small scale (200 µL) test reactions, imitating the conditions of the large-scale cascade reactions displayed in *Figure 107*, but with 200 mM acceptor (100 mM of 2(*R*),4-dihydroxybutanal) and 130, 150, or 170 mM L-serine. As shown above, 1.0 eq. (100 mM) D-serine did not lead to full conversion before. For analysis, HPLC/UV with PMP derivatization was used to monitor the ratio between (*R*)- and (*S*)-2,4-dihydroxybutanal, which was converted into 2(*R*),4-dihydroxybutanal (Figure 108).

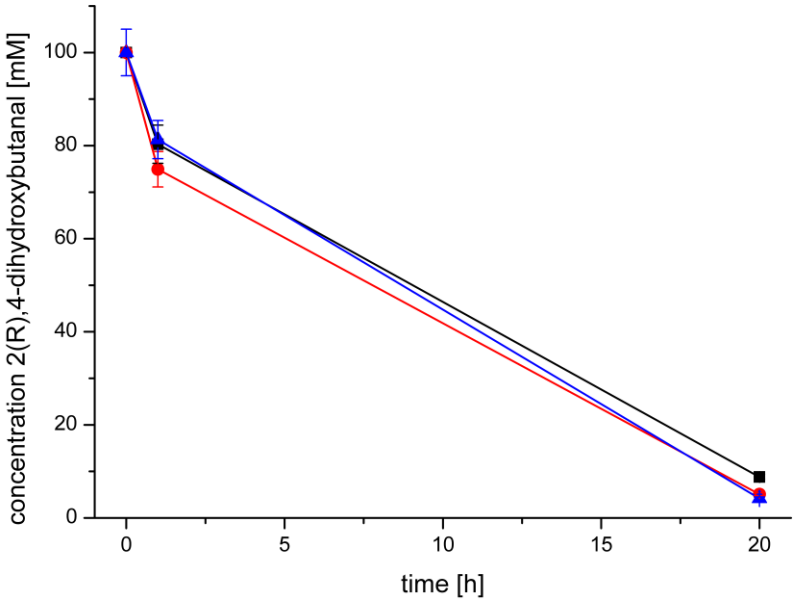


Figure 108: Time course of the ***R*-11** (2R,4-dihydroxybutanal) concentration cascade reactions between 130 mM (1.3 eq., ■), 150 mM (1.5 eq., ●), and 170 mM (1.7 eq., ▲) L-serine and 200 mM ***rac*-11** *rac*-2,4-dihydroxybutanal. Reaction conditions: 4 mg/mL ProRCMS#9, 2 mg/mL DAAO from *Rhodotorula gracilis*, 6 mg/mL transketolase from *Geobacillus stearothermophilus* in TEA, 100 mM, pH 7.0, 0.1 mM FAD, 0.1 mM PLP, 9.0 mM $MgCl_2$, 2.4 mM ThDP, 1 atm O_2, 25 °C.

The diagram shows that 1.3 eq. of L-serine is sufficient to reach more than 90% conversion of the acceptor aldehyde within 20 h, and that higher concentrations of L-serine make only little difference. For this reason, it was decided to use between 1.2 and 1.4 eq..

4.3.4.4 Final Cascade Reactions

Conditions for the final cascade reactions (a – d) are given in Table 15.

Table 15: Conditions of biocatalytic reaction cascades (20 mL, T: temperature during reaction, O_2: continuous stream of oxygen supplied to the head space) for the conversion of L-amino acid [L-aa, a - c: L-serine (ser); d: L-alanine (ala)] and 2,4-dihydroxybutanal [2,4-dhb, a: 2R,4-dihydroxybutanal (*R*); b - d: *rac*-2,4-dihydroxybutanal (*rac*)] to the corresponding deoxysugar (a - c: 5-deoxyfructose; d: 1,5-dideoxyfructose) with transketolase (a - c: *G.st.*: *Geobacillus stearothermophilus*, wild-type, heat shocked purified lyophilizate; d: Var: *Geobacillus stearothermophilus*, H102LH747S variant, heat shock-purified lyophilizate), amino acid racemase (AAR, Prozomix amino acid racemase panel ProRCMS#, enzyme number given), either DAAO (a, b: *Rhodotorula gracilis*, cell-free extract, d: Prozomix D-amino acid panel ProDAAO#4536, cell-free extract) and catalase (*Micrococcus lysodeicticus*, solution, Sigma-Aldrich (Merck) a, b: 5 kU at 0, 30, 60, 120, 240, 480 min, d: 5 kU at 0, 30, 60, 120, 240, 480, 1440 min), or transaminase (TA, Prozomix transaminase panel ProTA#17, (*R*)-selective, cell-free extract) and α-ketoglutarate (α-kg, 140 mM, disodium salt) in triethanolamine buffer(100 mM, pH 7.0).

	a	b	c	d
2,4-dhb [mM]	110 (2R)	200 (*rac*)	200 (*rac*)	100 (*R*)
L-aa [mM]	140 (ser)	140 (ser)	140 (ser)	110 (ala)
α-kg [mM]	-	-	140	-
TK [mg/mL]	8 (*G.st.*)	8 (*G.st.*)	8 (*G.st.*)	8 (Var)
DAAO [mg/mL]	2 (*R.g.*)	2 (*R.g.*)	-	4 (#4536)
TA [mg/mL]			10 (#17)	
AAR [mg/mL]	4 (#11)	4 (#9)	4 (#9)	4 (#31)
T [°C]	25	25	25	45
O_2	yes	yes	no	yes

Time dependent concentration curves are displayed in Figure 109.

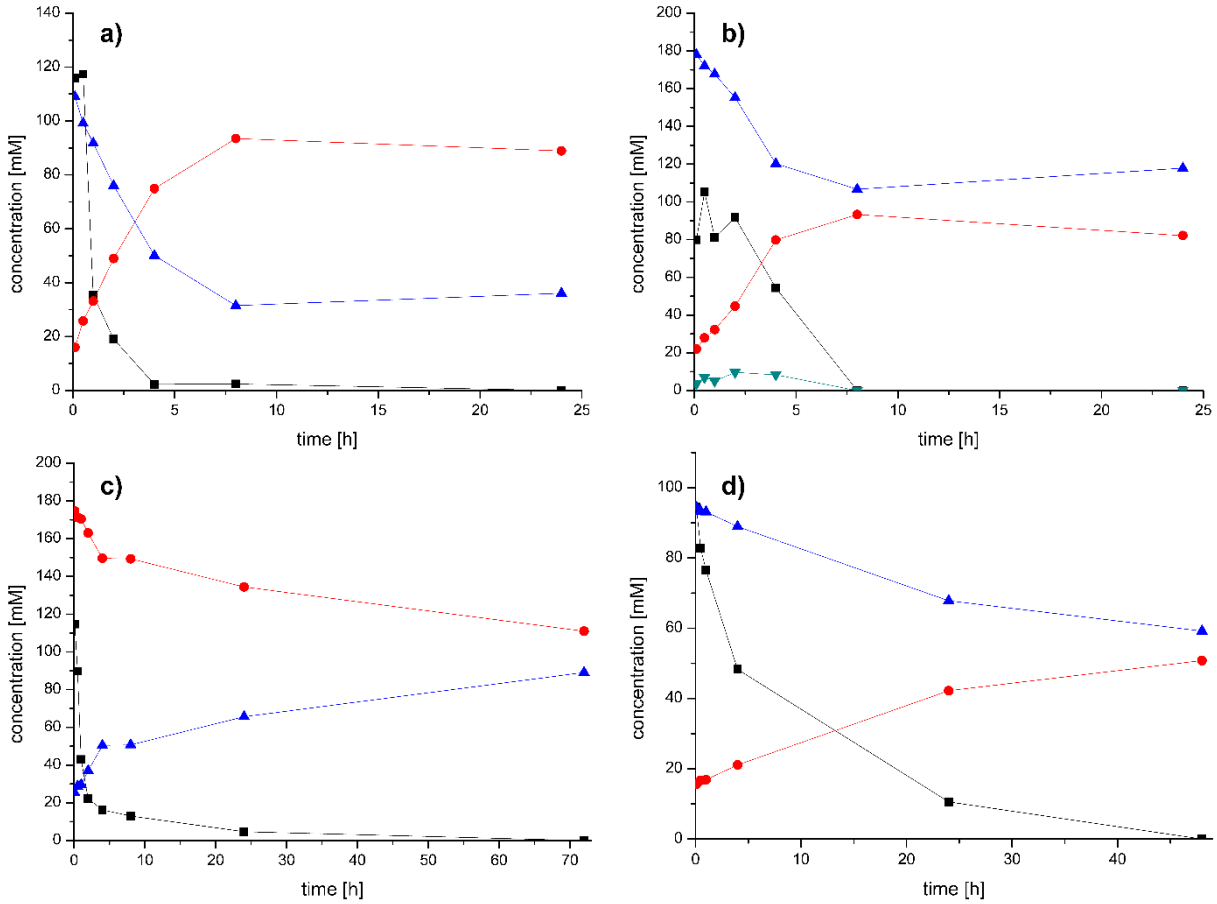


Figure 109: Time course of reaction cascade a - d: Conversion of L-amino acid (■: a - c: L-serine, 140 mM; d: L-alanin, 110 mM) and 2,4-dihydroxybutanal (▲: a: 2R,4-dihydroxybutanal; b - d: *rac*-2,4-dihydroxybutanal;) to the corresponding deoxysugar (●: a - c: 5-deoxyfructose; d: 1,5-dideoxyfructose) with transketolase (a - c: *Geobacillus stearothermophilus*, wild-type, heat shock-purified lyophilizate; d: *Geobacillus stearothermophilus*, H102LH747S variant, heat shocked purified lyophilizate), amino acid racemase (AAR, Prozomix amino acid racemase panel ProRCMS#, a: #11; b, c: #9; d: #31), either DAAO (a, b: *Rhodotorula gracilis*, cell-free extract, d: Prozomix D-amino acid panel ProDAAO#4536, cell-free extract) and catalase (*Micrococcus lysodeicticus*, solution, Sigma-Aldrich (Merck) a, b: 5 kU at 0, 30, 60, 120, 240, and 480 min, d: additionally 10 kU at 1440 min), or transaminase (Prozomix transaminase panel ProTA#17, (*R*)-selective, cell-free extract) and α-ketoglutarate (140 mM, disodium salt) in triethanolamine buffer(100 mM, pH 7.0). In b, the amount of D-serine (▼) is also given.

Table 16 displays yields of the corresponding ketohexose, corresponding to the 2(*R*),4-dihydroxybutanal, reaction times, employed amino acid equivalents and productivities in relation to the employed amino acid racemase cell-free extract of the four cascade reactions.

Table 16: Isolated yields, reaction duration, donor (L-serine, or L-alanine, forming hydroxypyruvate, or pyruvate during the amine oxidation) equivalents and productivities [(mg product)/(mg racemase cell-free extract)] of cascade reactions a – d (Figure 95, Table 15).

Cascade Reaction	Yield	Reaction duration [h]	Donor equivalents	Productivity [g/g]
a) Oxidase, L-serine, (*R*)-acceptor	quant.	24	1.3	4.1 (ProRCMS#11)
b) Oxidase, L-serine, *rac*-acceptor	quant.	24	1.4	4.5 (ProRCMS#9)
c)Transketolase, L-serine, *rac*-acceptor	76%	72	1.4	3.4 (ProRCMS#9)
d) Oxidase, L-alanine, *rac*-acceptor	55%	54	1.1	1.9 (ProRCMS#31)

All reactions produced high product yields. Shortest reaction times and best yields and productivities were reached with the ProRCMS/oxidase/transketolase setup (a, b), outperforming the reaction setup of L'Enfant *et al.* in terms of yield and product titers, even though a generally less well accepted transketolase acceptor was used in our synthesis. [233]

The cascades also deliver higher yields than the small-scale (2.5 mL) reaction performed by Lee *et al.*, who reacted 100 mM L-[3-^{14}C]serine with 35 mM D-ribose-5-phosphate, employing alanine racemase, DAAO and transketolase, yielding 69% of product. [280]

The yield of the ProRCMS/TK/TA cascade lies in the range of those conversions, given in the paper from Bawn et al., describing the one-pot synthesis of L-*gluco*-heptulose. [277] It should be kept in mind that the employed transaminase is a rare (*R*)-selective enzyme that can be outperformed by (*S*)-selective enzymes. This enzyme was selected for racemase application from L-serine, to extend the existing enzymatic toolbox. Also, the enzyme is one of our best transaminases for application at 25 °C, which was desired for better comparability of low temperature transketolase cascades.

With high activity (*S*)-selective transaminases and L-serine at higher temperatures, higher yields and productivities should be achieved. Combination of (*S*)-selective transaminases, racemase and racemic serine is another interesting application, due to low *rac*-serine prices (*vide supra*). However, transaminase catalysis always has the flaw of requiring another cosubstrate, leading to a lower atom economy.

The newly available thermostable ProDAAO#4536 has also been successfully applied for production of **31** (1,5-dideoxyfructose). Lower yield and productivity originate from the lower transketolase activity on pyruvate.

Product separation proved to be advantageous for cascades performed with enantioenriched acceptors, due to lower silica-gel column load.

An alternative biocatalytic route to **30** and **31** (5-deoxy and 1,5-dideoxyfructose, respectively) has been published in 2011 by Rale *et al.* applying the two aldolases fructose-6-phosphate aldolase (FSA) and a transaldolase B (TalB) variant (F178Y), both originated from *E. coli*. [332] The reactions employ **32** (3-hydroxypropanal) as acceptor and variation of the donor between dihydroxyacetone (R = CH_2OH) and hydroxyacetone (R = CH_3) yields either **30**, or **31** (5-deoxy, or 1,5-dideoxyacetone, respectively) as products (Figure 110).

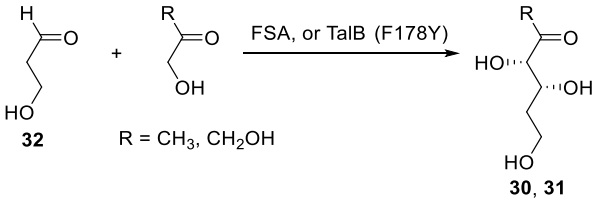


Figure 110: Alternative aldolase pathway to deoxysugars **30** and **31**. [332]

In their experiments, conversions above 75%, as monitored by densitometric thin layer chromatography, were achieved in 24 to 48 h, relative to **32** (3-hydroxypropanal), which was employed at 100 mM concentration with 1.5 equivalents of the corresponding donor.

The big advantage in this methodology is that both educts can be generated from glycerol, as demonstrated by Yang *et al.*. [333,334] Glycerol is not only cheap, but also available in large quantities, industrially not fully consumed at the moment, due to its availability from the production of biodiesel (transesterification of fats with methanol). [335,336] For this reason, processes that are suited to generate valuable products from glycerol are desired by industry.

A similar approach has been followed by Liu *et al.* already in 1991. [337] They used fructose-1,6-diphosphate aldolase and acid phosphatase to convert fructose 1,6-diphosphate and 3-hydroxypropanal into 5-deoxyfructose, reaching about 70% of isolated yield. However, they needed an almost 4-fold excess of **32** (3-hydroxypropanal), besides the expensive and unstable dihydroxyacetone phosphate nucleophile.
Even though the aldehyde can be generated from glycerol, it is problematic, because it is readily dehydrated to volatile and cancerogenic acrolein.

The products **30** (5-deoxyfructose) and **31** (1,5-dideoxyfructose) are rare sugar derivatives. In nature, specific reductases produce deoxy sugars in one step from the corresponding oxygenated sugar, while chemical production is difficult and requires multiple steps. However, only few deoxy sugars are abundant in nature, namely 2-deoxy-D-ribose, 6-deoxy-L-mannose (L-rhamnose), 6-deoxy-L-galactose (L-fucose), and 6-deoxy-D-glucose (D-quinovose). [338]

Rare sugars are in general interesting high-value compounds for food and pharmaceutical industries. Many of them taste sweet and thus are potential candidates for artificial sweeteners in diet foods with a low glycemic index. For example, **30** (5-deoxyfructose) has been reported to be nearly as sweet as fructose. [339] This property becomes increasingly interesting as pressure on national and international governments is increasing to regulate and limit sugar contents in food. [340] Another important market for sugars in food and medicine are prebiotic oligosaccharides, for example 2'-fucosyllactose, sold as high value ingredients in baby food. [341]

Sugars are important recognition elements in bioactive molecules, regulating physiological and pathological processes. As such, many sugars, their derivatives and glycosides are pharmaceutically active and conjugation to other bioactive molecules can potentially increase specificity and potency of derived drugs. For this reason, screening libraries for glycoconjugates are built up, to elucidate and harness this vastly undiscovered pharmaceutical potential. [342,343] One example for a glycoconjugate drug is the broad-spectrum antibiotic vancomycin. Aside from their use as glycoconjugates, fructose derivatives are substrate analogues and potential regulators for glucosidase and other sugar-processing enzymes.

Additionally, deoxy sugars are interesting mechanistic probes for sugar-processing enzymes, because their corresponding kinetic constants provide important information about the underlying molecular recognition. [344]

Comparing the NMR of **30** (5-deoxyfructose) with a D-fructose NMR from the literature [345] reveals the expected absence of furanose forms (Figure 111) in the 5-deoxyfructose (**30**) spectra.

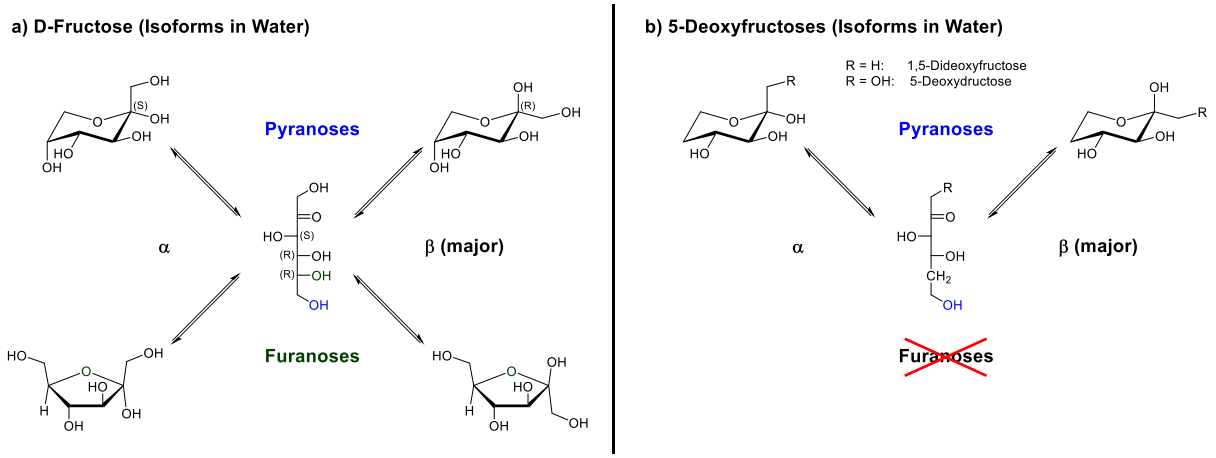


Figure 111: D-Fructose, 5-deoxy-, and 1,5dideoxyfructose isoforms in water.

Bearing no hemiacetal forming hydroxy group at C-5, 5-deoxyfructose (**30**) cannot form 5-ring constitutional isomers. The β-anomer with 3-OH and 4-OH in equatorial positions is the conformation with the lowest conformational energy (9.44 kcal/mol) as calculated in Chem3D. The energy was minimized by rotation of exocyclic bonds with angle optimization for minimized energy and subsequent MM2 energy minimization. The lowest energy found for the α-anomer was 12.4 kcal/mol.

The respective hydrogen bond setup with the lowest conformational energy found is displayed in Figure 112.

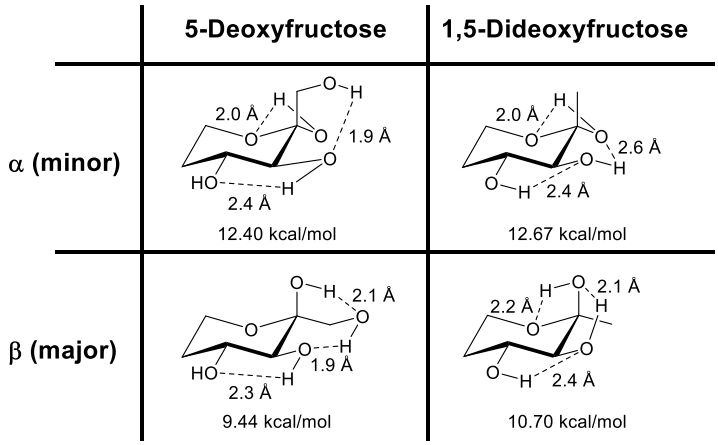


Figure 112: Computed 5-deoxyfructose (**30**) and 1,5-dideoxyfructose (**31**) conformational energy and intramolecular hydrogen bonds of MM2-optimized conformations.

The ^{13}C NMR spectra reveal a second isomer. Due to the absence of an aldehyde signal and with a second matching hemiacetal signal, this was concluded to be the less abundant α-anomer.

4.4 Application of Panel Racemases B: Preparation of α-Keto acids

As another example for the application of amino acid racemases, the production of 2-keto acids **33**, **34**, **35** was selected, representing the least complex cascade system, only employing amino acid racemase and DAAO together with catalase.

It was chosen to produce phenylpyruvate (**33**), 2-ketoisovalerate (**34**), and 2-ketoglutarate (**35**) from L-phenylalanine (L-Phe), L-valine (L-Val) and L-glutamic acid (L-Glu) as examples for the application of panel racemases in the combination with DAAO (Figure 113).

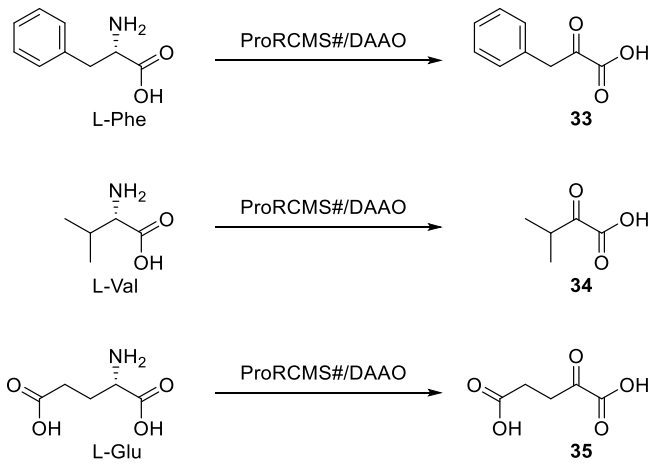


Figure 113: Targeted L-α-amino acid to α-keto acid conversions, through combination of suitable panel racemases (ProRCMS#) and the DAAO from *Rhodotorula gracilis*.

Keto acids are valuable building blocks in chemical synthesis, bearing the ability to undergo a multitude of useful transformations and substrates for many catalytic enzyme classes. Examples for useful applications, possibly carried out in a biocatalytic fashion, are the addition of nucleophiles, the generation of aldehydes *via* decarboxylation and reduction to enantiopure α-amino, or α-hydroxy acids (Figure 114).

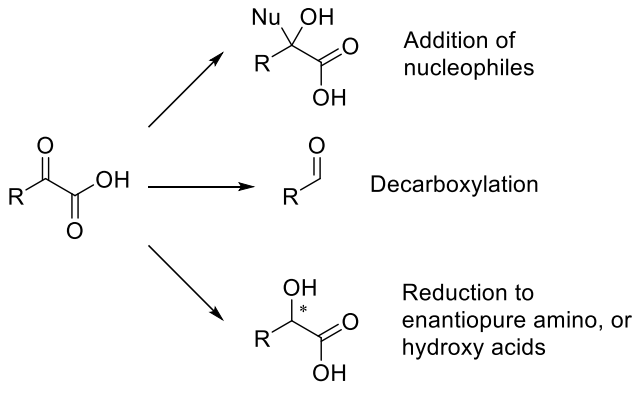


Figure 114: Exemplary application of α-keto acids

4.4.1.1 Selection of Amino acid Substrates

Phenylpyruvate (**33**) was selected, because its *in situ* production is especially useful due to its limited stability (*vide supra*). Even though α-keto acids are relatively resistant to decarboxylation as compared to their β-keto acid homologues, some of them undergo a decomposition reaction, especially if the developing carbanion is further stabilized such as in the case of **33** (phenylpyruvate). The formed benzaldehyde is easily oxidized when in contact with air and can be detected due to its characteristic smell and the characteristic needle-like crystals of the finally developing benzoic acid.

For the following discussion of prices, it should be kept in mind that L- and racemic canonical amino acids are in general bulk chemicals and therefore prices from research chemical vendors are little representative and largely overestimated as discussed in Chapter 4.3.3.

Nonetheless, **33** (2-keto acid) is in general more expensive than its easily accessible canonical L-amino acid counterpart L-phenylalanine (98%, 100 g: 127 €) [346], phenylpyruvate (sodium salt, 25 g: 97 €) [347].

34 (2-keto isovalerate) (95%, sodium salt, 25 g: 653 €) [348] on the other hand costs more than ten times of L-valine (25 g, 98%: 46,10 €) [349]. This exemplary transformation was also chosen, because it is an example of a bulkier amino acid and a relatively bad substrate for known racemases.

α-Ketoglutarate (**35**) (98%, disodium salt dihydrate 100 g: 145 €)[350] is not an expensive compound, however still significantly more expensive than its L-amino acid counterpart L-glutamic acid (99%, 100 g: 37,10 €) [351]. Also, *in situ* conversion of L-glutamic acid can be useful in biocatalytic cascades for α-ketoglutarate regeneration during transamination. L-Glutamic acid is an example of an acidic amino acid and racemase/oxidase catalysed conversion to α-ketoglutarate is highly unusual, because of the low substrate reactivity with these enzyme classes.

4.4.1.2 Envisioned Reaction System

The envisioned reaction system comprised of the chosen panel racemase and the DAAO from *Rhodotorula gracilis*, both as cell-free extract. To quench the developing hydrogen peroxide, avoiding product and catalyst degradation and recovering half of the used oxygen, a commercial catalase was used (FigureFigure 115).

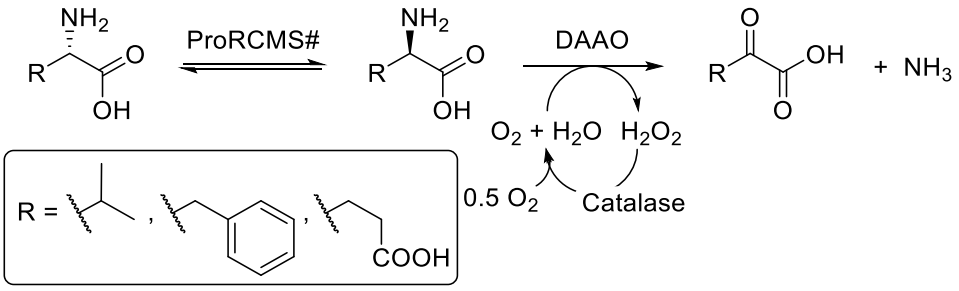


Figure 115: General reaction scheme for the production of α-keto acids from L-amino acids, employing a suitable panel racemase (ProRCMS#), the DAAO from *Rhodotorula gracilis* and commercial catalase.

A similar system has been presented in the literature, using immobilized enzymes, but unable to produce **33** phenylpyruvic acid. [352]

In this work, catalysts were employed in dilute potassium phosphate buffer (25 mM), L-phenylalanine used at 100 mM, L-valine at 200 mM, and L-glutamate at 250 mM concentration, respectively, and the reactions were applied on a 20 mL-scale. 100% Oxygen was supplied as a continuous stream above the agitated reaction mixture. To increase DAAO lifetime, the reactions were carried out at 25 °C overnight, without bubbling of oxygen and in a Falcon tube lying on its side. The chosen reaction setup was supposed to increase the surface of the liquid phase, to improve oxygen uptake (*vide supra*).

4.4.1.3 Selection of Racemases

From the initial screening, best panel racemase candidates for these three transformations were ProRCMS#11, #29, #36, #46, and #53, with measured conversions of around 15, 95, 35, 25, and 20%, respectively for L-phenylalanine, ProRCMS#13, #29, and #53 with conversions below the limit of quantification for L-valine and ProRCMS#4, #53 with conversions of around 15 and 10%, and ProRCMS#4 found to be active on L-glutamate with a conversion < 5%. It should be noted that, even though the relative oxidase activity on valine is stated to be between 33 and 95% in the literature (~20 mM, Table 3), valine delivered a very low response in our assay, even with the 0.4 mM calibration standard, and as such has a high intrinsic limit of detection and quantification. This is probably due to the high K_M of the DAAO from *Rhodotorula gracilis* for this substrate (Table 4).

The hits for glutamate racemization by ProRCMS#4 and #53 during the panel screening were unexpected and at first attributed to mistakes in handling during the assay procedure, since glutamate is no, or a very bad substrate for DAAO from

Rhodotorula gracilis and for known PLP-dependent amino acid racemases, even for broad spectrum racemases (Table 3 and Table 5). As such, D-glutamic acid also delivered a small concentration response during the assay and consequently high limits of detection. There is a separate class of glutamate racemases, which is PLP-independent, but these enzymes are not expected to be part of the ProRCMS panel (*vide supra*). [353,354]

4.4.1.4 Small-Scale Test Reactions

To confirm the screening results, choose the best racemase catalyst, and to avoid wasting material, the candidate racemases were applied in small-scale test reactions in Eppendorf vials at a 500 µL-scale. In this small-scale experiment, also the best hit for glutamate (ProRCMS#4) was included. All reaction components were added to the individual vial and then incubated at 25 °C. To enable oxygen uptake, the vial caps were punctured. After 20 h, reaction progress was checked by TLC, detecting the amino acid with ninhydrin and the oxo acid with Seebach reagent (Figure 116).

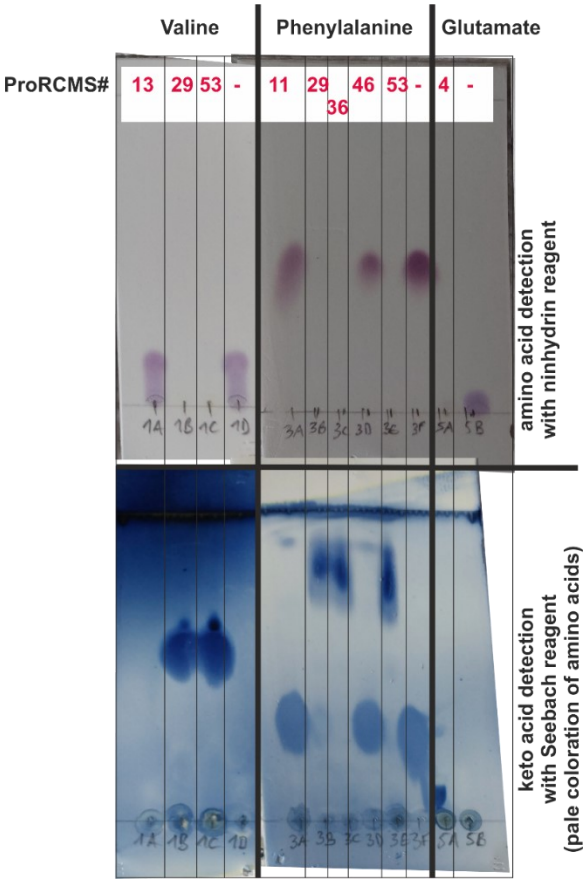


Figure 116: TLCs of the small-scale experiments, run in two different solvents (see experimental) and coloured with ninhydrin, or Seebach reagent. Amino acids colour red and keto acids colour dark blue. In the case of Seebach reagent colouration, amino acids also show in pale blue. Used panel racemases are provided with their number in red. Columns without number indicate the corresponding racemase blank.

TLCs show that ProRCMS#29 and #53 function in a cascade reaction, converting amino to keto acid. ProRCMS#13 is either a false hit for valine or did not work in the cascade for another reason. Likewise, for phenylalanine conversion, ProRCMS#29, #36, and #53 were suited, while ProRCMS#11, and #46 did not enable the biocatalytic transformation. Surprisingly, glutamate was converted to the corresponding α-ketoglutarate through the action of ProRCMS#4 and DAAO from *Rhodotorula gracilis*, which was very surprising due to aforementioned reasons.

4.4.1.5 Large-Scale Synthesis

From these small-scale experiments, ProRCMS#29, #53, and #4 were selected, in a 1, 4 mg/mL, and 10 mg/mL loading of cell-free extract, for the large-scale preparative synthesis of phenylpyruvate (**33**), 2-keto isovalerate (**34**), and 2-ketoglutarate (**35**), respectively.

Phenylpyruvate (**33**) and 2-keto isovalerate (**34**) could be obtained by acidic extraction into diethyl ether, yielding 71% of **34** (2-keto isovalerate) and **33** (phenylpyruvate) quantitatively. Concluding from the characteristic smell, some of the liquid 2-keto isovalerate (**34**) was lost during solvent evaporation.

The colourless amorphous **33** (phenylpyruvate) developed the characteristic benzaldehyde smell within hours, even under argon, and formed the characteristic benzoic acid needle-like crystals when left under air. These findings provide further proof for the lack of stability of free phenylpyruvic acid (**33**) and the usefulness of *in situ* generation.

2-Ketoglutaric (**35**) acid purification *via* extraction was not successful. Therefore, purification by column chromatography was chosen, unfortunately yielding only an impure product. The NMR after solvent roto-evaporation still contained a lot of formic acid. Another NMR after subsequent lyophilization before weighting the product was not recorded. Nonetheless, the two product triplets were visible in the ^{1}H NMR spectrum. Repeating the experiment and applying a different method for purification, like reverse phase chromatography or ion-exchange chromatography is advisable.

Concluding, the first application of panel racemases was achieved, producing valuable 2-keto acids **33**, **34**, and **35** with a productivity of 14.6 g/g for **33** (phenylpyruvic acid), 4.0 g/g for **34** (2-keto isovalerate), and about 1 g/g for **35** (2-ketoglutarate), calculated on the basis of the applied racemase cell-free extract. A grand value gain is

accomplished in the case of **34** 2-keto isovalerate. As such, these reactions are examples of the successful application of panel racemases, demonstrating also unusual amino acid racemase conversions and with the underlying reactions ready to be coupled to other catalysts, as demonstrated in the subsequent section of this work.

4.5 Application of Panel Racemases C: Preparation of α-Hydroxy Acids

As a second example for *in situ* racemisation, it was chosen to produce D-2-hydroxyphenylpropionic acid (**36**, D-phenyllactic acid) from L-phenylalanine and D-2-hydroxy-3-(4-hydroxyphenyl)propionic acid (**37**) from L-tyrosine (Figure 117).

Figure 117: Targeted L-α-amino acid to D-α-hydroxy acid conversions, through combination of suitable panel racemases (ProRCMS#), the DAAO from *Rhodotorula gracilis* and D-hydroxyisocaproic acid dehydrogenase (D-HicDH).

Enantiopure D-2-hydroxy acids in general are valuable building blocks for pharmaceuticals and polymers. For example, D-2-hydroxy-4-fluorphenyl propionic acid is a key building block of Ruprintivir™, an API for the treatment of acute viral infections. [355] Also, such compounds are added to cosmetics as fillers. For more information and publications, see the theory part about D-2-hydroxyisocaproate dehydrogenase.

D-Phenyllactic acid (**36**) was selected for the same reasons as **33** in the previous chapter, now with demonstration of *in situ* reaction of the unstable **33** (phenylpyruvate).

D-2-Hydroxy-3-(4-hydroxyphenyl)propionic acid (**37**) was chosen, because it is a very expensive compound, with 100 mg of the racemate being priced at almost 100 € on the Sigma-Aldrich website (D/L-*p*-hydroxyphenyllactic acid, 100 mg, 97%: 95,40 €). [356] Additionally, the supporting information of a publication regarding the production of 2-keto acids through combination of BsrV and DAAO claims that *p*-hydroxyphenylpyruvic acid is not available with this racemase. [352] As such, our results represent a valuable extension to the possible biocatalytic space.

Biocatalytic linear cascades for the production of α-hydroxy acids have been nicely summarized by Schrittwieser *et al.*, with L-amino acid deaminase and D- and L-2-hydroxy acid dehydrogenases as the most straightforward solution. [357]

Alternative chemical production of α-hydroxy acids from a corresponding α-amino acid usually proceeds *via* the formation of toxic diazonium salts and as typical S_N1 substitutions. However, there are also diazonium salt procedures available, leading to retention. [358]

4.5.1.1 Envisioned Reaction System

For this work, we selected the cascade reaction displayed in Figure 118.

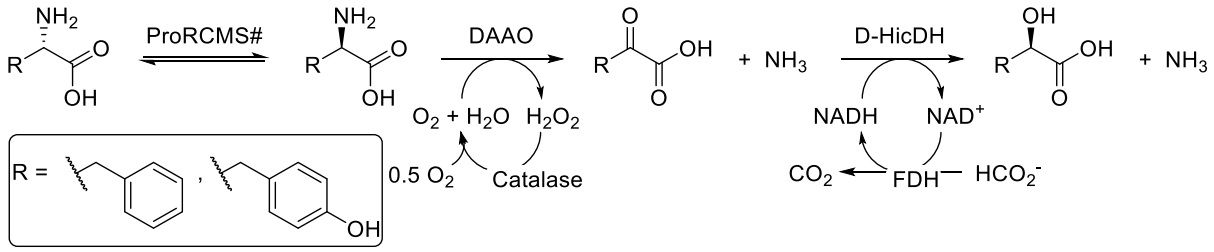


Figure 118: General reaction scheme for the production of D-α-hydroxy acids from L-amino acids, employing a suitable panel racemase (ProRCMS#), the DAAO from *Rhodotorula gracilis*, commercial catalase, D-hydroxyisocaproic acid dehydrogenase from *Lactobacillus paracasei* (D-HicDH) and commercial formate dehydrogenase.

The reactions were set up as described in the previous chapter. Likewise, the previous chapter describes included enzymes that are needed for the 2-keto acid production, but for α-hydroxy acid synthesis, the cascade system must be further extended. D-2-Hydroxyisocaproic acid dehydrogenase is employed for *in situ* reduction of the developing 2-keto acid to the corresponding D-2-hydroxy acid. As a dehydrogenase, the enzyme is dependent on a nicotinamide cofactor. This specific dehydrogenase can be driven with NADH and NADPH but prefers NADH.

Cofactor recycling is advised when using expensive nicotinamide cosubstrates and can be achieved chemically, electrochemically, photochemically, and enzymatically. Frequently used enzymatic NAD(P)H recycling systems include hydrogenase, glucose dehydrogenase, glucose-6-phosphate dehydrogenase, phosphite dehydrogenase, lactate dehydrogenase, formate dehydrogenase, or another alcohol dehydrogenase with for example ethanol, or isopropanol in excess. [359,360,361]

We chose NADH recycling with formate dehydrogenase, oxidizing formate to carbon dioxide. Advantages of this system are for example the high stability, activity, and commercial availability of the formate dehydrogenase. Disadvantage is its low turnover. Formate can be introduced into the reaction as inexpensive ammonium formate, which also represents a suitable buffer that even can be co-evaporated with water after the reaction under reduced pressure, facilitating product purification. Also, enzyme-

compatible ammonium formate can be added to the reaction mixture in excess, thereby steering the dehydrogenase equilibria to our advantage.

4.5.1.2 Selection of Racemases

The selection of amino acid racemases for the two biocatalytic cascade reactions was again based on the ProRCMS panel screening. ProRCMS#29 was selected for phenylalanine racemisation due to the positive results, described in the previous chapter. For tyrosine racemisation, candidates for the 20 mL-scale reactions were ProRCMS#29, #36, and #53, selected based on the initial panel screening and tested positive in a small-scale reaction as described in the previous chapter. Due to time constraints and, because tyrosine is a well-known substrate for D-HicDH we did not test the complete cascade setup at small-scale (200 µL) as usual, but directly at larger scale (20 mL). Instead, for the racemisation of tyrosine, ProRCMS#29 was selected from a small-scale conversion to the corresponding keto-acid and TLC analysis.

4.5.1.3 Large-Scale Synthesis

Tyrosine has a very low solubility in aqueous solutions (about 2 mM) and therefore, an amount equal to 50 mM was added to the reaction mixture as dry flakes, to be continuously dissolved during the reaction. Phenylalanine was applied at 100 mM concentration.

D-Phenyllactic acid (**36**) was isolated quantitatively as colourless needle-like crystals through etheric extraction and **37** (D-2-hydroxy-3(*p*-hydroxyphenyl)propionic acid) was isolated as a colourless fluffy solid *via* column chromatography and subsequent lyophilization with a yield of 70%.

Concluding, both **36** and **37** could be produced and isolated in good to excellent yield (quant. and 70% respectively) and purity with productivities of 15.3 and 0.9 g/g calculated on the basis of the applied racemase cell-free extract. The cascade delivered access to these valuable enantiopure aromatic hydroxy acids, and ProRCMS#29 is well-suited to racemise aromatic amino acids.

Enantiopurity was not specifically evaluated due to a lack of time and analytical options, but in particular because excellent enantioselectivity of HicDHs is well established. [89,357,362,363]

Complementary L-selective enzymes, like L-hydroxyisocaproic acid dehydrogenase from *Lactobacillus confusus* for L-2-hydroxy acid synthesis are available as well. [362,364,365,366]

It should be noted that ProRCMS#29 is also a candidate for the production of a wider variety of substituted hydroxyphenyllactic acids, like demonstrated by Busto *et al.* [367] These authors used tyrosine phenol lyase to generate a variety of substituted L-tyrosine derivatives and then transformed these in one-pot into the corresponding 2-hydroxy acids employing L-amino acid deaminase and L- or D-hydroxyisocaproic acid dehydrogenase. In these biocatalytic cascades, L-amino acid deaminase could likely be replaced by a combination DAAO and ProRCMS#29.

4.6 Summary and Outlook

As proven after adapting an assay and screening 56 amino acid racemases with 16 canonical, and 14 selected non-canonical amino acids, the ProRCMS library (Prozomix Ltd.) contains a wide variety of useful enzymes with amino acid racemase activity. As expected, many racemases were active on alanine and serine, but there were also enzymes for barely known lysine, tyrosine, and glutamate racemisation. Some of them accepted a broad substrate range, and we even found enzymes that racemise unusual substrates like allylglycine, N^{ε},N^{ε},N^{ε}-trimethyl lysine, or β-aspartame. Also, the library enzymes are suitable for synthetic use as proven in practical applications, and also because some of them proved to be thermostable. Due to these diverse and innovative properties, the ProRCMS racemase library represents a highly valuable extension to the biocatalytic toolbox, and is the ideal starting point when looking for a suitable amino acid racemase.

Unique structural features have been revealed upon modelling the sequences of ProRCMS#31 (thermostable enzyme with broad substrate scope) and ProRCMS#9 (high activity on serine), and comparing them amongst each other and with racemase structures from the literature. The drawn conclusions should be confirmed through X-ray crystallography and mutational studies. The found structural features are expected to contribute to the general structural knowledge about different kinds of fold-type II amino acid racemases, and their structural differentiation from other fold-type II PLP-dependent enzymes like dehydratases. This can be regarded as a starting point for future mechanistic studies and may contribute to the creation of new variants with increased thermostability, altered substrate scope, or even new catalytic functionalities.

It is expected that the library racemases will find broad application in the future assembly of biocatalytic cascades due to their thermostability, compatibility with other enzymes, and potential for the generation of high-value compounds from cheap, available, renewable, bio-based starting materials. It is also expected that many existing biocatalytic transformations can be extended by the racemisation module.

In this work, racemases have been exemplarily applied for several enzymatic, one-pot, sequential cascade reactions in combination with DAAO, transaminase, dehydrogenase and transketolase. Amino acid oxidases allow oxidation of specific amino acid enantiomers. Therefore, in combination with amino acid racemases, kinetic

dynamic resolution is achieved and the generated, reactive 2-keto acids are versatile building blocks, which can be isolated, or further converted *in-situ*. In the easiest case, they can be enzymatically reduced to form valuable α-amino, or hydroxy acid enantiomers. This has been demonstrated in the production of D-phenyllactic acid, and D-*p*-hydroxyphenyllactic acid. When combining the dynamic kinetic resolution-module with transketolase, diverse sugars can be built up, as demonstrated in the production of the two rare, non-natural carbohydrates 5-deoxyfructose, and 1,5-dideoxyfructose.

The well-known issue of enzyme inactivation upon bubbling of reaction mixtures has been addressed, specifically for the aeration of reactions employing DAAO. Innovative techniques have been developed and compared to increase this enzyme's stability during preparative reactions. Solutions have been found for microscale and 500 mg-scale with potential for further upscaling. This should contribute to lifting existing stability limitations on this broadly applied enzyme and help the development of future enzyme cascades.

During the exploration and optimization of different routes for the production of acceptor substrates for transketolases, we have discovered a new and more general approach for the reduction of lactones to lactols than hitherto known. This methodology bears the potential to extend the chemical space by providing a new handle to avoid overreduction to corresponding alcohols. Ways for further exploration of this methodology have been thoroughly and systematically discussed.

Concluding, this work delivers well-rounded and useful results, providing a large array of data about a new and innovative amino acid racemase library and examples of these enzymes' application in biocatalytic cascade reactions. It also dives deep into the theory of underlying principles and its technical implementation.

Additionally, it opens the door for further research in the fields of dynamic kinetic resolution in biocatalytic cascade reactions, enzyme engineering, and lactone reduction.

5 Experimental

5.1 Materials and Equipment

5.1.1 NMR Spectrometry

Spectra were recorded either on Bruker Avance 4 AC300, or DRX500 at room temperature (r.t.).

^{1}H NMR spectra were recorded either at 300, or 500 MHz frequency.

^{13}C NMR spectra were recorded either at 75, or 125.8 MHz frequency.

Samples were dissolved in fully deuterated solvents, given for each spectrum.

Spectra were referenced either on the corresponding solvent, or on a reference, given for the individual spectrum.

For quantitative NMR, samples were dissolved in D_2O containing an internal standard (maleic acid, *Trace*Cert®).

Spectra were analyzed with Mestlab MestReNova v.14.2.2-28739.

Multiplets were termed as follows: s (singlet), bs (broad singlet), d (doublet), t (triplet), q (quadruplet), m (multiplet), or combinations of these.

5.1.2 UV/Vis Spectroscopy

Samples were deployed in 96-well microtiter plates (Sarstedt, flat bottom, 200 µL) and analyzed on a Molecular Devices, LLC SpectraMax 190 Microplate reader photometer.

5.1.3 GC/MS Spectrometry

Samples were analyzed on an MD800 GCMS-system (Agilent), equipped with electron spray ionization mass spectrometry (EI-MS), and He as a carrier gas.

Silylated samples were separated on an Agilent J&W DB-5 (30 m, 0.25 mm ID, 0.25 µm df) column.

Acetylated samples were separated on an Restek Rt®-bDEXsm (30 m, 0.25 mm ID, 0.25 µm df) column.

Parameters (for both):

Injector temperature: 240 °C

Temperature program: Initial 35 °C, hold 4 min; 10 °C/min to 200 °C, hold 10 min.

5.1.4 HPLC Analysis

Samples were analyzed either on a Shimadzu Prominence LC20AT HPLC system (Shimadzu U.S.A. Manufacturing Inc. Canby, USA) with two chromatography pumps (solvent A and B), equipped with modules from the same manufacturer (communications bus module CBM-20A, autosampler prominence SIL-20A HT, degasser prominence DGU-10A_3, UV/Vis diode array detector prominence SPD-M20A with wolfram and deuterium lamp, and MS detector LCMS-2020) for gradient elution, or a Shimadzu Prominence LC-20AD HPLC system (Shimadzu U.S.A. Manufacturing Inc. Canby, USA) with one chromatography pump, equipped with modules from the same manufacturer (communications bus module CBM-20A, degasser prominence DGU-10A_{3R}, and UV/Vis diode array detector prominence SPD-20AV with wolfram and deuterium lamp) for isocratic elution. In the Shimadzu Prominence LC-20AD HPLC system, samples are injected manually *via* a six-way valve.

Columns were stored in an adjustable column oven.

5.1.5 Differential Scanning Fluorometry

Fluorometry changes upon heating of protein samples were quantified on a NanoTemper Technologies, Germany Prometheus NT.48.

5.1.6 Thin-Layer Chromatography

Thin layer chromatography (TLC) was carried out on pre-coated aluminum sheets (Supelco, TLC silica gel 60 W F_{254}s, Sigma-Aldrich (Merck), for aqueous eluents, and Supelco, TLC silica gel 60 F_{254}, Sigma-Aldrich, for non-aqueous eluents) with detection at UV (254 nm), or by staining with one of the following staining reagents. Eluents are given. If dipping into, or spraying with the reagent alone did not lead to the development of colored spots, the plate was heated.

Vanillin reagent (unspecific detection)

4-Hydroxy-3-methoxybenzaldehyde (vanillin, [152.14], 12 g, 78.9 mmol) is dissolved in a mixture of sulfuric acid (conc, 2 mL) and ethanol (95%, 200 mL).

Potassium permanganate reagent (detection of oxidizable)

Potassium permanganate ([158.03], 1.85 g, 11.7 mmol) and potassium carbonate ([138.20], 12.3 g, 89.1 mmol)) are dissolved in NaOH (200 mL, aq., 0.1 N).

Anisaldehyde reagent (detection of carbonyle)

p-Methoxybenzaldehyde (anisaldehyde, [136.15], 5.0 mL, δ = 1.119 g/mL, 5.6 g, 41 mmol) is added to a mixture (195 mL) of acetic acid (glacial, 2.06 mL) and sulfuric acid (conc., 1.88 mL) in ethanol (186 mL).

Nitrophenyl hydrazine reagent (detection of carbonyle)

Dragendorff's reagent (detection of quaternary ammonium) [368]

For 100 mL of reagent base, $Bi_5O(OH)_9(NO_3)_4$ ([1461.99], 8.0 g, 5.48 mmol) is dissolved in nitric acid (23 mL, aq., 25%). KI ([166.00], 20 g, 120 mmol) in HCl (1 mL, aq., 6 N) and water (d.i., 76 mL). The solution is filtered and stored in a brown bottle at - 4 °C.

200 mL of staining reagent are prepared by mixing water (d.i., 121 mL), HCl (30 mL, 6N), reagent base (12 mL) and NaOH (36 mL, 6 N). Under stirring, HCl (6 N) is added dropwise until a clear solution is formed. The solution is stored at 8 °C.

Ninhydrin reagent (detection of amine)

Ninhydrin ([178.14], 1.50 g, 8.42 mmol, 16.7 mM) is dissolved in acetic acid (glacial, 5 mL) in ethanol (95%, 500 mL).

Seebach's reagent (detection of alcohol)

Phosphomolybdic acid (5.0 g) and $Ce(SO_4)_2$*4 H_2O ([404.3], 2.0 g, 4.9 mmol) are dissolved in sulfuric acid (200 mL, aq., 1.0 M).

5.1.7 Other Apparatus and Devices

Freeze-dryer: Telstar, LyoQuest

Microplate Incubator: Heidolph, Incubator 1000 with lid

Automated flash chromatography with UV and ELS-detector: Teledyne ISCO, CombiFlash EZ Prep

Protein Purification by IMAC chromatography: GE Healthcare, HisTrap column

SDS-gelelectrophoresis: Bio-Rad, Mini Protean II

Ultrafiltration: Stedim Biotech AG, Vivacell 250 Sartorius

Eppendorf ultrafiltration: Amicon 8400, Membrane: YM 10

Centrifuges: Eppendorf, Centrifuge 5415D, and 5810R

pH-meter: SI Analytics, TitroLine 7000

5.1.8 Enzyme Sources

D-Amino acid oxidase from *Rhodotorula gracilis*, L-2-hydroxyisocaproic acid dehydrogenase from *Lactobacillus confusus*, D-2-hydroxyisocaproic acid dehydrogenase from *Lactobacillus paracasei*, ProDAAO#4536, and ProRCMS#1 - #56 were received from Prozomix Ltd. as freeze dried cell-free extract powder.

Catalase from bovine liver (Sigma-Aldrich), Lipase B from *candida antarctica* (immobilized, Sprin technologies) Formate dehydrogenase (FDH) from *Candida bodinii* (Boehringer Mannheim) and L-lactate dehydrogenase from rabbit muscle (Roche) were available in-house as commercial preparations.

Catalase from *Micrococcus lysodeicticus* was obtained from Sigma-Aldrich (Merck).

5.1.9 Chemical Sources

All chemicals were available at the in-house chemical library, or were obtained from Sigma-Aldrich, Merck, Fluka, Lancaster, Roth, Alfa Aesar, abcr, Acros Organics, Biosynth, BLDpharm, TCI, or Fisher scientific,

5.2 Standard Operation Procedures

5.2.1 IMAC Chromatography of Enzymes

D-amino acid oxidase (DAAO, *Rhodotorula gracilis*, 203 mg, 4.7 mg/mL, heterologously expressed in *E. coli* BL21DE3, N-term His^6-tag, Prozomix Ltd), ProRCMS#9, or #31 (from Prozomix racemase panel) is dissolved in potassium phosphate buffer (43 mL, 20 mM, pH 7.4, 0.5 M NaCl, 0.05% NaN_3) and centrifuged at 4 °C (20 min, 4500 g,). A Ni-NTA column (GE Healthcare, HisTrap, preconditioned with $NiSO_4$, 0.1 M) is conditioned with the same buffer before applying the filtered (syringe filter) enzyme supernatant. After washing the column with 10 volumes of

imidazole (30 mM) in potassium phosphate buffer (20 mM, pH 7.4, 0.5 M NaCl, 0.05% NaN_3) the eluting fraction is tested negative for proteins with Bradford reagent (Coomassie® Brilliant blue G 250, Sigma-Aldrich). The His^6-tagged enzyme is eluted with 5 volumes of EDTA (50 mM) in potassium phosphate buffer (20 mM, pH 7.4, 0.5 M NaCl, 0.05% NaN_3) until tested negative for proteins with Bradford reagent after washing the column with 3 volumes of the same buffer (20 mM, pH 7.4, 0.5 M NaCl, 0.05% NaN_3).

The column is regenerated by washing with 10 volumes of potassium phosphate buffer (20 mM, pH 7.4, 0.5 M NaCl, 0.05% NaN_3) and charged with 2 volumes of $NiSO_4$ (aq., 0.1 M).

The purified oxidase is either dialyzed (dialysis tubes, activated with NaCl, 7%, aq. at 60 °C) twice against potassium phosphate buffer (50 mM, pH 7.5) and directly used as a solution, or microfiltered, diluted (1:10) with potassium phosphate buffer (5 mM, pH 7.5, 0.1 mM FAD), microfiltered again and lyophilized to yield a yellow fluffy powder.

In the case of DAAO from *Rhodotorula gracilis*, all employed buffers are charged with flavin adenine dinucleotide (FAD, 0.1 mM) and cooled on ice.

5.2.2 Bradford-Test: Qualitative Test for Protein in Aqueous Samples

To assess the presence, or absence of protein in a stream of aqueous solution (IMAC chromatography), a sample (2 – 3 drops) is drawn into a ceramic spot plate and Bradford reagent [50 µL, Bio-Rad protein assay dye reagent concentrate, diluted (1:4) with d.i. water] is added. Formation of a blue color indicates the presence of protein.

5.2.3 Quantitative Analysis of Protein Concentrations in Aqueous Samples

Samples are analyzed according to the manual of the commercial kit (Thermo Scientific, Pierce BCA Protein Assay).

The assay mixture is prepared by mixing solution A (50 parts, bicinchoninic acid solution) solution B (1 part, $CuSO_4$ solution).

The assay mixture is added to a solution of lyophilized powder (25 µL, 1 mg/mL, cell-free extract), or a calibration (0, 20, 40, 60, 80, 100 µg bovine serum albumin standard) in the well of a microtiter plate. The well is sealed with an acetate foil, incubated (30 min, 37 °C) and spectrophotometrically analyzed (562 nm).

5.2.4 SDS-Page

5.2.4.1 SDS-Gel Preparation

Two gels are prepared by mixing L-asparagine (132 mg) with solutions of tris(hydroxymethyl)aminomethane (1.0 mL, 760 mM, pH 7.4, aq.), L-serine (1.0 mL, 1.0 M, pH 7.4, aq.), glycine (1.0 mL, 1.0 M, pH 7.4, aq.) and water (3.0 mL, d.i.). The mixture is incubated (10 min, 60 °C, 800 rpm). The prepared solution is cooled to r.t. and mixed with acrylamide (2.75 mL, 30%, aq.), *N,N'*-methylenebisacrylamide (1.16 mL, 2%, aq.), ammonium persulfate (60 µL, 10%, aq.) and *N,N,N',N'*-tetramethylethane-1,2-diamine (10 µL), stirred until homogeneous and immediately poured into gel molds (4.5 mL per mold). Casted gels are sufficiently hardened to be used after 60 min and can be stored in wet towels at 8 °C for up to two weeks.

5.2.4.2 SDS-Gel Electrophoresis

Enzyme cell-free extract (freeze dried powder, 1 mg/mL), or purified enzyme (freeze dried powder, 0.5 mg/mL) is dissolved in potassium phosphate buffer (20 mM, pH 7.5). A volume corresponding to 15 µg protein, according to BCA-assay, is mixed with Laemmli [369] buffer (1:1) and heated (5 min, 95 °C), before loading a volume, equivalent to 3 µg protein into a pocket of a previously prepared wet gel and run in a Biorad Mini Protean II (100 V) in the in-house SDS buffer (125 mM tris(hydroxymethyl)aminomethane, 0.95 M glycine, 17 mM sodium dodecyl sulfate in water).

5.2.4.3 Preparation of Coomassie Quick-Stain

The staining solution is prepared by dissolving (stirring for ~3 h) Coomassie Brilliant Blue G250 (75 mg) in water (1 L, d.i.) and acidifying (pH ~1.5) the solution with HCl (3 mL, conc.). The solution is stored in a brown bottle.

5.2.4.4 Staining of SDS-Gels

To develop an SDS-page, it is washed once in hot water (65 °C), submerged in the quick-stain solution in a beaker and heated in the staining solution (40 sec, ~ 90 °C).

The beaker with the submerged page is shaken for 15 min.

Heating and shaking cycles are repeated 2 – 3 times, until development of the bands is complete.

To decolor the blue background, the page is submerged in water (d.i.) and heated (20 sec, ~ 95 °C). Alternatively, the page is submerged in water (d.i.) at r.t. and shaken overnight.

5.2.5 Prometheus Differential Scanning Fluorometry

To determine protein melting points (T_m), protein samples (1 mg/mL, cell-free extract, or 0.5 mg/mL purified enzyme in a given buffer, or d.i. water) are filled into a glass capillary (10 µL). The filled capillary is placed on a sample holder. The excitation beam intensity is adjusted to fit the instruments threshold and then the fluorescence intensity is determined at 330 and 350 nm during a temperature ramp.

To determine a protein T_m, a constant temperature ramp of 1 °C/min (20 – 90 °C) is employed. Melting points are either determined as the inflection point of the F_{350}/F_{330} fluorescence/temperature curve, or if not determinable, of the F_{350}, or F_{330} curves.

To differentiate reversible and irreversible unfolding, alternating intervals of heating (7 °C/min, hold 1 min) and cooling (to 25 °C, hold 5 min) are applied with the peak temperature increasing for 1 °C each cycle, following a protocol from Svilenov *et al.*.[226] The first cycle heats to 26 °C and the last cycle heats to 95 °C peak temperature. The program yields two F_{350}/F_{330} fluorescence/temperature curves, one corresponding to the fluorescence at peak temperatures (reversible unfolding) and one corresponding to the fluorescence at 25 °C, in between the heating cycles (irreversible unfolding).

5.2.6 High Performance Liquid Chromatography (HPLC) Measurements

Reverse phase HPLC separations are carried out on a XBridge™ C18 analytical column (pore size: 3.5 µm, 3.0 · 150 mm) from Waters (Milford, USA) with gradient elution (see method).

Hydrophilic Interaction Liquid Chromatography (HILIC) is carried out on a XBridge™ Amide analytical column (pore size 3.5 µm, 4.6 · 250 mm) from Waters (Milford, USA),

Separation of enantiomers (chiral methods) is carried out on a Lux Cellulose-I analytical column (pore size: 5 µm, 250 · 4.6 mm) from Phenomenex (Torrace, USA).

The following separation methods are employed:

Method Reverse I: Flow rate: 0.5 mL/min, mobile phase A: 0.1% formic acid, B: 0.1% formic acid in ACN, column temperature 30 °C. For separation, the following linear gradient elution program is applied (% B): 0 – 2 min isocratic 5%, 2 - 26 min gradient to 95%, 26 – 28 min gradient to 5%, 28 – 30 min isocratic 5%.

Method Reverse II: Flow rate: 0.5 mL/min, mobile phase A: 0.1% formic acid, B: 0.1% formic acid in ACN, column temperature 30 °C. For separation, the following linear gradient elution program is applied (% B): 0 – 12 min isocratic 5%, 12 - 36 min gradient to 50%, 36 – 44 min gradient to 95%, 44 – 45 min gradient to 5%, 45 – 47 min isocratic 5%.

Method chiral reverse isocratic I: Flow rate: 1.0 mL/min, mobile phase A: 0.1% formic acid, B: 0.1% formic acid in ACN, column temperature 30 °C. For separation, the following isocratic elution program is applied (% B): 0 – 20 min 25%.

Method chiral reverse isocratic II: Flow rate: 1.0 mL/min, mobile phase A: 0.1% formic acid, B: 0.1% formic acid in ACN, column temperature 30 °C. For separation, the following isocratic elution program is applied (% B): 0 – 25 min 85%.

Method chiral reverse gradient: Flow rate: 0.5 mL/min, mobile phase A: 0.1% formic acid, B: 0.1% formic acid in ACN, column temperature 30 °C. For separation, the following linear gradient elution program is applied (% B): 0 – 16 min isocratic 5%, 16 - 17 min gradient to 30%, 17 – 24 min isocratic 30%, 24 – 25 min gradient to 5%, 25 – 27 min isocratic 5%.

Method HILIC I: Flow rate: 1.0 mL/min, mobile phase A: 7.6 mM NH_4HCO_2 and 0.15% formic acid in 87% ACN, 13% water, B: 10 mM NH_4HCO_2 and 0.15% formic acid in water, column temperature 35 °C. For separation, the following linear gradient elution program is applied (% B): 0 – 6 min isocratic 0%, 6 - 7 min gradient to 5%, 7 – 10 min gradient to 20%, 10 – 12 min gradient to 40%, 12 – 13 min gradient to 0%, 13 – 16 min isocratic 0%.

Method HILIC II: Flow rate: 0.5 mL/min, mobile phase A: 7.6 mM NH_4HCO_2 and 0.15% formic acid in 87% ACN, 13% water, B: 10 mM NH_4HCO_2 and 0.15% formic acid in water, column temperature 35 °C. For separation, the following linear gradient elution program is applied (% B): 0 – 10 min isocratic 0%, 10 - 12 min gradient to 5%,

12 – 18 min gradient to 20%, 18 – 22 min gradient to 40%, 22 – 24 min gradient to 0%, 24 – 30 min isocratic 0%.

5.3 ProRCMS-Panel Exploration

5.3.1 Two-Step Colorimetric Conversion Assay

The Screening of racemases was performed in two consecutive steps.

Step 1: Racemisation of L-amino acids

In an Eppendorf vial (1.5 mL), a solution of the respective amino acid (180 µL, see Figure 33, Figure 34, 11.1 mM, L-tryptophan and 4-benzoyl-L-phenylalanine as saturated solution) in potassium phosphate buffer (50 mM, pH 8.0) is added to a suspension of the respective amino acid racemase (20 µL, 40 mg/mL, ProRCMS#1 – 56, cell-free extract) in potassium phosphate buffer (50 mM, pH 8.0), containing pyridoxal-5'-phosphate (1 mM). This mixture is incubated (30 min, 30 °C, 600 rpm), then heated (6 min, 97 °C). The denatured enzyme is removed by centrifugation (6 min, 4 °C, 16100 *g*) and the supernatant employed in step 2.

Step 2: Coupled photometric assay

In a 96 well plate, an aliquot (50 µL) of the supernatant from step 1 is warmed (30 °C) and mixed with warm (30 °C), premixed colouring stock (50 µL), containing DAAO (either 4 mg/mL of Ni-NTA column chromatography purified, or 10 mg/mL dialysed cell-free extract), horseradish peroxidase (0.4 mg/mL, 220 U/mg, Thermo Fisher Scientific, lyophilized powder), vanillic acid (4.4 mM) and 4-aminoantipyrine (4.4 mM) in potassium phosphate buffer (50 mM, pH 7.5). The development of this mixture is continuously measured spectrophotometrically (60 min, every 7 sec, 498 nm).

5.3.2 β-Aspartame Conversion Experiments

In an Eppendorf vial (1.5 mL), a solution (180 µL) of methyl L-aspartyl-β-L-phenylalaninate (55.6 mM) in potassium phosphate buffer (50 mM, pH 7.5) is added to a solution (40 µL) of the respective racemase (20 mg/mL, ProRCMS#2, 5, 11, 28, 29, 33, 35, 36,) and PLP (0.5 mM) in the same buffer. The mixtures are incubated (18 h, 25 °C, 600 rpm) in duplicates.

The incubated reaction mixture is heated (3 min, 90 °C) and centrifuged (2 min, 4 °C, 16100 *g*). An aliquot (50 µL) of the supernatant is diluted (1:10) with water (d.i.) and an aliquot (200 µL) of the diluted solution is added *via* a syringe filter to 200 µL of ACN in an HPLC vial and subjected to HPLC/MS, employing commercial β-aspartame and the synthesized **3** (methyl D-aspartyl-β-L-phenylalaninate) as a reference standard.

Additionally, an aliquot (10 µL) of the sample is diluted (1:2) in water (d.i.). Sodium borate buffer (50 µL, 7 mM, pH 8.0) and FmocCl (75 µL, 60 mM in ACN) are added to the diluted sample. The mixture is incubated (5 min, 30 °C). L-Alanine (85 µL, 0.1 M, aq.) is added at 30 °C. The mixture is left to sit for 30 min at r.t. and applied to HPLC (method chiral normal isocratic II) after microfiltration

HPLC/MS (method chiral normal isocratic II, MS single ion measurement detection):

t_{ret}(methyl *N*$^{\alpha}$-(((*9H*-fluoren-9-yl)methoxy)carbonyl)-L-aspartyl-β-phenylalaninate) = 5.8 min (517.55 [M^+ + H], 539.55 [M + Na^+])

t_{ret}(methyl *N*$^{\alpha}$-(((9H-fluoren-9-yl)methoxy)carbonyl)-L-aspartyl-β-phenylalaninate) = 4.5 min (295 [M^+ + H], 293 [M^- – H]).

HPLC/MS (method chiral reverse gradient, MS detection):

t_{ret}(methyl L-aspartyl-β-L-phenylalaninate) = 12.7 min (295 [M^+ + H], 293 [M^- – H])

t_{ret}(**3**, methyl D-aspartyl-β-L-phenylalaninate) = 13.6 min (295 [M^+ + H], 293 [M^- – H]).

5.3.3 Screening for Serine Dehydratase Activity

In an Eppendorf vial (1.5 mL), a solution (200 µL) of L-serine (20 mM) and PLP (0.2 mM) in potassium phosphate buffer (50 mM, pH 7.5) is added to a solution (200 µL) of the respective racemase in the same buffer. The racemase concentration in the stem solution is adjusted according to the relative conversion in the substrate screening with ProRCMS#53 as the most active racemase, defined as 1 mg/mL. (ProRCMS#9, #10, #11, #12, #15, #46, #53 with 1.14, 1.41, 1.12, 1.54, 1.46, 1.34, 1.00 mg/mL, respectively). The mixtures are incubated (18 h, 25 °C, 600 rpm) in duplicates.

Potassium hydroxide (200 µL, 60% wt./V, aq.) is added to aliquots (200 µL) of each reaction mixture, left to sit for 10 min and centrifuged (5 min, 16100 *g*).

In a 96 well transparent microtiter plate, an ethanolic solution of salicylaldehyde (40 µL, 2% wt./V in ethanol, 96%) is added to an aliquot (160 µL) of each supernatant. Additionally, a duplicated calibration of pyruvate (0, 0.1, 0.2, 0.5, 1.0, 2.0, 5.0 mM) is treated in the same way as the reaction mixture and included in the microtiter plate. The plate is incubated (37 °C, 200 rpm, 35 min) and subsequently the absorption of individual wells is measured in a plate reader at 480 nm.

For values, exceeding the range of the linear calibration, the procedure is repeated with the alkaline, centrifuged mixture diluted (1:2) with potassium hydroxide (200 µL, 20% wt./V, aq.).

5.4 Racemase/Oxidase-Cascades Process Optimization

5.4.1 D-Amino acid oxidase Stabilizing Agents

5.4.1.1 Prometheus Melting Point Measurement

The following solutions were examined, employing the standard operation protocol 5.2.5, without temperature cycling.

DAAO from (131 µL, 7.6 mg/mL, *Rhodotorula gracilis*, cell-free extract) in bicine buffer (200 mM, pH 7.5) is added to different mixtures of water (119 µL, d.i.) and solutions of different stabilizing agents (250 µL, 200 mM of either NH_4SO_4, or glycyl betaine, or sorbitol, or urea). The solutions are compared with a solution of the same oxidase in water (d.i.) without additives.

5.4.1.2 Photometric Activity Assay

In a 96 well plate, of D-serine (100 µL, 20 mM) with different additives (none, or 200 mM sorbitol, or 200 mM glycyl betaine, or 200 mM sorbitol and 200 mM betain) in bicine buffer (50 mM, pH 7.5) are warmed to 30 °C and mixed with warm (30 °C), premixed colouring stock (80 µL), containing , horseradish peroxidase (1 mg/mL, Thermo Fisher Scientific, lyophilized powder, 220 U/mg), vanillic acid (15.5 mM), 4-aminoantipyrine (14.0 mM) in bicine buffer (50 mM, pH 7.5), and DAAO (20 µL, 30, or 10 mg/mL, cell-free extract) in bicine buffer (50 mM, pH 7.5) in triplicates. The colour change of this mixture is continuously measured spectrophotometrically (60 min, every 7 sec, 498 nm).

5.4.2 D-Amino acid oxidase Immobilization

The enzyme is immobilized, following the procedures from Guisán *et al.*. [254,255,256]

5.4.2.1 Step 1: Etherification with Glycidol

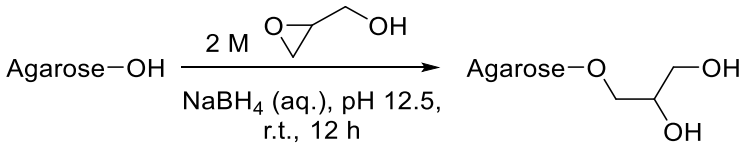


Figure 119: Etherification of SepharoseCL-6B with glycidol.

In a round bottom flask (50 mL) with a flying stirring bar, Sepharose CL-6B (15.0 mL, packed beads) is suspended in H_2O (4.0 mL, d.i.). Under gentle stirring, the pH is adjusted to 12.55 by dropwise addition of NaOH (1.0 M, aq.). The suspension is filled up to 23 mL with NaOH (0.16 M, aq.). $NaBH_4$ (138 mg, 3.65 mmol) and glycidol

(3.07 mL, 46.0 mmol, dropwise addition) are added under gentle stirring and the mixture is left to sit for 12 h at 25 °C.

The etherified beads are washed with H_2O (2 L, d.i.) until the filtrate becomes neutral.

5.4.2.2 Step 2: Oxidation of Glyceryl Groups

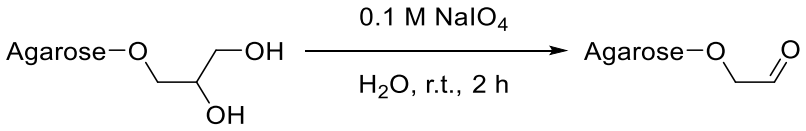


Figure 120: Oxidation of glyceryl groups.

In a round bottom flask (250 mL) with flying stirring bar, the washed beads from step 1 (15 mL packed beads) are suspended in water (150 mL, d.i.) and $NaIO_4$ (15 mL, 0.1 M, aq.) is added under stirring. The mixture is allowed to react for 2 h at 25 °C.

The etherified beads are washed with H_2O (2 L, d.i.).

5.4.2.3 Step 3: Preparation of Aminated Sepharose Beads

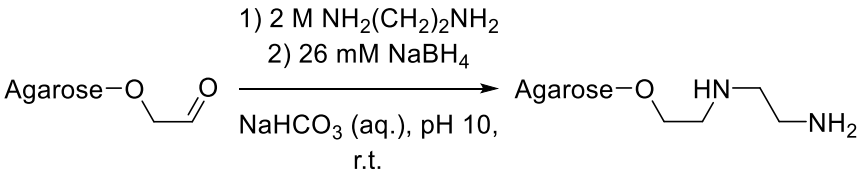


Figure 121: Two-step telescoped reaction sequence of imine-formation and reduction.

In a round bottom flask (250 mL) with flying stirring bar, oxidized beads from step 2 (10 mL, packed beads) are suspended in sodium bicarbonate buffer (100 mL, 100 mM, pH 10.0) and ethylene diamine (13.3 mL, 200 mmol) is added under stirring at r.t.. The suspension is stirred for 2 h and then, $NaBH_4$ (1.00 g, 26.4 mmol) is added under stirring. The suspension is allowed to sit for 2 h.

The aminated beads are subsequently washed with a solution of sodium acetate (0.1 M) and NaCl (1.0 M) in water (100 mL, pH 5.0), a solution of $NaHCO_3$ (0.1 M) and NaCl (1.0 M) in water (100 mL, pH 10.0), and water (500 mL, d.i.)

5.4.2.4 Step 4: Introduction of a Glyceraldehyde Spacer

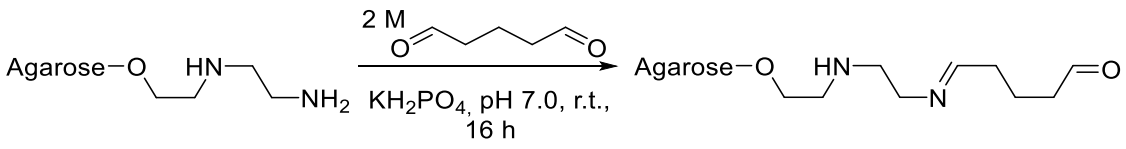


Figure 122: Imine formation between aminated beads and glutaraldehyde.

In a round bottom flask (50 mL) with flying stirring bar, aminated beads from step 3 (5 mL, packed beads) are suspended in potassium phosphate buffer (25 mL, 200 mM, pH 7.0). Glutaraldehyde (20 mL, 53 mmol, 25% wt./wt.) is added under stirring. The suspension is gently stirred overnight at r.t..

The beads are washed with water (500 mL, d.i.).

5.4.2.5 Step 5: Coimmobilization of D-Amino Acid Oxidase and Catalase

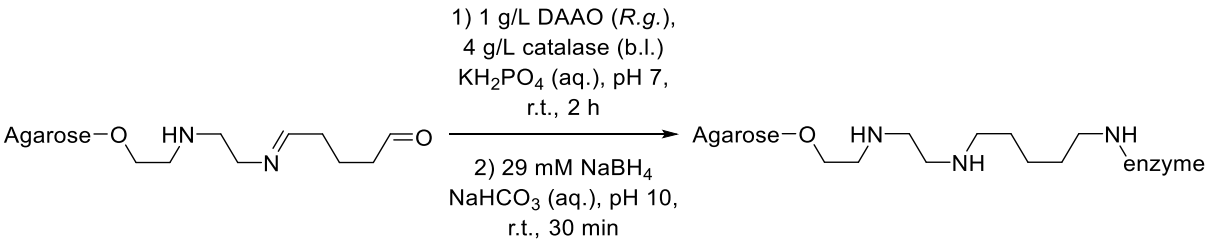


Figure 123: Coimmobilization of DAAO and catalase.

The beads from step 4 (5 mL, packed beads) are suspended in a round bottom flask (100 mL) in potassium phosphate buffer (15 mL, 20 mM, pH 7.0) immediately after washing. DAAO (15 mg, *Rhodotorula gracilis*, purified, lyophilized enzyme) and catalase (60 mg, 138 kU, bovine liver, Sigma-Aldrich) are added simultaneously and the mixture allowed to react under gentle stirring for 2 h at r.t.. Then $NaBH_4$ (30 mL, 1.3 mmol, 43 mM) in bicarbonate buffer (50 mM, pH 10) is added and the mixture stirred for 30 min.

The beads with immobilized enzyme are washed with water (2 L, d.i.), potassium phosphate buffer (100 mL, 50 mM, pH 7.5) and stored under the same buffer at 8 °C.

5.4.3 Evaluation of Enzyme Thermostability

5.4.3.1 Preincubation of Enzymes

Enzymes [packed beads of coimmobilized DAAO and catalase (0.2 mL), DAAO (0.4 mg/mL, cell-free extract) and catalase from bovine liver (1.6 mg/mL, 2296 U/mg, Sigma-Aldrich), or ProRCMS#31 (1 mg/mL, cell-free extract)] are incubated (600 rpm) at different temperatures (24, 30, 40, 50, 60, 68.5 °C) in potassium phosphate buffer (1 mL, 50 mM, pH 7.5) as triplicates.

5.4.3.2 Conversion of Serine

Aliquots (250 µL, homogeneous suspension in the case of immobilized enzymes) of each preincubated enzyme are mixed with either a solution (250 µL) of D-serine (100 mM) and catalase (0.4 mg/mL, 2297 U/mg, bovine liver, Sigma-Aldrich) in

potassium phosphate buffer (50 mM, pH 7.5) for preincubated oxidase preparations, or a solution (250 µL) of L-serine (100 mM), DAAO (1 mg/mL, *Rhodotorula gracilis*, cell-free extract), and catalase (4 mg/mL, 2297 U/mg, bovine liver, Sigma-Aldrich) in potassium phosphate buffer (50 mM, pH 7.5) for preincubated ProRCMS#31. The mixtures are incubated (30 min, 25 °C, 300 rpm) in open Eppendorf vials (1.5 mL).

5.4.3.3 Determination of Hydroxypyruvate with WST-1

The reaction mixtures are centrifuged (3 min, 4 °C, 16100 *g*) and an aliquot (20 µL) of each supernatant is added to solutions of WST-1 (2-(4-Iodophenyl)-3-(4-nitrophenyl)-5-(2,4-disulfophenyl)-2*H*-tetrazolium sodium salt, 180 µL, 124 µM) in NaOH (0.11 M, aq.) in the wells of a microtiter plate as duplicates. The microtiter plate is sealed with an acetate foil, incubated (30 min, 30 °C, 300 rpm) and the absorption measured in a microtiter plate reader (520 nm).

5.4.4 Comparative Measurement of Oxygen Concentrations

Five tubes (50 mL), equipped with a stirring bar, are filled with water (d.i.) to the 50 mL mark, closed air-tight and the caps are equipped with two cannulas each. Three of the tubes are each additionally provided with a submerged elastic tube (internal diameter 1.0, 1.0, 0.64 mm; wall thickness 0.30, 0.30, 0.28 mm; material fluoroethylene propylene, polytetrafluoroethylene, Silastic™, respectively) of 36 cm submerged length, leading through the cap and sealed at the submerged end. The tubes are sonicated for 8 min in a water bath at 30 °C, while bubbling the liquid inside the tubes with argon through one of the two cannulas. Directly after, cannulas leading to the argon balloon are pulled out of the liquid, into the headspace. For the fifth tube, an additional canula, leading to an oxygen balloon is stuck through the cap, into the liquid, bubbling it with oxygen. The three submerged flexible tubes are pressurized with 2 bars of oxygen. The liquid in the five tubes is held at 30 °C and stirred for 5 min. Flexible tubes are de-pressurized, disconnected from the oxygen bottle and the now loose ends sealed with parafilm. To each 50 mL tube, $MnCl_2$ (0.4 mL, 2.0 M, aq.) and NaI (0.4 mL, 4.0 M in 8.0 M NaOH, aq.) is added. To prevent excessive contamination with atmospheric oxygen, tubes are opened and closed gently and only as far as necessary. Also, the pipette tips are submerged 1 cm deep into the liquid. After closing the tubes, the small headspace is purged with argon. Then, all cannulas are removed and holes in the caps sealed with parafilm. The tubes are vigorously shaken and then, the solutions are stirred for 5 min. The precipitate is allowed to settle (5 min), then stirred

(1 min) and allowed to settle again (1 min). H_2SO_4 (0.4 mL, 8 M, aq.) is carefully added to each tube as described above. After stirring the closed tubes (1 min), aliquots (20 µL) of the formed solutions are diluted 1:10 with water (d.i.) in the wells of a microtiter plate as triplicates and the absorption of each well measured at 456 nm.

5.5 Application of Panel Racemases A: Transketolase Cascades

5.5.1 Heat Treatment Purification of Transketolase

5.5.1.1 General Procedure for Transketolase Purification

Transketolase is dissolved in triethanolamine buffer (20 mL, 5 mM, pH 7.0). The solution is distributed to 20 Eppendorf vials (1.5 mL) and incubated (30 min, 50 °C, 20 *g*).

The suspensions are allowed to cool to r.t., centrifuged (15 min, 16100 *g*) and the clear supernatant is lyophilized, yielding a colourless fluffy solid.

5.5.1.2 Purification of Wild-Type Transketolase

Transketolase (504 mg, *Geobacillus stearothermophilus*, wild-type, cell-free extract) is subjected to the general procedure 5.5.1.1, yielding a fluffy solid (336 mg, 66%).

5.5.1.3 Purification of Transketolase H102L/H474S Variant

Transketolase (483 mg, *Geobacillus stearothermophilus*, H102L/H474S variant, cell-free extract) is subjected to the general procedure 5.5.1.1, yielding a fluffy solid (280 mg, 58%).

5.5.2 General Procedures for the L-LDH Assays

5.5.2.1 Sample Preparation for L-LDH Assays

NaOH (20 µL, 0.3 M, aq.) is added to an aliquot (100 µL) of the unbuffered reaction mixture. The mixture is frozen in liquid nitrogen. Shortly before analysis, the sample is warmed to 25 °C, HCl (20 µL, 0.3 M, aq.), and potassium phosphate buffer (860 µL, 50 mM, pH 7.5) are added and the diluted (1:10) sample is centrifuged (3 min, 16100 *g*). For samples with prediluted analyte concentrations between 0 and 5 mM, the supernatant is directly applied in the assay, while samples with higher analyte concentrations (5 – 50 mM) are further diluted (1:10) with the same buffer.

5.5.2.2 Determination of Hydroxypyruvate (0.5 – 50 mM) with L-LDH

In the well of a multititer plate, duplicates of the diluted sample (50 µL, 1:10, or 1:100, see sample preparation) are mixed with L-LDH (10 µL, 6.9 – 10.3 U, of a 1:14 dilute of the commercial enzyme in the same buffer, L-lactate dehydrogenase from rabbit muscle type-II, 12 mg/mL ammonium sulphate suspension, Sigma-Aldrich). To start the reaction, NADH (40 µL, 1.75 mM in the same buffer) is added. The absorption is followed (60 min, every 8 sec, 340 nm). For determination of the concentration, the

absorption after the completed conversion of NADH is compared with a corresponding calibration.

5.5.2.3 Determination of D-Serine (0.5 – 50 mM) with L-LDH

To determine the total concentration of D-serine and hydroxypyruvate, the same method as in 5.5.2.2 is applied except, instead of adding only L-LDH, the same volume of a mixture, containing DAAO (94 µg, *Rhodotorula gracilis*, purified, lyophilised enzyme), catalase (33 µg, 76 U, bovine liver, Sigma-Aldrich), and L-LDH (6.9 – 10.3 U, L-lactate dehydrogenase from rabbit muscle type-II, 12 mg/mL ammonium sulphate suspension, Sigma-Aldrich) in the same buffer, is added. To calculate the D-serine concentration, the determined hydroxypyruvate concentration (5.5.2.2) is subtracted.

5.5.2.4 Determination of L-Serine (0.5 – 50 mM) with L-LDH

To determine the total concentration of D, L- serine, and hydroxypyruvate, the same method as in 5.5.2.2 is applied except, instead of adding only L-LDH, the same volume of a mixture, containing ProRCMS#9 (414 µg, purified *via* NiNTA), DAAO (375 µg, *Rhodotorula gracilis*, purified *via* NiNTA), catalase (33 µg, 76 U, bovine liver, Sigma-Aldrich), and L-LDH (6.9 – 10.3 U, L-lactate dehydrogenase from rabbit muscle type-II, 12 mg/mL ammonium sulphate suspension, Sigma-Aldrich) is added. To calculate the L-serine centration, the determined combined D-serine and hydroxypyruvate concentrations (5.5.2.3) are subtracted.

5.6 Deoxysugar Synthesis

5.6.1 General Procedure for Transketolase Reactions

5.6.1.1 Reaction Setup and Execution

A Falcon tube (50 mL) is painted black to exclude light. A 2 · 3 cm window is cut into the side of the tube, apprx. 1 cm below its thread for easy access. The window is kept closed and sealed with scotch tape.

A given amount of transketolase (*Geobacillus stearothermophilus*, variant given, purified, lyophilized enzyme) is weighted into the tube and slurried with a solution (5.0 mL) of ThDP (9.6 mM) and $MgCl_2$ (36 mM) in triethanolamine buffer (100 mM, pH 7.0) and incubated (20 min, 40 °C).

A given amount of amino acid racemase (exact enzyme given, cell-free extract), is slurried in a solution (2 mL) of PLP (1.0 mM) in the same buffer.

Catalase (50 µL, *Micrococcus lysodeicticus*, 20.2 kU, Sigma-Aldrich) is added to a solution (1.95 mL) of a given amount of DAAO (exact enzyme given, cell-free extract) and FAD (1.0 mM, disodium salt) in the same buffer.

The prepared racemase slurry and oxidase solution are added to the transketolase slurry, together with L-serine (volume given, 1.0 M) in the buffer and freshly deprotected and quantified acceptor aldehyde (volume and concentration given) in the buffer. The reaction mixture is filled up to 20 mL with the buffer.

The Falcon tube is closed with its lid, tilted to the side and secured into a crystallizing dish. The lid is punctured with two cannulas, of which both are kept above the suspension's surface and of which one is connected to a balloon filled with oxygen.

The crystallization dish, containing the Falcon tube is incubated (duration given, 25 °C, 600 rpm).

Samples (222 µL) are drawn and catalase (50 µL, *Micrococcus lysodeicticus*, 20.2 kU, Sigma-Aldrich) is added at 0, 30, 60, 120, 240, and 480 min. After 24 h, a final sample (222 µL) is drawn.

5.6.1.2 Sample Preparation

An aliquot (1 µL) of the sample is diluted with water (d.i., 1:1) and part (1 µL) of the dilution is applied on TLC (^{i}propanol/NH_3, 25%, aq., 3 + 2, ninhydrin reagent).

An aliquot (1 µL) of the sample is applied on TLC (ethyl acetate/^{n}butanol/^{i}propanol/H_2O/acetic acid, 7 + 1 + 1 + 1 + 1, anisaldehyde reagent).

An aliquot (10 µL) of the sample is diluted in water (1:40, d.i.) and alkalinized with NaOH (10 µL, 0.4 M, aq.). A solution of PMP (3-methyl-1-phenylpyrazolin-5-on, 10 mL, 0.5 M in methanol) is added, the sample is heated (60 min, 70 °C), neutralized with HCl (10 µL, 0.4 M, aq.), extracted with chloroform (400 µL) and applied to HPLC (method chiral reverse isocratic I) after microfiltration.

An aliquot (10 µL) of the sample is diluted in water (d.i., 1:5). Sodium borate buffer (50 µL, 7 mM, pH 8.0) and FmocCl (75 µL, 60 mM in ACN) are added to the diluted sample. The mixture is incubated (5 min, 30 °C). L-Alanine (85 µL, 0.1 M, aq.) is added at 30 °C. The mixture is left to sit for 30 min at r.t. and applied to HPLC (method chiral reverse isocratic II) after microfiltration.

The remaining sample (200 µL) is lyophilized. The lyophilized oil is dissolved in a solution (600 µL) of maleic acid (70.84 mM, *Trace*Cert®, standard for quantitative NMR, Sigma-Aldrich, in D_2O). A ^{1}H NMR (500 MHz, number of scans: 256, relaxation delay: 1.0000 s) is recorded.

Purification procedures vary and are given for individual products.

TLC (ethyl acetate/^{n}butanol/^{i}propanol/H_2O/acetic acid, 7 + 1 + 1 + 1 + 1, anisaldehyde reagent):

R_f(5-deoxyfructose) = 0.32 (brown or red),

R_f(1,5-dideoxyfructose) = 0.51 (red)

R_f(2,4-dihydroxybutanal) = 0.65 (black).

TLC (^{i}propanol/NH_3, 25%, aq., 3 + 2, ninhydrin reagent):

R_f(alanine) = 0.60 (red),

R_f(serine) = 0.44 (red),

R_f(glutamate) = 0.34 (red).

TLC (DCM/methanol/formic acid, 85 + 15 + 0.1, anisaldehyde reagent):

R_f(5-deoxyfructose) = 0.16 (brown or red),

R_f(2,4-dihydroxybutanal) = 0.40 (black).

TLC (DCM/methanol/formic acid, 9 + 1 + 0.1, anisaldehyde reagent):

R_f(1,5-dideoxyfructose) = 0.25 (red),

R_f(2(*R*),4-dihydroxybutanal) = 0.39 (black).

HPLC/UV/Vis (method chiral reverse isocratic I, 270 nm)

t_{ret}(PMP-derivatized (*R*)-2,4-dihydroxybutanal) = 14.6 min,

t_{ret}(PMP-derivatized (*S*)-2,4-dihydroxybutanal) = 16.0 min.

HPLC/MS (method reverse I, 254 nm, mass spectrometric)

t_{ret}((((*9H*-fluoren-9-yl)methoxy)carbonyl)serine) = 18.7 min (M^--H =326, M+$HCOO^-$ = 372).

t_{ret}((((*9H*-fluoren-9-yl)methoxy)carbonyl)alanine) = 19.3 min (M^--H =310, M+$HCOO^-$ = 356).

HPLC/UV/Vis/MS (method chiral reverse isocratic II, 254 nm)

t_{ret}((((*9H*-fluoren-9-yl)methoxy)carbonyl)-L-serine) = 5.3 min (326 [M – H^+]),

t_{ret}((((*9H*-fluoren-9-yl)methoxy)carbonyl)-D-serine) = 4.6 min (326 [M – H^+]).

5.6.2 Cascade a: 5-Deoxy-D-threo-hex-2-ulose (5-Deoxyfructose)

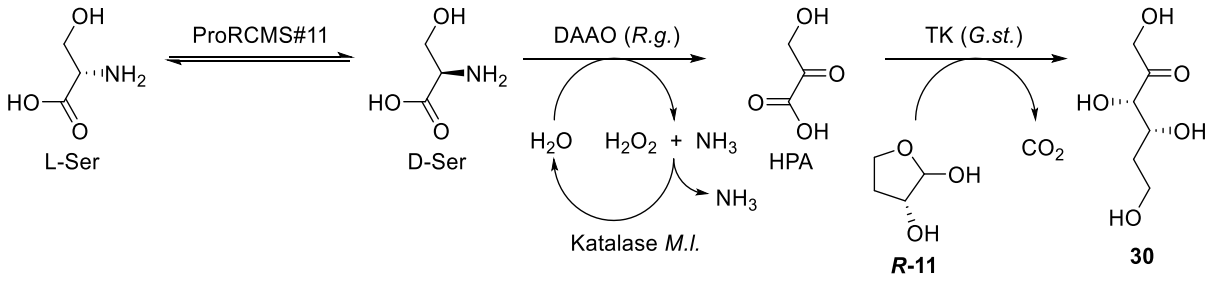


Figure 124: Cascade reaction to produce **30** from ***R*-11**, employing ProRCMS#11, DAAO, and TK.

Amounts of enzymes: 160 mg transketolase (*Geobacillus stearothermophilus*, wild-type), 79.9 mg ProRCMS#11, 40.0 mg DAAO (*Rhodotorula gracilis*, wild-type).

Amounts of substrates: L-serine (2.80 mL, 1.00 M, 1.27 eq.) in the buffer, ***R*-11** (2(*R*),4-dihydroxybutanal, 3.67 mL, 600 mM, 1.00 eq.) in the buffer.

Purification: The reaction mixture is extracted with 20 mL of chloroform. The aqueous layer is lyophilized. The lyophilized oil is dissolved in a mixture of methanol (10 mL) and water (1 mL) and subjected to flash chromatography (100 g silica gel, 0.035 – 0,07 mm, diameter 2.5 cm, length 35 cm, 800 mL DCM/methanol/formic acid 0.85 + 0.15 + 0,01, then 500 mL 0.75 + 0.25 + 0.01, 20 mL fractions collected).

Solvent is removed under reduced pressure. The liquid residue is diluted with water (15 mL, d.i.) and lyophilized to yield **30** (5-deoxy-D-threo-hex-2-ulose) as a slightly yellow glassy solid (386 mg, quant., fractions 35 – 57).

^{1}H NMR: (500 MHz, D_2O): 3.93 – 3.83 (m, 2H, C6H_2O), 3.74 – 3.64 (m, 1H, C4HOH), 3.68 (d, 1H, $^{2}J_H$ = 11.7 Hz, C1H_2OH), 3.46 (d, 1H, $^{2}J_H$ = 11.7 Hz, C1H_2OH), 3.42 (d, 1H, $^{2}J_H$ = 9.5 Hz, C3HOH), 2.0 – 1.92 (dm, 1H, $^{2}J_H$ = 12.8, C5H_2, equatorial), 1.63 ppm (dddd, 1H, $^{2}J_H$ = 12.8 Hz, $^{3}J_H$ ≈ 12 Hz, $^{3}J_H$ = 5.4 Hz , $^{2}J_H$ ≈ 1.5 Hz, C5H_2, axial).

^{13}C NMR: (126 MHz, D_2O): 100.33 (minor, C2HOOH), 97.92 (major, C2HOOH), 72.67 (minor, C4HOH) 71.93 (major, C4HOH), 68.03 (major, C3HOH), 67.97 (minor, C3HOH), 63.39 (major, C1H_2OH), 60.14 (minor, C1H_2OH), 58.96 (minor, C6H_2O), 58.44 (major, C6H_2O), 32.52 (major, C5H_2), 32.34 ppm (minor, C5H_2).

5.6.3 **Cascade b: 5-Deoxy-D-threo-hex-2-ulose (5-Deoxyfructose)**

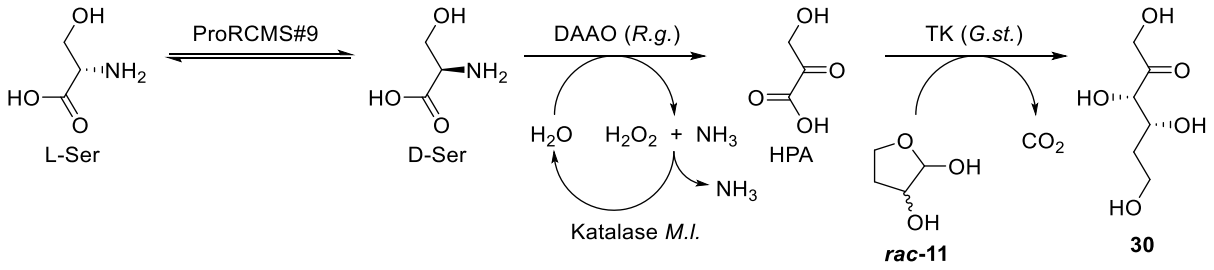


Figure 125: Cascade reaction to produce **30** from ***rac*-11**, employing ProRCMS#9, DAAO, and TK.

Amounts of enzymes: 160 mg transketolase (*Geobacillus stearothermophilus*, wild-type), 79.9 mg ProRCMS#9, 40.0 mg DAAO (*Rhodotorula gracilis*, wild-type).

Amounts of substrates: L-serine (2.80 mL, 1.00 M, 1.40 eq.) in the buffer, ***rac*-11** (*rac*-2,4-dihydroxybutanal, 2.67 mL, 1.50 M, 4.00 mmol) in the buffer.

Purification: The reaction mixture is extracted with chloroform (20 mL). The aqueous layer is lyophilized. The lyophilized oil is dissolved in a mixture of methanol (10 mL)

and water (1 mL) and subjected to flash chromatography (100 g silica gel, 0.035 – 0,07 mm, diameter 2.5 cm, length 35 cm, 600 mL DCM/methanol/formic acid 0.85 + 0.15 + 0,01, then 300 mL DCM/methanol/formic acid 0.75 + 0.25 + 0.01, 20 mL fractions collected).

Solvent is removed under reduced pressure. The liquid residue is diluted with water (15 mL, d.i.) and lyophilized to yield **30** (5-deoxy-D-threo-hex-2-ulose) as a brownish glassy solid (355 mg, quant., fractions 26 – 44).

^{1}H NMR: (500 MHz, D_2O): 3.97 – 3.86 (m, 2H, C6H_2O), 3.78 – 3.70 (m, 1H, C4HOH), 3.72 (d, 1H, $^{2}J_H$ = 11.8 Hz, C1H_2OH), 3.49 (d, 1H, $^{2}J_H$ = 11.6 Hz, C1H_2OH), 3.45 (d, 1H, $^{2}J_H$ = 9.5 Hz, C3HOH), 2.02 – 1.92 (dm, 1H, $^{2}J_H$ = 12.5, C5H_2, equatorial), 1.73 – 1.59 ppm (ddm, 1H, $^{2}J_H$ = 12.5 Hz, $^{3}J_H$ = 5.5 Hz, C5H_2, axial).

^{13}C NMR: (126 MHz, D_2O): 100.32 (minor, C2HOOH), 97.91 (major, C2HOOH), 72.68 (minor, C4HOH) 71.93 (major, C4HOH), 68.03 (major, C3HOH), 67.97 (minor, C3HOH), 63.40 (major, C1H_2OH), 60.14 (minor, C1H_2OH), 58.96 (minor, C6H_2O), 58.44 (major, C6H_2O), 32.52 (major, C5H_2), 32.34 ppm (minor, C5H_2).

5.6.4 **Cascade c: 5-Deoxy-D-threo-hex-2-ulose (5-Deoxyfructose)**

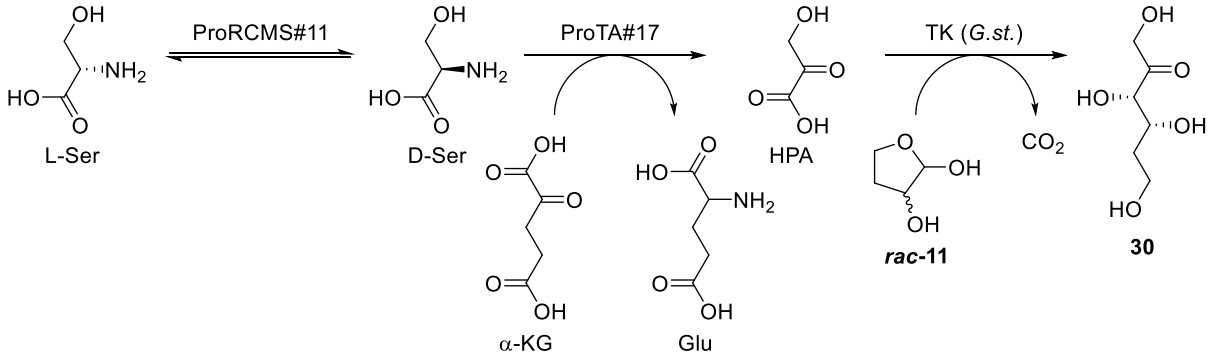


Figure 126: Cascade reaction to produce **30** from ***rac*-11**, employing ProRCMS#11, ProTA#17, and TK.

The synthesis follows the general procedure, except:

No solution of D-AAO, catalase and FAD is added. Instead, the amino acid racemase is weighed in together with the transaminase and slurried in a solution of PLP (4 mL, 0.5 mM) in the buffer.

No oxygen is supplied to the reaction, the vial is shut during the reaction.

α-Ketoglutarate is added to the reaction mixture.

Amounts of enzymes: 160 mg transketolase (*Geobacillus stearothermophilus*, wild-type), 79.9 mg ProRCMS#9, 200.0 mg transaminase (ProTA#17, cell-free extract).

Amounts of substrates: L-serine (2.80 mL, 1.00 M, 1.40 eq.) in the buffer, ***rac*-11** (*rac*-2,4-dihydroxybutanal, 2.67 mL, 1.50 M, 4.00 mmol) in the buffer, α-ketoglutarate (646 mg, disodium salt, 98%).

Purification: The reaction mixture is extracted with 20 mL of chloroform. The aqueous layer is lyophilized. The lyophilized oil is dissolved in a mixture of methanol (10 mL) and water (1 mL) and subjected to flash chromatography (100 g silica gel, 0.035 µm – 0,07 mm, diameter 2.5 cm, length 35 cm, 1 L DCM/methanol/formic acid 0.85 + 0.15 + 0,01, then 300 mL 0.75 + 0.25 + 0.01, 25 mL fractions collected).

Solvent is removed under reduced pressure. The liquid residue is diluted with water (15 mL, d.i.) and lyophilized to yield **30** (5-deoxy-D-threo-hex-2-ulose) as a brownish glassy solid (230 mg, 77%, fractions 41 – 56).

^{1}H NMR: (500 MHz, D_2O): 3.98 – 3.86 (m, 2H, C6H_2O), 3.77 – 3.68 (m, 1H, C4HOH), 3.70 (d, 1H, $^{2}J_H$ = 11.6 Hz, C1H_2OH), 3.48 (d, 1H, $^{2}J_H$ = 11.8 Hz, C1H_2OH), 3.44 (d, 1H, $^{2}J_H$ = 9.5 Hz, C3HOH), 2.02 – 1.93 (dm, 1H, $^{2}J_H$ = 12.8, C5H_2, equatorial), 1.72 – 1.58 ppm (ddm, 1H, $^{2}J_H$ = 12.8 Hz, $^{3}J_H$ = 5.5 Hz, C5H_2, axial).

^{13}C NMR: (126 MHz, D_2O): 100.33 (minor, C2HOOH), 97.92 (major, C2HOOH), 72.68 (minor, C4HOH) 71.94 (major, C4HOH), 68.03 (major, C3HOH), 67.98 (minor, C3HOH), 63.40 (major, C1H_2OH), 60.15 (minor, C1H_2OH), 58.96 (minor, C6H_2O), 58.44 (major, C6H_2O), 32.52 (major, C5H_2), 32.34 ppm (minor, C5H_2).

5.6.5 Cascade d: 1,5-Dideoxy-D-threo-hex-2-ulose (1,5-Dideoxyfructose)

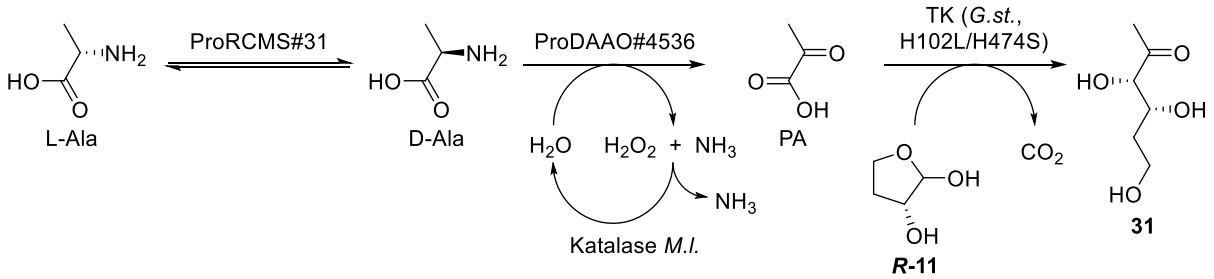


Figure 127: Cascade reaction to produce **31** from ***R*-11**, employing ProRCMS#31, ProDAAO#4536, and TK.

The synthesis follows the general procedure, except:

Instead of a solution of L-serine, a solution of L-alanine (1.0 M) in the buffer is added to the enzyme mixture.

The reaction mixture is incubated (54 h, 45 °C). After 47 h, another portion of L-alanine (36 mg, 40.4 mmol, 0.20 eq., 808 mM) in 0.5 mL buffer is added.

Amounts of enzymes: 160 mg transketolase (H102L/H747S variant), 80 mg ProRCMS#31, 80 mg ProDAAO#4536.

Amounts of substrates: L-alanine (2.20 mL, 1.00 M, 1.10 eq.), ***R*-11** (2(*R*),4-dihydroxybutanal, 3.33 mL, 600 mM, 1.00 eq.).

Purification: The reaction mixture is extracted with 20 mL of chloroform. The aqueous layer is lyophilized. The lyophilized oil is dissolved in a mixture of methanol (10 mL) and water (1 mL) and subjected to flash chromatography (130 g silica gel, 0.035 – 0,07 mm, diameter 2.5 cm, length 35 cm, 2 L DCM/methanol/formic acid 9.75 + 0.25 + 0,01, then 1 L 9 + 1 + 0.1, 25 mL fractions collected).

Solvent is removed under reduced pressure. The liquid residue is diluted with water (15 mL, d.i.) and lyophilized to yield **31** (1,5-dideoxy-D-threo-hex-2-ulose) as a brown sticky mass (149 mg, 55.0%, fractions 93 – 108).

^{1}H NMR: (500 MHz, D_2O, CD_3OD): 3.85 – 3.79 (m, 2H, C4HOH, C6H_2OH-a), 3.65 (dd, 1H, $^{2}J_H$ = 11.4, $^{3}J_H$ = 5.6 Hz, C6H_2OH-b), 2.96 (d, 1H, $^{3}J_H$ = 9.1 Hz, C3HOH), 1.86 (dd, 1H, $^{2}J_H$ = 12.5, $^{3}J_H$ = 5.3 Hz, C5H_2-eq.), 1.57 (ddd, 1H, $^{2}J_H$ = 12.6, $^{3}J_H$ = 12.2, $^{3}J_H$ = 5.0 Hz, C5H_2-ax.), 1.39 ppm (s, 3H, C1H_3).

^{13}C NMR: (126 MHz, D_2O, CD_3OD, major isomer): 98.9 (C2HOOH), 79.3 (C3HOH) 69.9 (C4HOH), 59.3 (C6H_2OH), 35.2 (C5H_2), 26.2 (C1H_3).

5.7 Application of Panel Racemases B: Preparation of α-Keto Acids

5.7.1 General Procedure for the Synthesis of 2-Oxo acids

A 2*3 cm window is cut into the side of a Falcon tube, apprx. 1 cm below its thread for easy access and painted black to exclude light. The window is kept closed and sealed with scotch tape.

Given amounts of DAAO (*Rhodotorula gracilis*, cell-free extract), catalase (2297 U/mg, bovine liver, Sigma-Aldrich), amino acid racemase (exact enzyme given, cell-free extract), FAD and PLP are weighed into the Falcon tube and suspended with 11 mL potassium phosphate buffer (25 mM, pH 7.3).

In another tube, given amounts of L-amino acid are dissolved in 5 mL potassium phosphate buffer (25 mM, pH 7.3), the pH adjusted to 7.3 with KOH (5 M, aq.) and made up to 9 mL with the buffer. This solution is added to the enzyme suspension.

The Falcon tube is closed, tilted to the side to increase the suspension's surface, and taped into a crystallizing dish. The tube's lid is punctured with two cannulas, of which both are kept above the suspension's surface and of which one is connected to a balloon filled with oxygen.

The crystallization dish, containing the yellow reaction suspension is incubated (duration given, 25 °C, 700 rpm).

Purification procedures varied and are given for individual products.

5.7.2 3-Methyl-2-oxobutyric acid

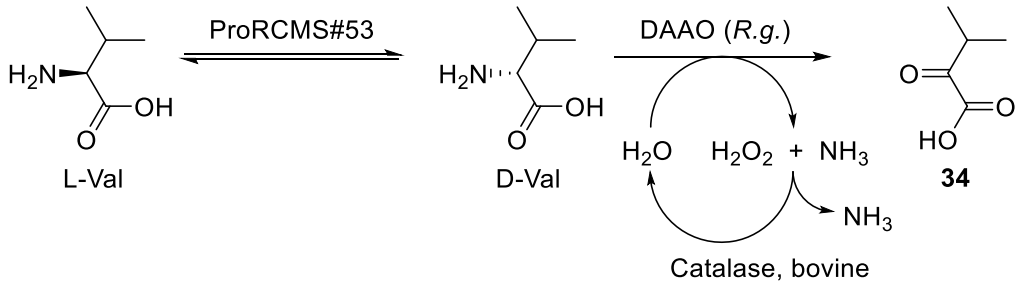


Figure 128: Cascade reaction to produce **34** from L-Val, employing ProRCMS#53 and DAAO.

The synthesis follows the general procedure, except:

The reaction tube is not painted black.

After 22 h, another portion of DAAO (4 mg, *Rhodotorula gracilis*, purified, lyophilized enzyme) and catalase (12 mg, 27.4 kU, bovine liver, Sigma-Aldrich) is dissolved in

potassium phosphate buffer (200 µL, 50 mM, pH 7.5) and added to the reaction mixture.

L-Valine (471 mg, 4.02 mmol) is employed at 200 mM concentration.

Amounts of enzymes: 10.3 mg DAAO, 31.4 mg of catalase, and 79.5 mg ProRCMS#53, 0.6 mg FAD (0.03 mM), 2.1 mg PLP (0.39 mM).

Purification: Another 19 h later, the reaction is stopped by addition of ethanol (15 mL) and 2 cycles of freezing and thawing.

The mixture is centrifuged (12 min, 3100 *g*), the clear supernatant collected, the pellet suspended with ethanol (20 mL) and centrifuged again.

United solutions are reduced to 5 mL at reduced pressure, acidified (pH 1) with HCl (2 M, aq.), extracted with Et_2O (3*25 mL), and dried with Na_2SO_4.

Removal of solvents yields **34** (3-methyl-2-oxobutyric acid) as a colourless, clear oil (471 mg, 67%, 67% purity, impurities: ether, ethanol).

TLC (*i*propanol/NH_3, 25%, aq., 3 + 2, ninhydrin reagent):

R_f(phenylalanine) = 0.14 (red).

TLC (ethyl acetate/*n*butanol/*i*propanol/H_2O/acetic acid, 7 + 1 + 1 + 1 + 1, Seebach's reagent):

R_f(3-methyl-2-oxobutyric acid) = 0.64 (blue).

^{1}H NMR: (300 MHz, acetone-d^6) 3.27 (hept., 1H, ^3J_H = 7.0 Hz, CH), 1.14 (s, 3H, CH_3), 1.11 ppm (s, 3H, CH_3).

^{13}C NMR: (75 MHz, acetone-d^6) 200.0 (CO), 163.2 (COOH), 37.1 (CH), 17.5 ppm (CH_3).

5.7.3 2-Oxo-3-phenylpropionic acid

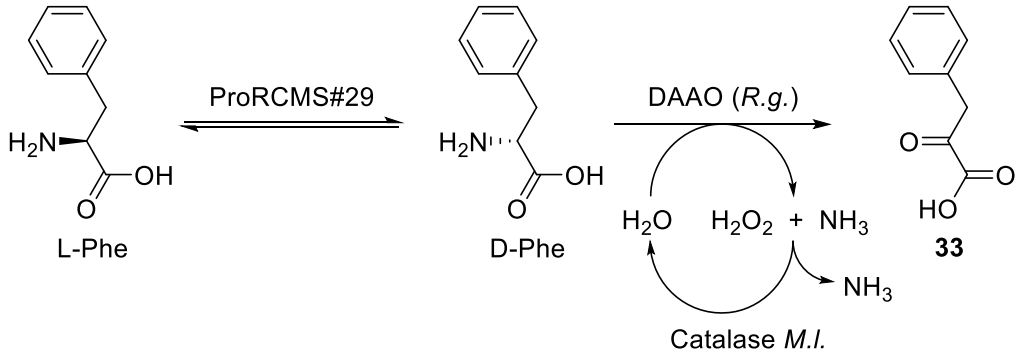


Figure 129: Cascade reaction to produce **33** from L-Phe, employing ProRCMS#29 and DAAO.

The synthesis follows the general procedure, except:

20 mL of a solution of L-phenylalanine (336.1 mg, 2,03 mmol, 101.7 mM) in potassium phosphate buffer (25 mM, pH 7.5) and 200 µL catalase (*Micrococcus lysodeicticus*, 20.2 kU, Sigma-Aldrich) are added to the dry enzymes.

Amounts of enzymes: 30.6 mg DAAO, 11.7 mg catalase, 22.8 mg ProRCMS#29, 1.8 mg FAD (0.10 mM), 1.5 mg PLP (0.28 mM).

Purification: After 17 h, the reaction is stopped by the addition of HCl (3.2 mL, 2 M, aq.). The suspension is centrifuged (12 min, 3100 *g*), the aqueous phase is removed, the pellet suspended with Et_2O (10 mL) and water (10 mL, d.i.) and centrifuged again.

United aqueous phases are extracted with Et_2O (6*20 mL). Elimination of persistent suspensions is eased by saturation with Na_2SO_4 and centrifugation.

United etheric phases are washed with Na_2SO_4 (sat., aq., 1*20 mL) and dried with Na_2SO_4 (anh.).

Removal of the solvent yields **33** (phenylpyruvate) as a colourless amorphous solid (412 mg, quant.).

TLC (^{i}propanol/NH_3, 25%, aq., 3 + 2, ninhydrin reagent):

R_f(phenylalanine) = 0.42 (red).

TLC (ethyl acetate/^{n}butanol/^{i}propanol/H_2O/acetic acid, 7 + 1 + 1 + 1 + 1, Seebach's reagent):

R_f(2-oxo-3-phenylpropionic acid) = 0.87 (blue).

^{1}H NMR (300 MHz, acetone-d^{6}): 7.84 (m, 2H, 2 CH enol arom., *o*), 7.32 (m, 8H, 8 CH arom., 3*enol, *m*, *p*, 5*ketone), 6.56 (s, 1H, CHOH), 4.21 ppm (s, 2H, CH_2).

^{13}C NMR (300 MHz, acetone-d^{6}): 193.5 (CO), 167.1 (COOH, enol), 162.4 (COOH, ketone), 141.5 (CHOH), 135.8 (C, arom., *I*, enol), 134.1 (C, arom., *I*, ketone), 130.8 (2CH, arom., *o*, ketone), 130.6 (2CH, arom., *o*, enol), 129.3 (2CH, arom., *m*, ketone), 129.2 (2CH, arom., *m*, enol), 128.4 (CH, arom., *p*, enol), 127.9 (CH, arom., *p*, ketone), 111.1 (CH, enol), 45.5 ppm (CH_2, ketone).

5.7.4 2-Oxo-glutaric acid (α-Ketoglutaric acid)

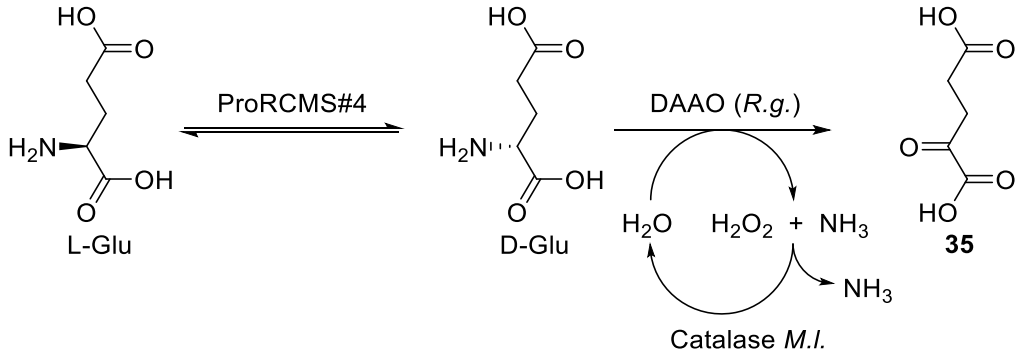


Figure 130: Cascade reaction to produce **35** from L-Glu, employing ProRCMS#4 and DAAO.

The synthesis follows the general procedure, except:

Enzymes (ProRCMS#4 and DAAO) are dissolved in the buffer (8.5 mL), containing equivalent amounts of coenzymes (FAD and PLP).

L-Glutamic acid (736 mg; 5.00 mmol) is slurried in the same buffer (10.5 mL) and added to the enzymes, together with catalase (200 µL, 20.2 kU, *Micrococcus lysodeicticus*, Sigma-Aldrich).

The mixture is filled up to 20 mL with buffer.

Amounts of enzymes: 120 mg DAAO, 200 mg ProRCMS#4, 1.8 mg FAD (0.10 mM), 1.5 mg PLP (0.28 mM).

Purification: After 17 h, the reaction is stopped by the addition of chloroform (20 mL) and vigorous mixing. The mixture is centrifuged (12 min, 3100 *g*) and the aqueous phase removed. Water (5 mL, d.i.) is added to the organic phase and pellet, mixed and centrifuged again.

United aqueous phases are lyophilized and subjected to flash chromatography (75 g silica gel, 40 – 200 µm, diameter 4.4 cm, length 12 cm, 400 mL chloroform/formic acid 100 + 1, then 200 mL 9 + 1 + 0,1, then 200 mL 8 + 2 + 0,1, then 200 mL 7 + 3 + 0,1, then 300 mL ethyl acetate/methanol/formic acid 6 + 4 + 0,1, 20 mL fractions collected).

Removal of the solvent yields **35** (2-oxo glutaric acid) as a colourless amorphous solid (224 mg, < 31%, fractions 51 – 67).

TLC (*i*propanol/NH_3, 25%, aq., 3 + 2, ninhydrin reagent):

R_f(glutamate) = 0.33 (red).

TLC (ethyl acetate/*n*butanol/*i*propanol/H_2O/acetic acid, 7 + 1 + 1 + 1 + 1, Seebach's reagent):

R_f(2-oxo-glutaric acid) = 0.05 (blue).

^1H NMR (300 MHz, methanol-d^4): 2.62 (t, 2H, CH_2), 2.43 ppm (t, 2H, CH_2).

5.8 Application of Panel Racemases C: Preparation of α-Hydroxy acids

5.8.1 General Procedure for the Synthesis of D-2-Hydroxy acids

A 2*3 cm window is cut into the side of a Falcon tube, apprx. 1 cm below its thread for easy access and painted black to exclude light. The window is kept closed and sealed with scotch tape.

Given amounts of D-hydroxyisocaproic acid dehydrogenase (D-HicDH, *Lactobacillus paracasei*, cell-free extract), DAAO (*Rhodotorula* gracilis, cell-free extract), amino acid racemase (exact enzyme given, cell-free extract), FDH (*Candida bodinii*, Boehringer Mannheim), FAD and PLP are weighed into the Falcon tube and suspended with water (8 mL, d.i.).

In another tube, given amounts of NH_4HCO_2 and NADH are dissolved in water (10 mL d.i.), the pH adjusted to 7.2 with KOH (5 M, aq.) and the tube made up to 11.9 mL with water (d.i.). This solution is added to the enzyme suspension.

A given amount of L-amino acid and catalase (100 µL, 10.1 kU, *Micrococcus lysodeicticus*, Sigma-Aldrich) are added to the suspended enzymes.

The Falcon tube is closed, tilted to the side to increase the suspensions surface, and taped into a crystallizing dish. The tube's lid is punctured with two cannulas, of which both are kept above the suspension's surface and of which one is connected to a balloon filled with oxygen.

The crystallization dish, containing the yellow reaction suspension is incubated (duration given, 25 °C, 700 rpm).

After 18 h, the reaction is stopped by the addition of chloroform (20 mL) and mixing. The emulsion is centrifuged (15 min, 3100 *g*) and the aqueous phase removed. The remains are extracted with water (5 mL, d.i.) and centrifuged again.

Purification procedures varied and are given for individual products.

5.8.2 D-2-Hydroxy-3-(4-hydroxyphenyl)propionic acid

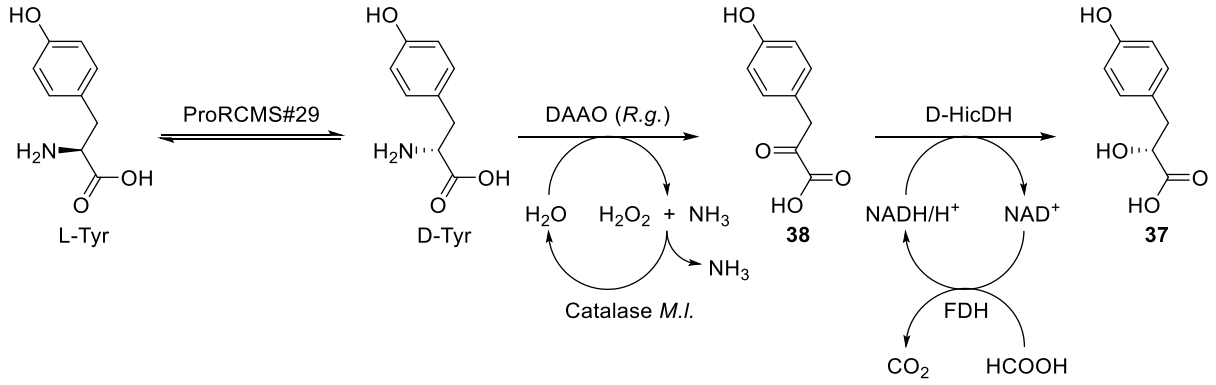


Figure 131: Cascade reaction to produce **37** from L-Tyr, employing ProRCMS#29, DAAO, and D-HicDH.

The synthesis follows the general procedure.

Amounts of enzymes: 112 mg D-HicDH, 208 mg DAAO, 140 mg ProRCMS#29, 13 mg FDH (6.5 U, *Candida bodinii*, Boehringer Mannheim), FAD (1.7 mg, 96 µM), 1.3 mg PLP (0.25 mM).

Amounts of Substrates: 181 mg L-tyrosine (49.9 mM), 354 mg NH_4HCO_2 (281 mM, 5.63 eq.), 66.4 mg NADH (4.45 mM, 0.089 eq.).

Purification: Apprx. 5 g of silica gel (60 Å, 70 – 200 µm) are added to the clear yellow solution, water is removed and then the dry yellow powder is loaded onto a silica column to carry out flash chromatography (50 g silica gel, 60 Å, 70 – 200 µm, diameter 4.5 cm, length 12 cm, 400 mL ethyl acetate/methanol/formic acid 9+1+0.1, 20 mL fractions collected).

The eluent is removed. Formic acid is removed from the residue by addition of water (10 mL, d.i.) and lyophilization, yielding **37** (D-2-hydroxy-3-(4-hydroxyphenyl)propionic acid) as a colourless fluffy solid (129 mg, 70%, fractions 8 – 15).

^{1}H NMR: (500 MHz, MeOD-d^4) 7.08 (d, 2H, ^3J_H = 8.4 Hz, 2 CH arom.), 6.70 (d, 2H, ^3J_H = 8.5 Hz, 2 CH arom.), 4.26 (dd, 1H, ^3J_H = 7.8 Hz, ^3J_H = 4.4 Hz, CHOH), 2.99 (dd, 1H, ^3J_H = 4.4 Hz ^2J_H = 14.0 Hz, CH_2), 2.81 ppm (dd, 1H, ^3J_H = 7.8 Hz, ^2J_H = 14.0 Hz, CH_2).

^{13}C NMR: (500 MHz, MeOD-d^4) 177.4 (COOH), 157.0 (COH, arom., *p*), 131.5 (2 CH, arom, *o*), 129.6 (C, arom., *i*), 116.0 (2 CH, arom., *m*), 73.1 (CHOH), 40.8 ppm (CH_2).

5.8.3 D-2-Hydroxy-3-phenylpropionic acid

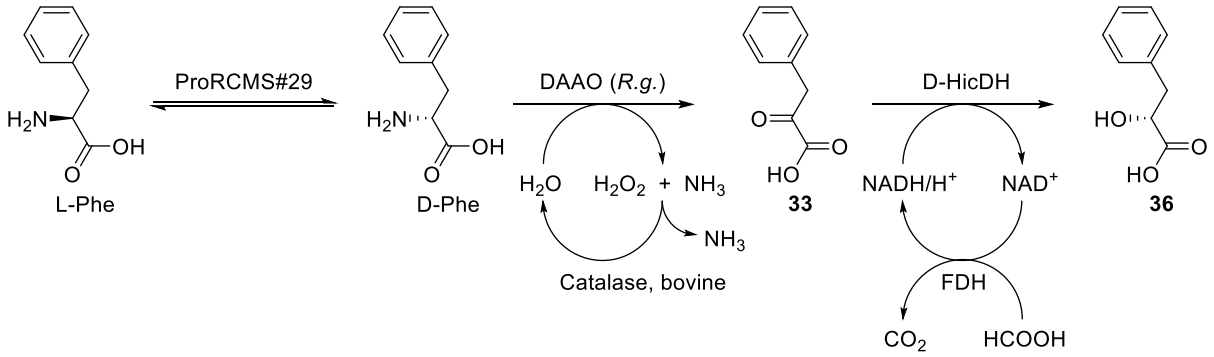


Figure 132: Cascade reaction to produce **36** from L-Phe, employing ProRCMS#29, DAAO, and D-HicDH.

The synthesis follows the general procedure, except:

Phenylalanine is added together with NADH and NH_4HCO_2 as a stem solution, containing L-phenylalanine (878.9 mg, 5.32 mmol), NH_4HCO_2 (785.9 mg, 12.46 mmol, 2.3 eq.) and NADH (35.5 mg, 47.5 µmol, disodium salt, 95%, 1 mol% eq.) in water (40 mL, d.i.). pH is adjusted to 7.5 with KOH (5 M, aq.)].

To the dry enzymes, 14.3 mg of catalase (bovine liver, 32.6 kU, lyophilisate, Sigma-Aldrich) is added.

The substrate solution is directly added to the dry enzymes and then the mixture is made up to 20 mL with d.i. water.

The mixture is given 21 h to react.

Instead of precipitating the enzymes with chloroform, the reaction mixture (pH 8) is acidified by addition of 5.5 mL of HCl (2 M, aq.).

After centrifugation (12 min, 3100 *g*), the solution is collected, the pellet suspended with Et_2O (20 mL) and water (1 mL, d.i.) and centrifuged again.

Amounts of enzymes: 49.4 mg D-HicDH, 41.3 mg DAAO, 23.1 mg ProRCMS#29, 1.2 mg FDH (0.6 U, *Candida bodinii*, Boehringer Mannheim), 1.0 mg FAD (57 µM), 1.2 mg PLP (0.22 mM).

Amounts of Substrates: 16 mL of the substrate solution, containing L-phenylalanine (133 mM), NH_4HCO_2 (312 mM, 2.34 eq.), NADH (476 µM, 8.95 meq.).

Purification: United aqueous phases are extracted with Et_2O (7*20 mL). Elimination of persistent suspensions is facilitated by saturation with Na_2SO_4. United etheric phases are washed with Na_2SO_4 (1*20 mL, sat., aq.) and dried with Na_2SO_4 (anh.).

Removal of the solvent yields **36** (D-2-hydroxy-3-phenylpropionic acid) as colourless needles (388 mg, quant.).

^{1}H NMR: (300 MHz, acetone-d^6) 7.49 – 7.00 (m, 5H, arom.), 4.39 (dd, 1H, ^3J_H= 7.8 Hz, ^3J_H = 1.0 Hz, CHOH), 3.12 (dd, 1H, ^3J_H = 4.2 Hz, ^2J_H = 13.9 Hz, CH_2), 2.91 ppm (dd, 1H, ^3J_H = 7.8 Hz, ^2J_H = 13.9 Hz, CH_2).

^{13}C NMR: (300 MHz, acetone-d^6) 175.3 (COOH), 138.7 (CH, arom., *i*), 130.4 (2 CH, arom., *m*), 128.9 (2 CH, arom., *m*), 127.2 (CH, arom., *p*), 72.1 (CHOH), 41.2 ppm (CH_2).

5.9 Synthesis of Aldehyde Acceptors for Transketolase Cascades

Routes (Ia): Routes Involving Reduction of a Lactone Precursor

5.9.1 Routes (Ia, A): Protection of 2-Hydroxy-γ-butyrolactone

5.9.1.1 *rac*-2-(*tert*-Butyldiemthylsilyl)-γ-butyrolactone

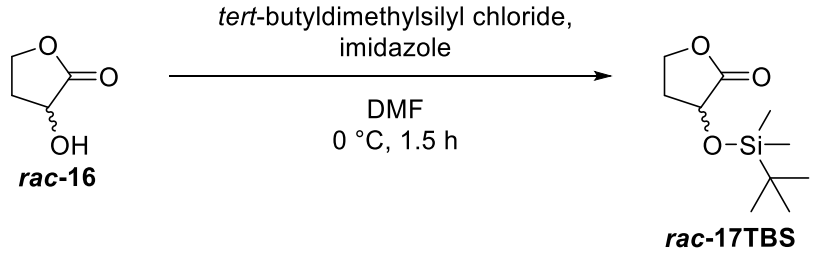


Figure 133: TBDMS-Protection of ***rac*-16**.

Imidazol (5.22 g, 76.7 mmol, 2.00 eq.) is dissolved in a solution of ***rac*-16** (*rac*-2-hydroxy-γ-butyrolactone, 3.92 g, 38.4 mmol) in DMF (50 mL, dry). *Tert*-butyldimethylsilyl chloride (7.53 g, 49.9 mmol, 1.30 eq.) is added in small portions over 5 min under stirring at 0 °C. After 1.5 h at 0 °C, NH_4Cl (30 mL, sat., aq.) is added at r.t..

The suspension is extracted with Et_2O (3*40 mL), the united etheric phases washed with NH_4Cl (2*30 mL, sat., aq.), the solvent is removed and the crude subjected to flash chromatography (120 g silica gel, 60 Å, 40 – 63 µm, diameter 7 cm, length 12 cm, DCM, 20 mL fractions collected).

Removal of the eluent yields ***rac*-17TBS** (*rac*-2-(*tert*-butyldiemthylsilyl)-γ-butyrolactone) as a colourless liquid with low viscosity and mouldy odour (7.45 g, 89%, fractions 6 – 16).

TLC (DCM/Et_2O 9 + 1, Seebach reagent):

R_f(*rac*-2-(*tert*-butyldiemthylsilyl)-γ-butyrolactone) = 0.40 (blue).

^{1}H NMR: (300 MHz, $CDCl_3$): 4.45 – 4.33 (m, 2H, CH_2O), 4.19 (td, ^3J_H = 9.2 Hz, ^3J_H = 6.4 Hz, CHO), 2.46 (dddd, 1H, ^3J_H = 7.7, ^3J_H = 6.4, ^3J_H = 3.3, ^2J_H = 12.7 Hz, CH_2), 2.29 – 2.13 (m, 1H, CH_2), 0.91 (s, 9H, 3*CH_3, *t*butyl), 0.17 (s, 3H, CH_3), 0.14 ppm (s, 3H, CH_3).

^{13}C NMR: (126 MHz, $CDCl_3$): 176.0 (COO, lactone), 68.4 (CHO), 64.8 (CH_2O), 32.5 (CH_2), 25.8 (3*CH_3, *t*butyl), 18.4 (CO, *t*butyl), - 4.6 (CH_3), - 5.1 ppm (CH_3).

5.9.1.2 *rac*-2(Methoymethoxy)-γ-butyrolactone

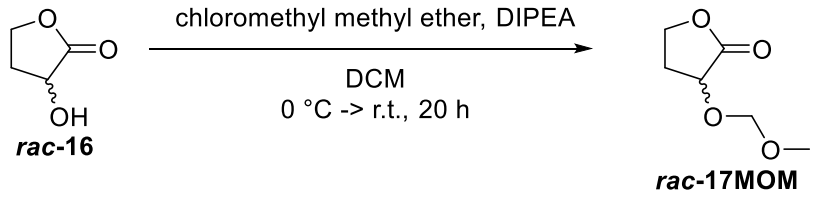


Figure 134: MOM-Protection of ***rac*-16**.

Diisopropylethylamine (3.6 mL, 20.7 mmol, 1.76 eq.) is added to ***rac*-16** (2-hydroxy-γ-butyrolactone, 1.20 g, 11.8 mmol) in DCM (30 mL, dry) at 0 °C under argon. Chloromethyl methyl ether (1.05 mL, 13.8 mmol, 1.18 eq.) is added at 0 °C, over 2 min to the stirred solution. After 2 h the mixture is warmed to r.t. and stirred for 18 h.

NH_4Cl (20 mL, sat. aq.) is added and the suspension extracted with Et_2O (3*50 mL). The combined etheric phases are washed with brine (50 mL), Et_2O is removed and the crude subjected to flash chromatography (120 g silica gel, 60 Å, 40 – 63 µm, diameter 7 cm, length 12 cm, 700 mL DCM/Et_2O 97 + 3, 20 mL fractions collected).

Removal of the eluent yields ***rac*-17MOM** (*rac*-2(methoymethoxy)-γ-butyrolactone) as a clear, colourless liquid (640 mg, 37%, fractions 18 – 28).

TLC (DCM/Et_2O 9 + 1, Seebach reagent):

R_f(*rac*-2-(methoxymethyl)-γ-butyrolactone) = 0.32 (blue).

^1H NMR: (300 MHz, $CDCl_3$): 4.83 (dd, 2H, J_1 = 69.5, J_2 = 6.8 Hz, CH_2O_2), 4.42 (td, 2H ^3J_H = 8.8 Hz, ^3J_H = 3.3 Hz, CH_2O), 4.23 (td, 1H,^3J_H = 9.3, ^3J_H = 6.4 Hz, CHO), 3.42 (s, 3H, CH_3O), 2.57 (dddd, 1H, ^3J_H = 7.9, ^3J_H = 6.4 Hz, ^3J_H = 3.1, ^2J_H = 12.8 Hz, CH_2), 2.28 ppm (dq, 1H, ^3J_H = 8.9, ^2J_H = 12.8 Hz, CH_2).

^{13}C NMR: (75 MHz, $CDCl_3$): 175.0 (COO, lactone), 96.1 (CH_2OO), 70.4 (CHO), 65.3 (CH_2O), 56.0 (CH_3O), 30.0 ppm (CH_2).

5.9.1.3 *rac*-2-Acetoxy-γ-butyrolactone

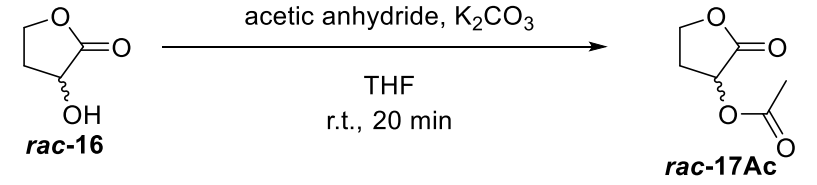


Figure 135: Acetyl-Protection of ***rac*-16**.

Acetic anhydride (950 µL, 10.0 mmol, 1.08 eq.) and ***rac*-16** (2-hydroxy-γ-butyrolactone, 950 mg, 9.31 mmol) are added to a stirred suspension of potassium carbonate (2.93 g, 21.2 mmol, 2.28 eq.) and molecular sieves (120 mg, powder, 4 Å) in THF (5 mL, dry) at r.t..

After 20 min, ethanol (5 mL, 97%) is added and the suspension is filtered over silica gel (100 g, apprx. d = 8 cm, l = 4 cm, pore size 60 Å, particle size 70 – 200 µm, wetted with DCM) and the filter flushed with DCM and Et_2O (10 mL each). Removal of the solvents yields ***rac*-17Ac** as a slightly yellow fluid (840 mg, 62%).

This product is synthesized as a racemic standard for CG/MS. A standard of the (*R*)-2-acetoxy-γ-butyrolacton is synthesized accordingly from a commercial sample of (*R*)-2-hydroxy-γ-butyrolacton.

TLC (DCM/Et_2O 97 + 3, Seebach's reagent):

R_f(2-acetoxy-γ-butyrolactone) = 0.27 (blue)

GC/MS:

t_{ret}(2(*R*)-acetoxy-γ-butyrolactone) = 14.91 min (43 [CH_3CO^+]),

t_{ret}(2(*S*)-acetoxy-γ-butyrolactone) = 14.53 min (43 [CH_3CO^+]).

^{1}H NMR: (300 MHz, $CDCl_3$): 5.41 (dd, 1H, ^3J_H = 9.7, ^3J_H = 8.6 Hz, CHOAc), 4.46 (td, 1H, $^{2,3}J_H$ = 9.1, ^3J_H = 2.5 Hz, CH_2O), 4.29 (td, 1H, $^{2,3}J_H$ = 9.6, ^3J_H = 6.4 Hz, CH_2O), 2.70 (dddd, 1H, ^3J_H = 8.8, ^3J_H = 6.4, ^3J_H = 2.5, ^2J_H = 12.9 Hz, CH_2), 2.28 (dtd, 1H, ^3J_H = 9.8, ^3J_H = 8.8, ^2J_H = 12.9 Hz, CH_2), 2.15 ppm (s, 3H, CH_3).

^{13}C NMR: (75 MHz, $CDCl_3$): 172.8 (COO, lactone), 169.9 (COO, acetyl), 67.8 (CHO), 65.1 (CH_2O), 29.0 (CH_2), 20.7 ppm (CH_3).

5.9.2 Routes (Ia, B): Reduction of Protected Lactones

5.9.2.1 2-*O*-(*tert*-Butyldimethylsilyl)-3-deoxy-D/L-*glycero*-tetrose

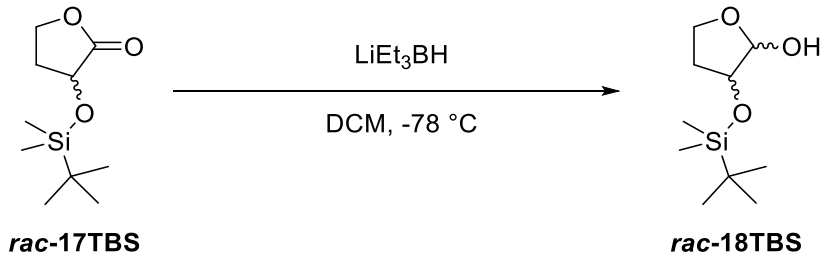


Figure 136: Lactone-reduction of ***rac*-17TBS**.

Over 2 min, a solution of lithium triethylborohydride (1.65 mL, 1.2 eq., 1 M in THF) is dropped into a stirred solution of ***rac*-17TBS** (2-((*tert*-butyldimethylsilyl)oxy)-γ-butyrolactone (299 mg, 1.39 mmol) in DCM (20 mL, dry, degassed) at - 78 °C. After 60 min, a mixture of methanol and water (3 mL, 3 + 1) is added. The reaction mixture is allowed to warm to r.t. over 30 min and stored overnight at 8 °C.

The phases are separated, the organic phase is reduced to 3 mL (300 mbar, 40 °C) and subjected to flash chromatography (120 g silica gel, 60 Å, 40 – 63 µm, diameter 2 cm, length 65 cm, 400 mL DCM/Et_2O 100 + 3, 10 mL fractions collected). Molar partitions within the two collected products are calculated according to ^{1}H-NMR peak integrals.

Removal of the eluent yields a colourless oil (157 mg, fractions 17 – 28, product A, containing 2-*O*-(*tert*-butyldimethylsilyl)-3-deoxy-D/L-*glycero*-tetrose, isomer I, 2-*O*-(*tert*-butyldimethylsilyl)-butane-1,2,4-triol, and *tert*-butyldimethylsilanol, molar ratio of hemiacetal and triol is 1:0.9) and another colourless oil (164 mg, fractions 40 – 55, containing two more hemiacetal isomers II and III, and the triol in a ratio of 1:0.7:0.2, respectively).

TLC (DCM/Et_2O 100 + 3, anisaldehyde reagent):

R_f(2-*O*-(*tert*-butyldimethylsilyl)-3-deoxy-D/L-*glycero*-tetrose, hemiacetal I) = 0.46 (black),

R_f(2-*O*-(*tert*-butyldimethylsilyl)-3-deoxy-D/L-*glycero*-tetrose, hemiacetal II and III) = 0.21 (black).

^{1}H NMR (500 MHz, $CDCl_3$, product A):

Peaks of **18TBS** (2-*O*-(*tert*-butyldimethylsilyl)-3-deoxy-glycero-tetrose), hemiacetal I: 5.29 (d, 1H, ^3J_H = 4.2 Hz, C-1HOOH), 4.17 (~dq, 1H, ^3J_H = 7.1, ^2J_H = 4.2 Hz, C-2HO), 4.11 – 3.92 (m, 1H, C-4H_2O-a, ~ 4.03, ddd, 1H, ^2J_H = 8.3, ^3J_H = 8.3, ^3J_H = 5.5 Hz)], 3.81 (ddd, 1H, ^2J_H = 8.2, ^3J_H = 8.2, ^3J_H = 6.9 Hz, C-4H_2O-b), 2.16 (dddd, 1H, ^2J_H = 13.0, ^3J_H = 7.6, ^3J_H = 7.5, ^3J_H = 5.5 Hz, C-3H_2-a), 1.91 – 1.76 (m, 1H, C-3H_2-b), 0.94 (s, 9H, *t*butyl), 0.18 ppm (s, 6H, 2*CH_3Si).

Peaks of 2-*O*-(*tert*-butyldimethylsilyl)-butane-1,2,4-triol: 4.11 – 3.92 (m, 3H, C-2HO, C-4H_2O-a,b, [~4.07 (dt, 1H, ^2J_H = 11.3, ^3J_H = 4.3 Hz, C-4H_2OH-a), 4.01 – 3.96 (m, 1H, C-3HO), 3.97 (~qd, 1H, ^2J_H = 11.1, ^3J_H = 3.2 Hz, C-4H_2OH-b)], 3.74 (dd, 1H, ^2J_H = 10.3, ^3J_H = 4.4 Hz, C-1H_2OH-a), 3.57 (dd, 1H, ^2J_H = 10.3, ^3J_H = 6.5 Hz, C-1H_2OH-b), 1.99 (ddt, 1H, ^2J_H = 14.1, ^3J_H = 3.7, ^3J_H = 3.7 Hz, C-3H_2-a, 1.91 – 1.76 (m, 1H, C-3H_2-b), 0.91 (s, 9H, *t*butyl), 0.09 ppm (s, 6H, 2*CH_3Si).

^{13}C NMR (75 MHz, $CDCl_3$, product A):

Peaks of **18TBS** (2-*O*-(*tert*-butyldimethylsilyl)-3-deoxy-glycero-tetrose), hemiacetal I: 96.67 (C-1HOOH), 72.64 (C-2HO), 65.46 (C-4H_2O), 32.26 (C-3H_2), 25.86 (3**t*butyl-CH_3), 18.11 (*t*butyl-C), - 4.17 (CH_3Si-a), - 4.99 ppm (CH_3Si-b).

Peaks of 2-*O*-(*tert*-butyldimethylsilyl)-butane-1,2,4-triol: 71.28 (C-2HO), 66.59 (C-1H_2OH), 60.87 (C-4H_2OH), 29.42 (C-3H_2), 26.00 (3**t*butyl-CH_3), 18.45 (*t*butyl-C), - 5.18 (CH_3Si-a), - 5.23 ppm (CH_3Si-b).

^{1}H NMR (500 MHz, $CDCl_3$, product B):

Peaks of **18TBS** (2-*O*-(*tert*-butyldimethylsilyl)-3-deoxy-glycero-tetrose), hemiacetal II: 9.66 (s, 1H, C-1HO), 5.22 (dd, 1H, J_H = 8.5, J_H = 4.0 Hz, C-1HOOH), 4.23 (~dt, 1H, J_H = 6.3, J_H = 3.7 Hz, C-2HO), 4.01 (~td, 1H, J_H = 8.5, J_H = 6.5 Hz, C-4H_2Oa), 3.95 (d, 1H, J_H = 8.6 Hz, C-4H_2Ob), 3.80 (ddd, 1H, J_H = 8.3, J_H = 8.2, J_H = 3.9 Hz, C-3HO), 2.06 (dddd, 1H, ^2J_H = 12.8, ^3J_H = 8.5, ^3J_H = 8.3, ^3J_H = 6.1 Hz, CH_2-a), 1.86 (dddd, ^2J_H = 12.9, ^3J_H = 6.7, ^3J_H = 3.4, ^3J_H = 3.4 Hz, CH_2-b), 0.91 (s, 9H, *t*butyl), 0.13 (s, 3H, CH_3Si-a), 0.12 ppm (s, 3H, CH_3Si-b).

Peaks of **18TBS** (2-*O*-(*tert*-butyldimethylsilyl)-3-deoxy-glycero-tetrose), hemiacetal III: 10.02 (s, 1H, C-1HO), 5.19 (s, 1H, C-1HOOH), 4.19 (dd, 1H, ^3J_H = 5.1, ^3J_H = 1.7 Hz, C-2H_2O), 4.06 (dd, 2H, ^3J_H = 8.8, ^3J_H = 5.1 Hz, C-4H_2O), 2.25 – 2.14 (dddd, 1H, ^2J_H = 12.6, ^3J_H = 8.7, ^3J_H = 8.7, ^3J_H = 4.9, CH_2-a), 1.77 (dddd, 1H, ^2J_H = 12.6, ^3J_H = 5.2,

$^{3}J_{H}$ = 5.1, $^{3}J_{H}$ = 1.8 Hz, CH_2-b), 0.88 (s, 9H, *t*butyl), 0.08 (s, 3H, CH_3Si-a), 0.07 ppm (s, 3H, CH_3Si-b).

Peaks of 2-*O*-(*tert*-butyldimethylsilyl)-butane-1,2,4-triol:4.08 – 3.89 (m, 3H, C-2HO, C-4H_2O-a,b), 3.71 (dd, 1H, $^{2}J_{H}$ = 10.3, $^{3}J_{H}$ = 4.5 Hz, C-1H_2OH-a), 3.54 (dd, 1H, $^{2}J_{H}$ = 10.3, $^{3}J_{H}$ = 6.4 Hz, C-1H_2OH-b), 1.96 (ddt, 1H, $^{2}J_{H}$ = 14.1, $^{3}J_{H}$ = 3.7, $^{3}J_{H}$ = 3.7 Hz, C-3H_2-a, 1.88 – 1.73 (m, 1H, C-3H_2-b), 0.88 (s, 9H, *t*butyl), 0.06 ppm (s, 6H, 2*CH_3Si).

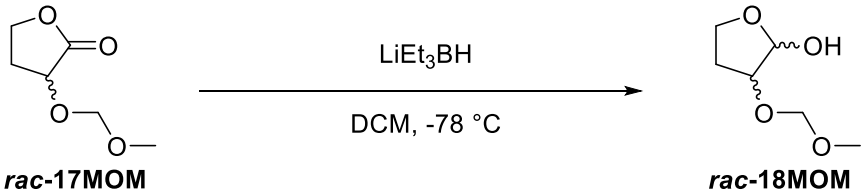
^{13}C NMR (75 MHz, $CDCl_3$, product B):

Peaks of **18TBS** (2-*O*-(*tert*-butyldimethylsilyl)-3-deoxy-glycero-tetrose), hemiacetal II: 103.36 (C-1HOOH), 72.26 (CHO), 64.88 (CH_2O), 33.84 (CH_2), 25.82 (3**t*butyl-CH_3), 18.18 (*t*butyl-C), - 4.60 (CH_3Si-a), - 4.99 ppm (CH_3Si-b).

Peaks of **18TBS** (2-*O*-(*tert*-butyldimethylsilyl)-3-deoxy-glycero-tetrose), hemiacetal III: 97.40 (C-1HOOH), 77.02 (CHO), 67.32 (CH_2O), 32.87 (CH_2), 25.89 (3**t*butyl-CH_3), 18.20 (*t*butyl-C), - 4.68 (CH_3Si-a), - 4.71 ppm (CH_3Si-b).

Peaks of 2-*O*-(*tert*-butyldimethylsilyl)-butane-1,2,4-triol: 71.26 (C-2HO), 66.57 (C-1H_2OH), 60.85 (C-4H_2OH), 29.39 (C-3H_2), 25.98 (3**t*butyl-CH_3), 18.43 (*t*butyl-C), - 5.21 (CH_3Si-a), - 5.26 ppm (CH_3Si-b).

5.9.2.2 3-Deoxy-2-*O*-(methoxymethyl)-D/L-*glycero*-tetrose



***rac*-17MOM** ***rac*-18MOM**

Figure 137: Lactone-reduction of ***rac*-17MOM**.

Over 2 min, a solution of lithium triethylborohydride (1.65 mL, 1.65 mmol, 1.2 eq., 1 M in THF) is dropped into a stirred solution of ***rac*-17MOM** (2-(methoxymethoxy)-γ-butyrolactone, 202 mg, 1.39 mmol) in DCM (20 mL, dry, degassed), at - 78 °C, under argon. After 60 min, a mixture of methanol and water (3 mL, 3 + 1) is added. The reaction mixture is allowed to warm to r.t. over 30 min and stored overnight at 8 °C.

The phases are separated, the organic phase reduced to 3 mL (300 mbar, 40 °C) and subjected to flash chromatography (120 g silica gel, 60 Å, 40 – 63 µm, diameter 2 cm, length 65 cm, 400 mL DCM/Et_2O 100 + 3, 10 mL fractions collected).

Removal of the eluent yields ***rac*-18MOM** (3-deoxy-2-*O*-(methoxymethyl)-D/L-*glycero*-tetrose) as a colourless oil (77 mg, 38%, containing three hemiacetal isomers in a ratio of 1:0.5:0.3, according to their ^{1}H NMR integrals).

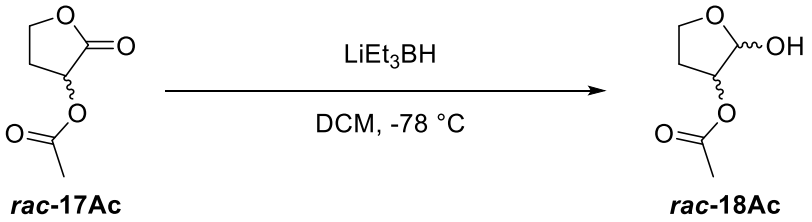
TLC (DCM/Et_2O 3 + 1, anisaldehyde reagent):

R_f(3-deoxy-2-*O*-(methoxymethyl)-D/L-*glycero*-tetrose) = 0.37 (black).

^{1}H NMR (500 MHz, $CDCl_3$, peaks of major isomer): 9.67 (s, 1H, C-1HO), 5.39 (s, 1H, C-1HOOH), 4.90 (bs, 1H, OH), 4.68 (s, 2H, CH_2OO), 4.15 – 4.02 (m, 2H, C-4H_2O), 4.13 (dd, 1H, $^{3}J_H$ = 5.5, $^{3}J_H$ = 1.8 Hz, C-2HO), 3.37 (s, 3H, CH_3O), 2.26 (dddd, 1H, $^{2}J_H$ = 13.7, $^{3}J_H$ = 8.5, $^{3}J_H$ = 8.4, $^{3}J_H$ = 5.5 Hz, C-3H_2), 1.95 ppm (dddd, 1H, $^{2}J_H$ = 13.5, $^{3}J_H$ = 7.4, $^{3}J_H$ = 4.3, $^{3}J_H$ = 1.9 Hz, C-3H_2-b).

^{13}C NMR (126 MHz, $CDCl_3$): 101.22 (C-1HOOH, major isomer), 96.57 (C-1HOOH, one of minor isomers), 96.38 (CH_2OO, one of minor isomers), 95.76 (CH_2OO, major isomer), 81.43 (C-2HO, major isomer), 76.65 (C-2HO, one of minor isomers), 66.99 (C-4H_2O, major isomer), 64.95 (C-4H_2O, one of minor isomers), 56.01 (CH_3O, one of minor isomers), 55.64 (CH_3O, major isomer), 30.11 ppm (C-3H_2, one of minor isomers), 30.03 ppm (C-3H_2, major isomer).

5.9.2.3 2-*O*-Acetyl-3-deoxy-D/L-*glycero*-tetrose



***rac*-17Ac** ***rac*-18Ac**

Figure 138: Lactone-reduction of ***rac*-17Ac**.

Over 2 min, a solution of lithium triethylborohydride (1.65 mL, 1.2 eq., 1 M in THF) is dropped into a stirred solution of ***rac*-17Ac** (2-acetoxy-γ-butyrolactone, 200 mg, 1.39 mmol) in DCM (20 mL, dry, degassed) at - 78 °C, under argon. After 60 min, a mixture of methanol and water (3 mL, 3 + 1) is added. The reaction mixture is allowed to warm to r.t. over 30 min and stored overnight at 8 °C.

The phases are separated, the organic phase reduced to 3 mL (300 mbar, 40 °C) and subjected to flash chromatography (120 g silica gel 60 Å, 40 – 63 µm, diameter 2 cm, length 65 cm, 400 mL DCM/Et_2O 100 + 3, 10 mL fractions collected).

Removal of the eluent yields ***rac*-18Ac** (2-*O*-acetyl-3-deoxy-D/L-*glycero*-tetrose) as a colourless oil (140 mg, mixture).

TLC (DCM/Et_2O 100 + 3, anisaldehyde reagent):

R_f(2-*O*-acetyl-3-deoxy-D/L-*glycero*-tetrose) = 0 - 0.32 (black).

^1H NMR (500 MHz, $CDCl_3$, peaks of major product **18Ac**): 10.04 (s, 1H, C-1), 5.96 (d, 1H, ^3J_H = 4.2 Hz, C-1HOOH), 4.84 (dd, 1H, ^3J_H = 6.0, ^3J_H = 4.3 Hz, C-2HO), 3.99 (t, 1H, ^3J_H = 8.4 Hz, C-4H_2O-a), 3.77 – 3.72 (m, 1H, C-4H_2O), 2.14 (dddd, 1H, ^2J_H = 14.3, ^3J_H = 7.1, ^3J_H = 5.6, ^3J_H = 4.3 Hz, C-3H_2-a), 2.03 (s, 3H, CH_3), 2.01 – 1.88 ppm (m, 1H, C-3H_2-b).

^{13}C NMR (126 MHz, $CDCl_3$, peaks of major product **18Ac**): 171.0 (COO), 105.7 (C-1HOOH), 80.9 (C-2HO), 65.1 (C-4H_2O), 33.8 (C-3H_2), 21.0 ppm (CH_3).

5.9.2.4 2,3(*R*)-Diacetoxytetrahydrofuran

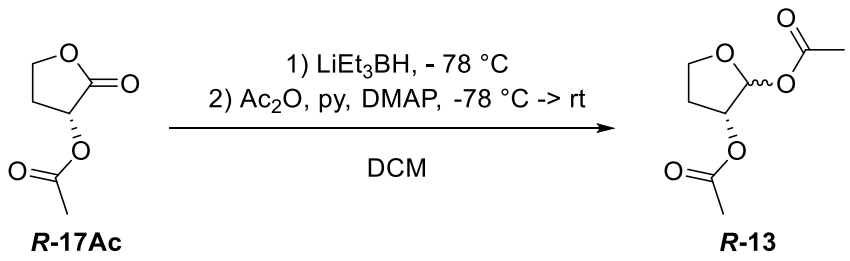


Figure 139: Telescoped reaction sequence of reduction and acetyl-protection of ***R*-17Ac**.

Over 10 min, a solution of lithium triethylborohydride (36.6 mL, 1.06 eq., 1 M in THF) is dropped into a stirred solution of ***R*-17Ac** (2(*R*)-acetoxy-γ-butyrolactone, 5 g, 34.7 mmol) in DCM (450 mL, dry, degassed) at - 78 °C, under argon. After 30 min, acetic anhydride (4.40 mL, 46.5 mmol, 1.34 eq.), pyridine (2.80 mL, 34.8 mmol, 1.00 eq.) and a solution of DMAP (13.5 mL, 23.2 mmol, 0.668 eq., 1.72 M in dry DCM) are added. After stirring for 2 h, NH_4Cl (15.0 mL, 104 mmol, sat., aq.) is added and the solution allowed to warm to r.t. over 1.5 h.

The organic phase is washed with NH_4Cl (2*20 mL, sat., aq.). After removing the solvent, the crude (yellow oil) is stored overnight at 7 °C.

The crude is subjected to flash chromatography (200 g silica gel, 60 Å, 40 – 63 µm, diameter 6 cm, length 15 cm, 600 mL cyclohexane/DCM/Et_2O 20 + 77 + 3, 20 mL fractions collected).

Removal of the eluent yields a colourless oil (2.41 g, 37%, fraction 12, product A, consisting of a mixture of *anti*- and *syn*-2,3(*R*)-diacetoxytetrahydrofuran in a ratio of 1.4:1), another colourless oil (1.74 g, 26%, fractions 13 – 15, product B, consisting of a mixture of *anti*- and *syn*-2,3(*R*)-diacetoxytetrahydrofuran in a ratio of 1:9.7), and another colourless oil (1.19 g, 14%, fractions 16 – 31, product C, (*R*)-1,2,4-butanetriyl acetate). Molar ratios are given according to ^{1}H NMR acetal peak integrals. Total yield of the desired acetal is 4.15 g, 63%.

^{1}H NMR (500 MHz, $CDCl_3$, product A, peaks of major isomer): 6.11 (s, 1H, C-1HOO), 5.14 (d, $^{3}J_H$ = 5.6 Hz, C-2HO), 4.09 (dd, 2H, $^{3}J_H$ = 8.7, $^{3}J_H$ = 5.5 Hz, C-4H_2O), 2.38 - 2.29 (m, 1H, C-3H_2-a), 2.04 (s, 3H, CH_3, acetyl-a), 2.02 (s, 3H, CH_3, acetyl-b), 2.01 - 2.95 ppm (m, 1H, CH_3, acetyl-b).

^{13}C NMR (75 MHz, $CDCl_3$, product A): 170.16 (COO, acetyl-a), 169.63 (COO, acetyl-b), 99.69 (C-1HOO), 77.14 (C-2HO), 68.19 (C-4HO), 29.56 (C-3H_2), 21.16 (CH_3, acetyl-a), 20.94 ppm (CH_3, acetyl-b).

^{13}C NMR (DEPT, $CDCl_3$, $CHCl_3$ at 77.26 ppm, product A): 99.7 (C-1HOO), 77.1 (C-2HO), 68.2 (C-4HO), 29.5 (C-3H_2), 21.2 (CH_3, acetyl-a), 20.9 ppm (CH_3, acetyl-b).

^{1}H NMR: (500 MHz, $CDCl_3$, product B, peaks of major isomer): 6.26 (d, 1H, $^{3}J_H$ = 4.3 Hz, C-1HOO), 5.12 (td, 1H, $^{3}J_H$ = 8.5, $^{3}J_H$ = 4.2 Hz, C-2HO), 4.11 (td, 1H, $^{3}J_H$ = 9.0, $^{2}J_H$ = 8.9, $^{3}J_H$ = 4.5 Hz, C-4H_2O-a), 3.93 (td, 1H, $^{2}J_H$ = 8.0, $^{3}J_H$ = 8.0 Hz, C-4H_2O-b), 2.31 (dddd, 1H, $^{2}J_H$ = 11.1, $^{3}J_H$ = 7.3, $^{3}J_H$ = 6.6, $^{3}J_H$ = 3.8 Hz, C-3H_2-a), 2.05 - 2.00 (dm, 1H, $^{3}J_{HH}$ = 5.5 Hz, C-3H_2-b), 2.04 (s, 3H, CH_3, acetyl-a), 2.03 ppm (s, 3H, CH_3, acetyl-b).

^{13}C NMR (75 MHz, $CDCl_3$, product B, peaks of major isomer): 170.25 (COO, acetyl-a), 169.94 (COO, acetyl-b), 94.17 (C-1HOO), 72.37 (C-2HO), 66.29 (C-4H_2O), 27.49 (C-3H_2), 21.12 (CH_3, acetyl-a), 20.65 ppm (CH3, acetyl-b).

^{13}C NMR (DEPT, $CDCl_3$, $CHCl_3$ at 77.26 ppm, product B, peaks of major isomer): 94.16 (C-1HOO), 72.36 (C-2HO), 66.29 (C-4H_2O), 27.47 (C-3H_2), 21.12 (CH_3, acetyl-a), 20.65 ppm (CH_3, acetyl-b).

^{1}H NMR: (500 MHz, $CDCl_3$, product C): 5.14 (tt, 1H, $^{3}J_H$ = 6.0, $^{3}J_H$ = 5.8, $^{3}J_H$ = 3.4 Hz, C-2HO), 4.24 (dd, 1H, $^{2}J_H$ = 12.0, $^{3}J_H$ = 3.5, C-1H_2O-a), 4.08 (t, 2H, $^{3}J_H$ = 6.3 Hz, C-4H_2O), 4.02 (dd, 1H, $^{2}J_H$ = 12.0, $^{3}J_H$ = 6.1, C-1H_2O-b), 2.03 (s, 6H, 2*CH_3, acetyl,

terminal), 2.01 (s, 3H, CH_3, acetyl, C-2), 1.90 ppm (td, 2H, ^3J_H = 6.6, ^3J_H = 6.4 Hz, C-3H_2).

^{13}C NMR: (75 MHz, $CDCl_3$, product C): 171.0 (COO, acetyl-a), 170.7 (COO, acetyl-b), 170.4 (COO, acetyl-c), 68.6 (C-2HO), 64.9 (C-1H_2O), 60.3 (C-4H_2O), 29.9 (C-3H_2), 21.0 (CH_3, acetyl-a), 20.9 (CH_3, acetyl-b), 20.8 ppm (CH_3, acetyl-c).

^{13}C NMR: (DEPT, 75 MHz, $CDCl_3$, $CHCl_3$ - 77.26 ppm, product C) 68.5 (C-2HO), 64.8 (C-1H_2O), 60.2 (C-4H_2O), 29.8 (C-3H_2), 20.9 (CH_3, acetyl-a), 20.8 (CH_3, acetyl-b), 20.7 ppm (CH_3, acetyl-c).

Product A and B are united and further purified *via* flash chromatography (400 g, silica gel, 60 Å, 40 – 63 µm, diameter 6 cm, length 30 cm, 400 mL DCM/cyclohexane/Et_2O 50 + 47 + 3, 20 mL fractions collected).

Removal of the eluent yields ***R*-13** (2,3(*R*)-diacetoxytetrahydrofuran) as a colourless oil (3.79 g, 58%, mixture of *anti*- and *syn*-2,3(*R*)-diacetoxytetrahydrofuran in a ratio of 3:2).

TLC (DCM/Et_2O 97 + 3, Seebach reagent):

R_f(*anti*-2(*R*),3-diacetoxytetrahydrofuran) = 0.47 (blue)

R_f(*syn*-2(*R*),3-diacetoxytetrahydrofuran) = 0.38 (blue)

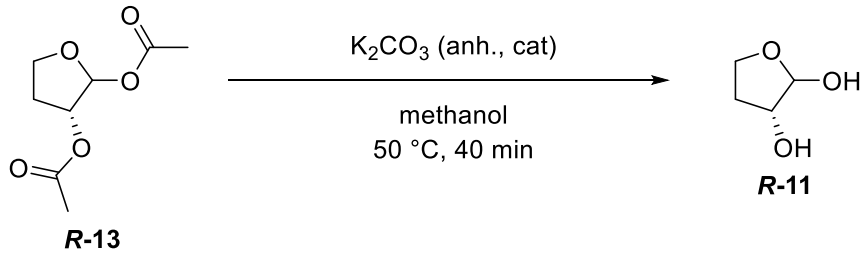
R_f((R)-1,2,4-butanetriyl acetate) = 0.23 (blue).

^{1}H NMR (500 MHz, $CDCl_3$): 6.31 – 6.27 (m, 1H, C-1HOO, minor isomer), 6.16 – 6.12 (m, 1H, C-1HOO, major isomer), 5.19 – 5.11 (m, C-2-HO, major and minor isomers), 4.16 – 4.08 (m, C-4H_2O-a, b, major isomer, C-4H_2O-a, minor isomer), 4.00 – 3.93 (m, C-4H_2O-b, minor isomer), 2.40 – 2.29 (m, C-3H_2-a, major and minor isomers), 2.11 – 1.95 ppm (m, C-3H_2-b, major and minor isomers, CH_3, acetyl-a, b, major and minor isomers).

^{13}C NMR (126 MHz, $CDCl_3$): 170.27 (COO, acetyl-a, minor isomer), 170.19 (COO, acetyl-a, major isomer), 169.96 (COO, acetyl-b, minor isomer), 169.66 (COO, acetyl-b, major isomer), 99.71 (C-1HOO, major isomer), 94.21 (C-1HOO, minor isomer), 77.16 (C-2HO, major isomer), 72.40 (C-2HO, minor isomer), 68.23 (C-4H_2O, major isomer), 66.34 (C-4H_2O, minor isomer), 29.59 (C-3H_2, major isomer), 27.52 (C-3H_2,

minor isomer), 21.20 (CH_3, acetyl-a, major isomer), 21.18 (CH_3, acetyl-a, minor isomer), 20.99 (CH_3, acetyl-b, major isomer), 20.71 ppm (CH_3, acetyl-b, minor isomer).

5.9.2.5 3-Deoxy-D-glycero-tetrose (2(*R*),4-Dihydroxybutanal, Deprotection)



***R*-13** **R-11**

Figure 140: Deprotection of ***R*-13**.

2,3(*R*)-Diacetoxytetrahydrofuran (1.53 g, 8.13 mmol) is dissolved in methanol (4 mL, dry) and K_2CO_3 (135 mg 0.977 mmol, 0.12 eq., anh) is added. The containing flask is closed and submerged into an ultrasonic bath at 50 °C for 30 min. Water (10 mL, d.i.) is added and methanol is removed at reduced pressure. The resulting solution is frozen until usage in transketolase cascade reactions.

For quantification of the concentration, an aliquot (200 µL) is lyophilized and the residue taken up in D_2O (600 µL), containing maleic acid (70.84 mM) as a standard.

The concentration of ***R*-11** (3-deoxy-D-*glycero*-tetrose) is determined, comparing the ^{1}H NMR areas of the maleic acid peak (s, 6.39 ppm, 2H, CH, area = 2.000) and one of the product's peaks (m, 2.23 ppm, 1H, CH_2-a, area = 2.833). The calculated concentration equals 600 mM (equivalent to 73% yield) in the original solution.

TLC (DCM/methanol/formic acid, 85 + 15 + 0.1, anisaldehyde reagent):

R_f(3-deoxy-D-glycero-tetrose) = 0.40 (black).

1H NMR (500 MHz, D_2O): 9.62 (d, 1H, ^3J_H = 3.0 Hz, CHO, free aldehyde) 6.33 (s, 2H, maleic acid CH), 5.24 (s, 1H, CHOOH, lactol), 4.27 – 4.19 (m, 1H, CHOH), 4.14 – 3.99 (m, 2H, CH_2OH), 2.30 – 2.16 (m, 1H, CH_2-a), 1.94 – 1.87 ppm (m, 1H, CH_2-b).

5.9.3 Routes (Ib): Kinetic Resolution of 2-Hydroxy-γ-butyrolactone

5.9.3.1 (*R*)-2-Acetoxy-γ-butyrolactone

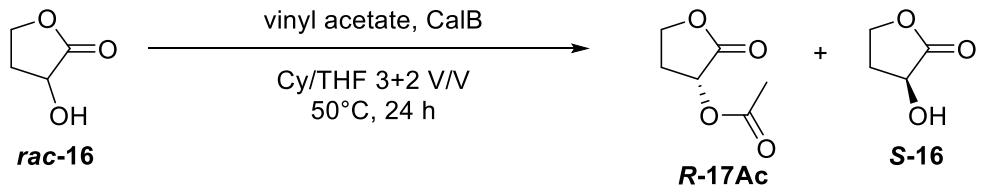


Figure 141: Kinetic resolution of ***rac*-16**.

To lipase (6.50 g, immobilized *Candida antarctica* lipase B, beads, Sprin technologies) and molecular sieves (1 g, powder, 4 Å) under a mixture (600 mL, 3 + 2) of cyclohexane (dry) and THF (dry), vinyl acetate (36.0 mL, 391 mmol, 2.0 eq.) is added. This mixture is warmed to 50 °C under stirring. Then, ***rac*-16** (*rac*-2-hydroxy-γ-butyrolactone, 20.00 g, 196 mmol, 1.0 eq.) is added to the suspension. After 1 h, another portion of molecular sieves (1.2 g) is added.

After 24 h, the mixture is filtered and reduced to 40 mL of a two-phasic mixture. The top-phase is reduced to 1 mL (yellow oil) and united with the bottom-phase (20 mL, yellow oil). United phases are subjected to flash chromatography (2 kg silica gel, 60 Å, 40 – 63 µm, diameter 12 cm, length 35 cm, cyclohexane/DCM/Et_2O 50 + 45 + 5, then cyclohexane/DCM/Et_2O 20 + 75 + 5, then cyclohexane, 50 mL fractions collected).

Removal of the eluent yields ***R*-17Ac** (2(*R*)-acetoxy-γ-butyrolactone) as a clear fluid (10.6 g, 37%, fractions 96 – 140) and ***S*-16** (2(*S*)-hydroxy-γ-butyrolactone) as a slightly red fluid (8.3 g, 41%). A side product, with similar polarity as the acetylated alcohol, supposedly with opened lactone is discarded.

TLC (DCM/Et_2O 97 + 3, Seebach's reagent):

R_f(2(*R*)-acetoxy-γ-butyrolactone) = 0.27 (blue),

R_f(2(*S*)-hydroxy-γ-butyrolactone) = 0.08 (blue).

^{1}H NMR: (500 MHz, $CDCl_3$, 2(*R*)-Acetoxy-γ-butyrolactone): 5.40 (t, 1H, ^3J_H = 9.1 Hz, CHOAc), 4.45 (t, 1H, ^3J_H = 9.1 Hz, CH_2O), 4.32 – 4.23 (m, 1H, CH_2O), 2.74 – 2.64 (m, 1H, CH_2), 2.34 – 2.21 (m, 1H, CH_2), 2.14 ppm (s, 3H, CH_3).

^{13}C NMR: (126 MHz, $CDCl_3$, 2(*R*)-Acetoxy-γ-butyrolactone): 172.9 (COO, lactone), 169.9 (COO, acetyl), 67.7 (CHO), 65.1 (CH_2O), 29.0 (CH_2), 20.7 ppm (CH_3).

^{1}H NMR: (500 MHz, $CDCl_3$, 2(*S*)-Hydroxy-γ-butyrolactone): 4.51 (t, 1H, $^{3}J_H$ = 9.2 Hz, CHOH), 4.41 (t, 1H, $^{3}J_H$ = 9.0 Hz, CH_2O), 4.26 – 4.17 (m, 1H, CH_2O), 2.58 (dddd, 1H, $^{3}J_H$ = 12.7, $^{3}J_H$ = 8.3, $^{3}J_3$ = 6.1, $^{2}J_H$ = 1.8 Hz, CH_2), 2.32 – 2.20 ppm (m, 1H, CH_2).

^{13}C NMR: (126 MHz, $CDCl_3$, 2(*S*)-Hydroxy-γ-butyrolactone): 178.4 (COO), 67.5 (CHOH), 65.4 (CH_2O), 30.9 ppm (CH_2).

5.9.3.2 Determination of the CalB *E*-Factor towards 2-Hydroxy-γ-butyrolactone

***rac*-16** (3-hydroxy-γ-butyrolactone, 198 mg, 1.94 mmol) is dissolved in THF (20 mL). Lipase (54 mg, immobilized *Candida Antarctica* lipase B, beads, Sprin technologies), molecular sieves (102 mg, powder, 4 Å) and vinyl acetate (0.38 mL, 4.1 mmol, 2.1 eq.) are added. The suspension is warmed to 50 °C and stirred for 16 h and 35 min.

A calibration of ***rac*-16** (2-hydroxy-γ-butyrolactone, 50, 60, 70, 80, 90, 100 mM) is prepared in THF.

Reaction samples and calibration solutions (100 µL each) are mixed with the internal standard 2-butanol (50 µL, 30 mM) and then derivatized as *tert*-butyldimethylsilyl ether, by mixing the sample, containing internal standard, with a derivatization solution (150 µL, *tert*-butyldimethylsilyl chloride, 1 M and imidazole, 2.5 M in dry THF) and heating (30 min, 60 °C). The derivatized mixture is allowed to cool to r.t.. Excess derivatization reagent is reacted with ethanol (99%, 10 µL).

Neat reaction samples are injected into GC/MS (DexCB chiral column, electron ionization) and derivatized samples into GM/MS (DB5, electron ionization).

Relative areas of (*R*)- and (*S*)-2-acetoxy-γ-butyrolactone peaks are measured as (*R*): 90.37% and (*S*): 9.63%.

The concentration of remaining 2-hydroxy-γ-butyrolactone is determined to be 57.1 mM. Starting the reaction with a concentration of 96 mM, this equals a conversion of 40.5%.

The *ee* of the product (*R*)-2-acetyl-γ-butyrolactone is calculated as

$$ee\big((R) - Product\big) = \frac{|[R - Product] - [S - Product]|}{100} = \frac{|90.37\% - 9.63\%|}{100}$$
$$= 80.74\%$$

The *E*-Factor is calculated using the following equation:

$$E = \frac{\ln\left[1 - c \cdot \left(1 + ee\left((R) - product\right)\right)\right]}{\ln\left[1 - c \cdot \left(1 + ee\left((R) - product\right)\right)\right]} = \frac{\ln\left[1 - 0.405 \cdot (1 + 0.8074)\right]}{\ln\left[1 - 0.405 \cdot (1 - 0.8074)\right]} = 16$$

GC/MS (DexCB, electron ionization):

t_{ret}((*R*)-2-acetoxy-γ-butyrolactone) = 19.62 min (M^+_{acetyl} = 43),

t_{ret}((S)-2-acetoxy-γ-butyrolactone) = 19.87 min (M^+_{acetyl} = 43).

GC/MS (DB5, electron ionization):

t_{ret}(2-hydroxy-γ-butyrolactone) = 22.06 min,

t_{ret}((S)-2-acetoxy-γ-butyrolactone) = 19.95 min (M^+_{acetyl} = 43).

5.9.4 Routes (II): Production of a Lactone Precursor from L-Homoserine

5.9.4.1 D-2,4-Dihyroxybutyric acid

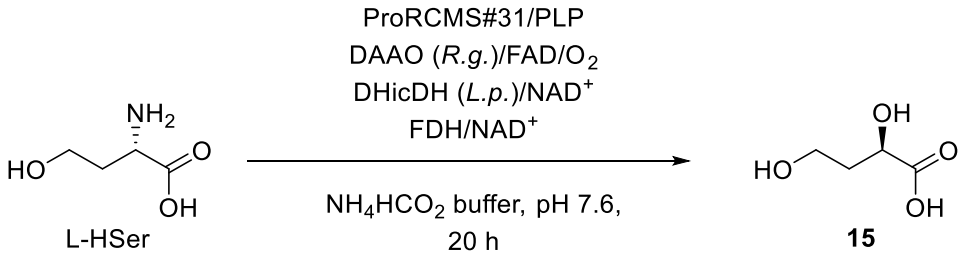


Figure 142: Cascade reaction to produce **15** from L-homoserine (L-HSer), employing ProRCMS#31, DAAO, and D-HicDH.

Small-scale test reactions:

Test reactions (500 µL each) are incubated (25 °C, 600 rpm) in Eppendorf vials (1.5 mL, caps punctured) with L-homoserine (L-HSer), varied amounts of ProRCMS#31 (cell-free extract), DAAO (*Rhodotorula gracilis*, cell-free extract), D-HicDH (*Lactobacillus paracasei*, cell-free extract), FDH (*Candida bodinii*, Boehringer Mannheim), NAD^+ and ammonium formate buffer (pH 7.6). The respective concentrations are given in Table 12. All reactions contain PLP and FAD (0.1 mM each).

A sample (100 µL) is taken from each reaction after 20 h. Samples are acidified with HCl (0.3 M, aq., 30 µL) and heated (3 min, 90 °C). Samples are neutralized with NaOH (0.4 M, aq., 22.5 µL). L-alanine (149 mM, aq., 20 µL) and eluent B (10 mM ammonium formate in ACN/water/formic acid 87 + 13 + 0.1, 127.5 µL) are added. The mixture is

filtered through a syringe filter and analysed on HPLC/MS (hydrophilic liquid interaction chromatography).

Preparative reaction:

A 2*3 cm window is cut into the side of a Falcon tube, apprx. 1 cm below its thread for easy access. The tube is painted black to exclude light. The window is kept closed and sealed with scotch tape.

D-HicDH (50 mg, *Lactobacillus paracasei*, cell-free extract), DAAO (103 mg, *Rhodotorula* gracilis, cell-free extract), ProRCMS#31 (101 mg, cell-free extract), FDH (12 mg, 6 U *Candida bodinii*, Boehringer Mannheim), catalase (110 mg, 253 kU, bovine liver, Sigma-Aldrich), FAD (2.1 mg, 0.1 mM) and PLP (0.8 mg, 0.1 mM) are weighed into the Falcon tube and suspended with 10 mL water (d.i.).

In another tube, NH_4HCO_2 (394 mg, 6.25 mmol, 2.55 eq.) and NADH (disodium salt, hydrate, 68 mg, 96 µmol, 3.6 mol% eq.), and L-homoserine (755 mg, 6.34 mmol) are dissolved in 10 mL water (d.i.), the pH adjusted to 7.2 with KOH (5 M, aq.) and made up to 15 mL with d.i. water. This solution is added to the enzyme suspension.

The Falcon tube is closed, tilted to the side to increase the suspensions surface and taped into a crystallizing dish. The tube’s lid is punctured with two cannulas, of which both are kept above the suspension's surface and of which one is connected to a balloon filled with oxygen.

The crystallization dish, containing the yellow reaction suspension is incubated (duration given, 25 °C, 700 rpm).

After 24 h, another portion of DAAO (103 mg, *Rhodotorula* gracilis, cell-free extract) and catalase (110 mg, 253 kU, bovine liver, Sigma-Aldrich) in water (2 mL, d.i.) is added and the oxygen balloon is refilled.

After another 24 h, the reaction is stopped by the addition of 20 mL of methanol and frozen for 1 week. The thawed mixture is centrifuged (20 min, 3100 *g*). The supernatant is collected and the solvent removed at reduced pressure, yielding a light brown highly viscous liquid (1.17 g, quant, 65% purity).

A portion of the crude (166 mg) is subjected to flash chromatography (20 g C-18 modified reverse phase silica gel, 90 Å, 40 – 63 µm, diameter 1 cm, length 15 cm, 100 mL H_2O/formic acid 100 + 0.1, 5 mL fractions collected).

Water evaporation yields **15** as a clear, colourless liquid (131 mg, quant, 82% purity, fractions 10 – 16).

TLC (ethyl acetate/^{n}butanol/^{i}propanol/H_2O/acetic acid, 7 + 1 + 1 + 1 + 1):

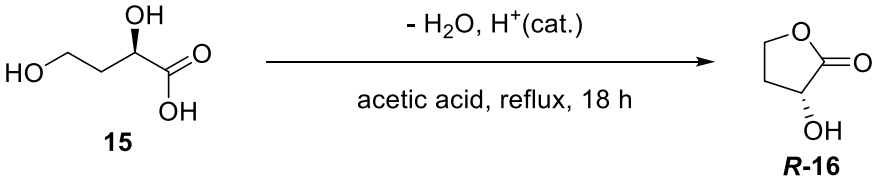
R_f(homoserine) = 0.14 – 3.3 (ninhydrin reagent, crimson)

R_f(2,4-dihydroxybutyric acid) = 0 – 0.29 (Seebach's regent, blue).

^{1}H NMR: (300 MHz, $CDCl_3$): 4.07 (dd, 1H, $^{3}J_H$ = 8.4, $^{3}J_H$ = 4.0 Hz, CHOH-a), 3.66 (dd, 2H, $^{3}J_H$ = 7.7, $^{3}J_H$ = 6.0 Hz, CH_2OH), 1.94 (ddt, 1H, $^{2}J_H$ = 11.5, $^{3}J_H$ = 7.5, $^{3}J_H$ = 4.0, CH_2-a), 1.76 ppm (ddt, 1H, $^{2}J_H$ = 14.3, $^{3}J_H$ = 8.3, $^{3}J_H$ = 6.0 Hz, CH_2-b).

^{13}C NMR: (75 MHz, $CDCl_3$): 182.0 (COO, lactone), 70.4 (CHO), 59.4 (CH_2O), 37.3 ppm (CH_2).

5.9.4.2 2(*R*)-Hydroxy-γ-butyrolactone (I)



15 ***R*-16**

Figure 143: Lactone formation in acetic acid.

15 (2(*R*),4-Dihydroxybutyric acid, 56.8 mg, 307 µmol, 65%) is dissolved in a mixture of acetic acid (glacial) and formic acid (1 mL, 9 + 1, V + V). Molecular sieves (50 mg, powder, 4 Å) is added and the mixture heated to reflux for 18 h.

The brown resin is subjected to flash chromatography (120 g silica gel 60 Å, 40 – 63 µm, diameter 7 cm, length 12 cm, 250 mL DCM/Et_2O 3 + 1, 5 mL fractions collected).

Removal of the eluent yields ***R*-16** as a clear, colourless liquid (18 mg, 57%, fractions 18 – 28).

TLC (DCM/Et_2O 3 + 1, Seebach reagent):

R_f(2(*R*)-hydroxy-γ-butyrolactone) = 0.34 (blue).

^{1}H NMR: (300 MHz, $CDCl_3$): 4.50 (d, 1H, $^{3}J_{H2\text{-}3b}$ = 10.1, $^{3}J_{H2\text{-}3a}$ = 8.3 Hz, CHOH), 4.43 (~td, $^{2}J_{H,}$ $^{3}J_{H4a\text{-}3b}$ = 9.0, $^{3}J_{H4a\text{-}3a}$ = 2.1 Hz, 1H, CH2O-a), 4.23 (ddd, 1H, $^{3}J_{H4b\text{-}3b}$ = 10.2 Hz, $^{2}J_{H}$ = 9.3, $^{3}J_{H4b\text{-}3a}$ = 6.0 Hz, CH2O-b), 2.60 (dddd, 1H, $^{2}J_{H}$ = 12.6, $^{3}J_{H3a\text{-}2}$ = 8.2, $^{3}J_{H3a\text{-}4b}$ = 6.0, $^{2}J_{H3a\text{-}4a}$ = 2.1 Hz, CH2-a), 2.26 ppm (dddd, 1H, $^{2}J_{H}$ = 12.6, $^{3}J_{H3b\text{-}4b}$ = 10.2, $^{3}J_{H3b\text{-}2}$ = 10.1, $^{3}J_{H3b\text{-}4a}$ = 8.7 Hz, CH2-b).

^{13}C NMR: (75 MHz, $CDCl_3$): 178.2 (COO), 67.5 (CHOH), 65.4 (CH2O), 30.9 ppm (CH2).

5.9.4.3 2(*R*)-Hydroxy-γ-butyrolactone (II)

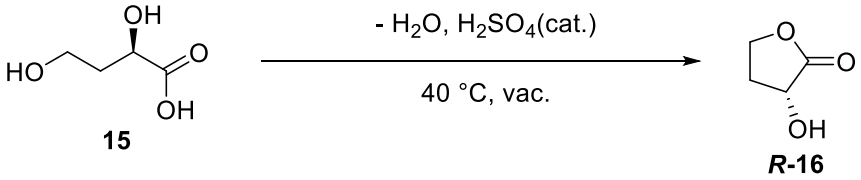


Figure 144: Lactone formation with H_2SO_4.

In a microdistillation apparatus, a droplet of H_2SO_4 (conc.) is added to **15** (2(*R*),4-dihydroxybutyric acid, 470 mg, 2.54 mmol, 65%). After 1 h, the condenser is washed with ACN and the formed emulsion stored overnight at 8 °C. The solvent is removed at reduced pressure, yielding 6 mg (2.5%) of a clear fluid.

TLC (DCM/Et_2O 3 + 1, Seebach reagent):

Rf(2(*R*)-hydroxy-γ-butyrolactone) = 0.34 (blue).

^{1}H NMR: (500 MHz, CDCl3): 4.69 (d, 1H, $^{3}J_{H2\text{-}3b}$ = 10.6, $^{3}J_{H2\text{-}3a}$ = 8.4 Hz, CHOH), 4.51 (~td, $^{2}J_{H,}$ $^{3}J_{H4a\text{-}3b}$ = 9.0, $^{3}J_{H4a\text{-}3a}$ = 1.8 Hz, 1H, CH2O-a), 4.35 (ddd, 1H, $^{3}J_{H4\text{-}3b}$ = 10.6 Hz, $^{2}J_{H}$ = 9.1, $^{3}J_{H4b\text{-}3a}$ = 5.9 Hz, CH2O-b), 2.68 (dddd, 1H, $^{2}J_{H}$ = 12.5, $^{3}J_{H3a\text{-}2}$ = 8.1, $^{3}J_{H3a\text{-}4b}$ = 6.0, $^{2}J_{H3a\text{-}4a}$ = 1.8 Hz, CH2-a), 2.26 ppm (dddd, 1H, $^{2}J_{H}$ = 12.4, $^{3}J_{H3b\text{-}4b}$ = 10.7, $^{3}J_{H3b\text{-}2}$ = 10.4, $^{3}J_{H3b\text{-}4a}$ = 8.7 Hz, CH2-b).

^{13}C NMR: (126 MHz, D_2O, acetonitrile-std: 1.47 ppm): 180.9 (COO), 67.8 (CHOH), 66.9 (CH2O), 31.2 ppm (CH2).

5.9.5 Routes (III): *syn*-Dihydroxylation

5.9.5.1 *syn*-3-Hydroxy-2-methoxytetrahydrofuran-3-ol

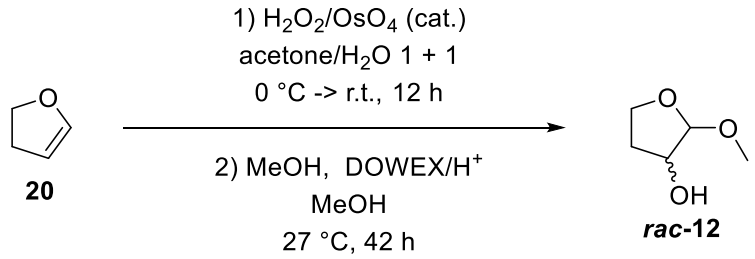


Figure 145: OsO_4-catalyzed *syn*-dihydroxylation and protection as methyl glycoside.

An aluminium foil wrapped round bottom flask (100 mL) is charged with acetone (20 ml, degassed). OsO_4 (10.0 mL, 1.64 mmol, 1.15 mol% eq., 4% aq.) and **20** (2,3-dihydrofurane, 10.8 mL, 143 mmol) are added at 0 °C. H_2O_2 (14.0 mL, 137 mmol, 0.961 eq., 30%, aq.) is added over 2 h under stirring. The suspension is stirred for 10 h at r.t..

Activated charcoal is added and the suspension stirred for 0.5 h. The clear supernatant of the settled suspension is decanted and the remaining suspension filtered over celite (diameter 6 cm, length 0.8 cm) and the filter cake is washed with methanol (10 mL). The combined filtrate is reduced to 15 mL two times and the loss in volume made up with methanol, then reduced to 12 mL. Dry methanol (20 mL), Dowex 50W-X8 (1 g, 1.2 mol% eq. H^+, acidic form, residual water removed by washing with Et_2O and then dried under reduced pressure) and molecular sieves (1.5 g, powder, 4 Å) are added and the mixture incubated (42 h, 27 °C, 300 rpm).

The mixture is filtered, reduced and subjected to flash chromatography (550 g silica gel, 60 Å, 40 – 63 µm, diameter 7 cm, length 30 cm, 1 L DCM/Et_2O 4 + 1, 20 mL fractions collected).

Removal of the eluent yields ***rac*-12** as a clear, colourless liquid (6.04 g, 35%, fractions 22 – 43).

TLC (DCM/Et_2O 4 + 1, anisaldehyde reagent):

R_f(*syn*-3-Hydroxy-2-methoxytetrahydrofuran) = 0.25 (black).

^{1}H NMR: (300 MHz, $CDCl_3$): 4.81 (s, 1H, CHO_2), 4.21 (dd, ^3J_H = 5.6, ^3J_H = 1.6 Hz, CHO), 4.09 (q, 1H, $^{2,3}J_H$ = 7.8 Hz, CH_2O), 3.94 (ddd, 1H, ^3J_H = 9.0, ^3J_H = 4.5, ^2J_H = 8.3 Hz, CH_2O), 3.33 (s, 3H, CH_3O), 2.24 (dddd, 1H, ^2J_H = 13.1, ^3J_H = 9.2,

$^{3}J_{H}$ = 7.4, $^{3}J_{H}$ = 5.6 Hz, CH_2), 1.82 ppm (dddd, 1H, $^{2}J_{H}$ = 13.4, $^{3}J_{H}$ = 7.8, $^{3}J_{H}$ = 4.5, $^{3}J_{H}$ = 1.6 Hz, CH_2).

^{13}C NMR: (75 MHz, $CDCl_3$): 109.1 (CHOO), 75.6 (CHO), 66.3 (CH_2O), 54.6 (CH_3O), 32.5 ppm (CH_2).

5.9.5.2 3-Deoxy-D/L-*glycero*-tetrose (2,4-Dihydroxybutanal, Deprotection)

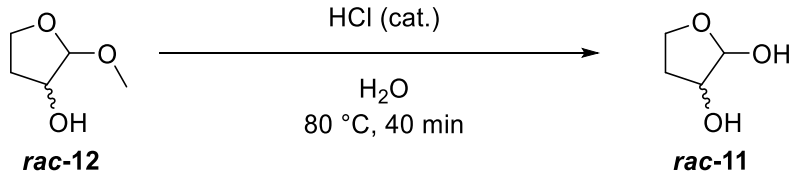


Figure 146: Deprotection of ***rac*-12**.

***rac*-12** (s*yn*-3-Hydroxy-2-mehoxytetrahydrofuran, 3.58 g, 30.3 mmol) is suspended in water (10 mL, d.i.) and HCl (50 µL, 0.48 mmol, 1.6 mol% eq., 37% aq.) is added under shaking. The suspension is warmed to 80 °C and methanol evaporated under reduced pressure (900 mbar) for 40 min.

The formed solution is neutralized with sodium hydrogen carbonate (42 mg, 0.50 mmol) and frozen until usage in transketolase cascade reactions.

For quantification of the concentration, an aliquot (200 µL) is lyophilized and the residue taken up in D_2O (600 µL), containing maleic acid (70.84 mM) as a standard.

The concentration of ***rac*-11** (3-deoxy-D/L-*glycero*-tetrose) is quantified, comparing the ^{1}H NMR areas of the maleic acid peak (s, 6.39 ppm, 2H, CH, area = 2.000) and one of the product's peaks (m, 2.23 ppm, 1H, CH_2, area = 6.832). The calculated concentration equals 1.45 M (equivalent to 48% yield) in the original solution of ***rac*-11**.

TLC (ethyl acetate/^{n}butanol/^{i}propanol/H_2O/acetic acid, 7 + 1 + 1 + 1 + 1, anisaldehyde reagent):

R_f(3-deoxy-D/L-*glycero*-tetrose) = 0.60 (black),

R_f(*syn*-3-hydroxy-3-methoxytetrahydrofuran) = 0.78 (black).

^{1}H NMR (500 MHz, D_2O): 9.69 (s, 1H, CHO, free aldehyde) 6.39 (s, 2H, maleic acid CH), 5.23 (s, 1H, CHOOH, lactol), 4.27 – 4.17 (m, 1H, CHOH), 4.12 – 3.94 (m, 2H, CH_2OH), 2.31 – 2.15 (m, 1H, CH_2), 1.93 – 1.79 ppm (m, 1H, CH_2).

^{13}C NMR (126 MHz, D_2O, methanol at 49.50 ppm, spectrum from previous synthesis, June 2022): 102.3 (C-1HOOH), 75.9 (C-2HOH) 67.2 (C-4H_2O) 31.1 ppm (C-3H_2).

5.10 Synthesis of Non-Natural Amino Acids

5.10.1 N_α-(*tert*-Butoxycarbonyl)N_ε,N_ε,N_ε-trimethyl-L-lysine

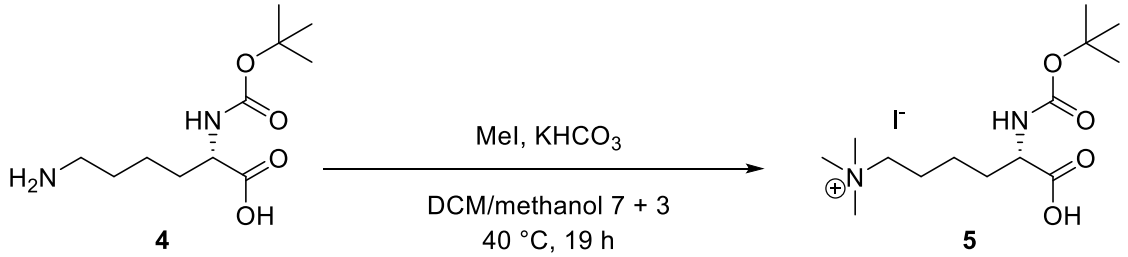


Figure 147: N^ε-Trimethylation of N^α Boc-protected L-lysine.

A two neck round bottom flask (100 mL) with septum, drying tube ($CaCl_2$, anh.) and stirring bar is charged with **4** (N_α-(*tert*-butoxycarbonyl)-L-lysine, 985 mg, 4.00 mmol) in DCM (85 mL, dry) and methanol (35.5 mL, dry) is added. The solution is warmed to 40 °C and $KHCO_3$ (4.02 g, 40.2 mmol, 10.3 eq.) is added. Iodomethane (4.0 mL, 64 mmol, 16 eq.) is dropped (67 µL/min) into the solution over 1 h under stirring.

After 18 h, the mixture is filtered and volatile compounds are removed under reduced pressure. The crude is purified *via* automate flash chromatography (C-18 modified silica gel MultoKrom® 100 – 5 C18, diameter 20 mm, length 150 mm, gradient elution with 10 mL/min, eluent A: water, 0.1% formic acid; eluent B: ACN, 0.1% formic acid, 0 – 5 min: 5% B; 5 – 25 min: linear gradient to 100% B, 30 –32 min: 5% B, 10 mL fractions collected).

Removal of the eluent yields **5** (N_α-(*tert*-butoxycarbonyl)N_ε,N_ε,N_ε-trimethyl-L-lysine) as a colourless, hygroscopic, amorphous solid (1.24 g, 74%, iodide salt [416.3], fractions 16, 17).

TLC (methanol/triethylamine 97 + 3, Dragendorff's reagent):

R_f(N_α-(*tert*-butoxycarbonyl)N_ε,N_ε,N_ε-trimethyl-L-lysine) = 0.07 (orange).

HPLC/MS (method reverse I, mass spectrometric):

t_{ret}(N_α-(*tert*-butoxycarbonyl)N_ε,N_ε,N_ε-trimethyl-L-lysine) = 8.7 min (M^+ = 303),

t_{ret}(N_α-(*tert*-butoxycarbonyl)N_ε,N_ε-dimethyl-L-lysine) = 7.1 min (M^++H = 289),

t_{ret}(N_α-(*tert*-butoxycarbonyl)-L-lysine) = 6.6 min (M^++H = 247, M^--H = 245).

HPLC/MS (method reverse II, mass spectrometric):

t_{ret}(*N*$_\alpha$-(*tert*-butoxycarbonyl)*N*$_\varepsilon$-methyl-L-lysine) = 16.3 min (M^++H = 275, M^--H = 273).

^{1}H NMR (500 MHz, D_2O): 3.98 – 3.83 (m, 1H, CHNH), 3.38 – 3.31 (m, 2H, CH_2N), 3.12 (s, 9H, 3*CH_3N), 1.88 – 1.87 (m, 3H, 2CH_2), 1.76 – 1.66 (m, 1H, CH_2), 1.52 – 1.38 (m, 2H, CH_2), 1.44 ppm (s, 9H, 3CH_3).

^{13}C NMR: (126 MHz, D_2O): 179.4 (COO), 157.6 (COON), 81.0 (CO), 66.4 (t, *J* = 2.9 Hz, CH_2N), 55.5 (CHN), 52.8 (CH_3N), 52.8 (CH_3N), 52.8 (CH_3N), 31.2 (CH_2), 27.7 (3CH_3), 22.1 (CH_2), 21.9 ppm (CH_2).

5.10.2 *N*$_\varepsilon$,*N*$_\varepsilon$,*N*$_\varepsilon$-Trimethyl-L-lysine dihydroformate

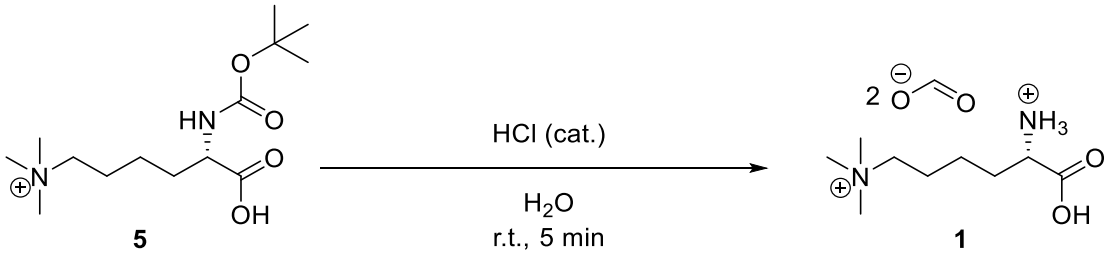


Figure 148: Deprotection of **5**.

HCl (aq. 3 M, 3 mL) is added to **5** (*N*$_\alpha$-(*tert*-butoxycarbonyl)*N*$_\varepsilon$,*N*$_\varepsilon$,*N*$_\varepsilon$-trimethyl-L-lysinium iodide, 1.24 g, 2.98 mmol).

The clear solution is subjected to flash chromatography (50 g silica gel, 60 Å, 40 – 63 µm, diameter 2.5 cm, length 10 cm, 2 L chloroform/methanol/formic acid 1 + 9 + 0.1, 10 mL fractions collected).

Removal of the eluent yields a colourless amorphic solid (480 mg, 58%, dihydroformate salt [276.4], fractions 90 – 160).

TLC (chloroform/methanol/formic acid 1 + 9 + 0.01, Ninhydrin reagent):

R_f(*N*$_\varepsilon$,*N*$_\varepsilon$,*N*$_\varepsilon$-trimethyl-L-lysine) = 0 (purple).

HPLC/MS (method reverse I, mass spectrometric):

t_{ret}(*N*$_\varepsilon$,*N*$_\varepsilon$,*N*$_\varepsilon$-trimethyl-L-lysine) = 1.5 min (M^+ = 189).

^{1}H NMR: (500 MHz, D_2O/methanol-d^5): 8.34 (s, 2H, 2HCOO), 3.78 – 3.71 (m, 1H, $CHNH_2$), 3.32 (dt, 2H, ^2J_H = 12.3, ^3J_H = 3.8 Hz, CH_2N), 3.09 (s, 9H, 3*CH_3N), 1.91 (dt, 2H, ^2J_H = 12.3, ^3J_H = 3.8 Hz, CH_2), 1.88 – 1.79 (m, 2H, CH_2), 1.56 – 1.34 ppm (m, 2H, CH_2).

^{13}C NMR: (126 MHz, D_2O/methanol-d^5): 175.3 (COO), 169.6 (HCOO), 67.0 (t, *J* = 3.1 Hz, CH_2N), 55.3 (CHN), 53.8 (CH_3N), 53.8 (CH_3N), 53.7 (CH_3N), 30.8 (CH_2), 23.0 (CH_2), 22.3 ppm (CH_2).

5.10.3 N_ε-Acetyl-N_α-(((9*H*-fluoren-9-yl)methoxy)carbonyl-D-lysine

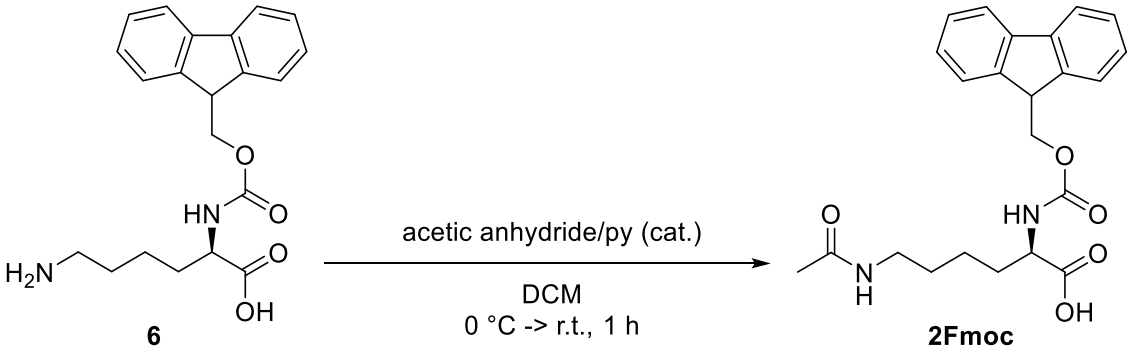


Figure 149: N^ε-Acetylation of N^α-Fmoc-protected D-lysine.

$AlCl_3$ (6.5 mg, 49 µmol, 4.9 mol% eq.) is added to a suspension of **6** (N_α-(((9*H*-fluoren-9-yl)methoxy)carbonyl)-D-lysine, 370 mg, 1.00 mmol) in DCM (5 mL, dry). Acetic anhydride (0.2 mL, 2 mmol, 2 eq.) is added, to the stirred suspension at 0 °C. After 10 min at 0 °C, the mixture is allowed to warm to rt. After 30 min at r.t., pyridine (10 µL, 9.8 mg, 0.12 mmol, 0.12 eq.) is added to the mixture.

After 1 h at r.t., volatile components are removed under reduced pressure, and the crude subjected to flash chromatography (100 g silica gel, 60 Å, 40 – 63 µm, diameter 2.5 cm, length 20 cm, 200 mL chloroform/methanol/formic acid 9 + 1 + 0.2, 10 mL fractions collected).

Removal of the eluent yields **2Fmoc** (N_ε-acetyl-N_α-(((9*H*-fluoren-9-yl)methoxy)carbonyl-D-lysine) as a colourless amorphous solid (370 mg, 90%, fractions 12 – 15).

HPLC/MS (method reverse I, 190 nm, 254 nm, mass spectrometric detection):

t_{ret}(N_α-(((9*H*-fluoren-9-yl)methoxy)carbonyl)-D-lysine) = 10.0 min (M^++H = 369, M^--H = 367),

t_{ret}(N_ε-acetyl-N_α-(((9*H*-fluoren-9-yl)methoxy)carbonyl)-D-lysine) = 13.5 min (M^++H = 411, M^--H = 409).

5.10.4 N^{ε}-Acetyl-D-lysine

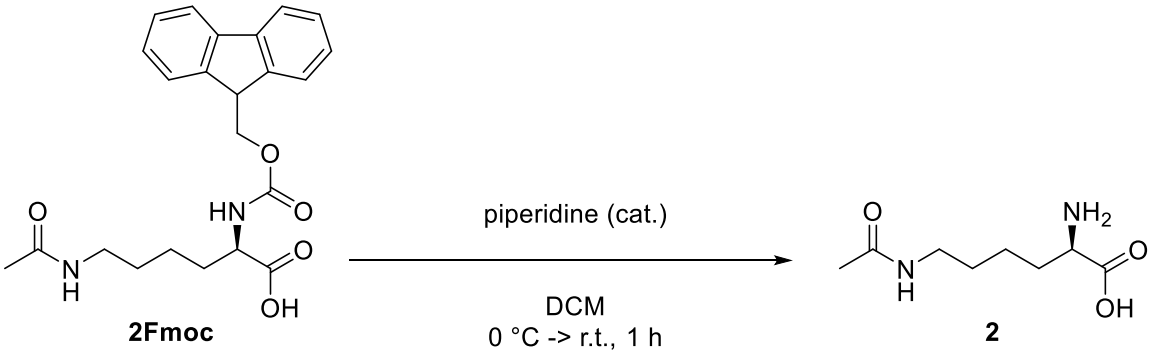


Figure 150: Deprotection of **2Fmoc**.

A portion of **2Fmoc** (N_{ε}-acetyl-N_{α}-(((9*H*-fluoren-9-yl)methoxy)carbonyl-D-lysine, 100 mg, 244 µmol) is dissolved in DCM (0.8 mL) and piperidine (215 mg, 2.5 mmol, 10 eq.) is added under stirring.

After 30 min, the solution is filtered through a short patch of C-18 modified silica gel.

Removal of volatiles yields **2** (N^{ε}-Acetyl-D-lysine) as a colourless amorphous solid (41 mg, 89% deprotection yield, 80% yield over two steps).

TLC (chloroform/methanol/formic acid: 90 + 10 + 2, UV-detection):

R_f(N_{ε}-acetyl-D-lysine) = 0.32 (256 nm).

HPLC/MS (method reverse I, 190 nm, 254 nm, mass spectrometric detection):

t_{ret}(N_{ε}-acetyl-N_{α}-(((9*H*-fluoren-9-yl)methoxy)carbonyl)-D-lysine) = 13.5 min
(M^++H = 411, M^--H = 409),

t_{ret}(N_{ε}-acetyl-D-lysine) = 1.6 min (M^++H = 189, M^--H = 187).

^{1}H NMR: (500 MHz, D_2O): 3.76 – 3.68 (m, 1H, CHN), 3.22 – 3.13 (m, 2H, CH_2N), 1.97(s, 3H, CH_3, 1.60 – 1.49 (m, 2H, CH_2), 1.47 – 1.32 ppm (CH_2).

^{13}C NMR: (126 MHz, $CDCl_3$): 175.3 (COO), 169.6 (HCOO), 67.0 (t, *J* = 3.1 Hz), 55.3 (CHN), 53.8 (CH_3N), 53.8 (CH_3N), 53.7 (CH_3N), 30.8 (CH_2), 23.0 (CH_2), 22.3 ppm (CH_2).

5.10.5 Methyl D-α-aspartyl-D-phenylalaninate

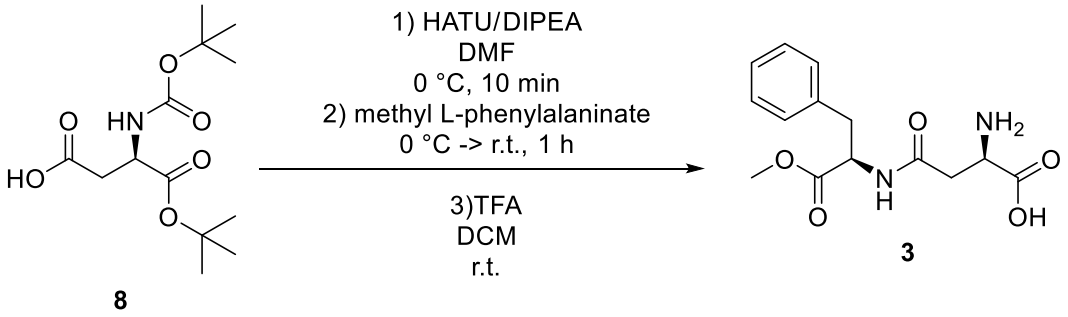


Figure 151: Telescoped reaction sequence of activation, coupling and deprotection.

1-[bis(dimethylamino)methylene]-1*H*-1,2,3-triazolo[4,5-*b*]pyridinium 3-oxide hexafluorophosphate (570 mg, 1.50 mmol, 1.50 eq.) and *N*,*N*-diisopropylethylamine (1 mL, 5.74 mmol, 5.74 eq.) are added to a cooled (0 °C) solution of **8** (α-*tert*-butyl *N*-*tert*-butyloxycarbonyl-D-aspartate, 347 mg, 1.20 mmol, 1.20 eq.) in DCM (5 mL, dry). The mixture is stirred at 0 °C for 10 min and then, methyl L-phenylalaninate hydrochloride (216 mg, 1.00 mmol) is added. The mixture is stirred at 0 °C for 30 min, warmed to rt and stirred for another 30 min.

The mixture is washed with HCl (2*10 mL, 0.1 M). The organic phase is reduced and subjected to automated flash chromatography (C-18 modified silica gel MultoKrom® 100 – 5 C18, diameter 20 mm, length 150 mm, gradient elution with 10 mL/min, eluent A: water, 0.1% formic acid, eluent B: ACN, 0.1% formic acid, 0 – 15 min: 5% B; 15 – 18 min: linear gradient to 12% B, 18 – 19 min: 12% B, 19 – 35 min: linear gradient to 75% B, 35 – 60 min: 75% B, 60 – 65 min: linear gradient to 100% B, 65 –70 min: 5% B, 10 mL fractions collected).

The eluent is removed, the colourless residue (fractions 50 – 52) dissolved in 5 mL of DCM and acidified with trifluoracetic acid (pH 1).

The solvent is removed, the residue dissolved in acetonitrile and subjected to automated flash chromatography (C-18 modified silica gel MultoKrom® 100 – 5 C18, diameter 20 mm, length 150 mm, gradient elution with 10 mL/min, eluent A: water, 0.1% formic acid, eluent B: ACN, 0.1% formic acid, 0 – 8 min: 5% B; 8 – 10 min: linear gradient to 10% B, 10 – 12 min: 10% B, 12 – 17 min: linear gradient to 25% B, 17 – 18 min: 25% B, 18 – 20 min: linear gradient to 30% B, 20 – 23 min: 30% B, 23 – 32 min: linear gradient to 50% B, 32 – 36 min: linear gradient to 100% B, 36 – 40 min: 100% B, 40 –42 min: 5% B, 10 mL fractions collected), yielding **3** (methyl

D-α-aspartyl-L-phenylalaninate) as a colourless, amorphous solid (110 mg, 31%, fractions 25 – 27).

TLC (chloroform/methanol/triethylamine 45 + 55 + 2, Ninhydrin reagent):

R_f(methyl D-α(*tert*-butyl N-*tert*-butyloxycarbonyl-aspartyl-L-phenylalaninate) = 0.75 (red),

R_f(methyl D-α-aspartyl-L-phenylalaninate) = 0.42 (red).

HPLC/MS (method reverse I, mass spectrometric)

t_{ret}(methyl D-α(*tert*-butyl N-*tert*-butyloxycarbonyl-aspartyl-L-phenylalaninate) = 17.9 min (M^++H = 451, M+Na^+ = 473),

t_{ret}(methyl D-α-aspartyl-L-phenylalaninate) = 6.7 min (M^++H = 295, M^--H = 293).

^{1}H NMR (500 MHz, methanol-d^4): 8.61 (d, 1H, ^3J_H = 10.1 Hz, NH), 8.25 - 8.16 (m, 2H, NH_2), 7.32 – 7.24 (m, 2H, CH arom., *m*), 7.24 – 7.18 (m, 3H, CH arom., *o*, *p*), 4.76 – 4.68 (m, 1H, CHN), 3.80 – 3.73 (m, 1H, CHN), 3.67 (s, 3H, CH_3O), 3.19 – 3.10 (m, 1H, CH_2), 3.02 – 2.87 (m, 2H, CH_2), 2.69 ppm (dd, 1H, 16.1, 9.5 Hz, CH_2).

^{13}C NMR (126 MHz, methanol-d^4): 173.5 (COOH), 173.1 (CON), 172.2 (COO), 138.0 (C arom., *i*), 130.3 (CH arom., *m*), 129.5 (CH arom. *o*), 127.9 (CH arom., *p*), 55.3 (CHNH), 52.8 (CH_2), 52.7 (CH_3O), 38.6 (CH_2Ph), 36.1 ppm (CNH_2).

5.11 Hydrophilic Interaction Liquid Chromatography

Two aqueous mixtures of hydrophilic compounds (mixture A: L-serine and lithium hydroxypyruvate, mixture B: L-leucine and L-isoleucine) each compound at 0.5 mg/mL are analysed on HPLC/MS to test the limits of the HILIC method.

HPLC/MS (HILIC method II, mass spectrometric detection):

Mixture A:

t_{ret}(L-serine) = 27.5 min (M^++H = 106, M^--H = 104),

t_{ret}(lithium hydroxypyruvate) = 23.6 min (M^++H = 105, M^--H = 103).

Mixture B:

t_{ret}(leucine isomer A) = 22.0 min (M^++H = 132, M^--H = 130),

t_{ret}(leucine isomer B) = 22.8 min (M^++H = 132, M^--H = 130).

6 References

[1] https://www.fda.gov/regulatory-information/search-fda-guidance-documents/development-new-stereoisomeric-drugs (accessed 01.09.2023).

[2] Yu, J. et al. *Enzym Catal* **2019**, *2019*, 6618 – 6625.

[3] Fischer, E. *Ber Dtsch Chem Ges* **1894**, *27*, 2985 – 2993.

[4] Koshland, D. E. J. *P Natl A Sci USA* **1958**, *44*, 98 – 104.

[5] Juraszek, J. et al. *Nat Commun* **2021**, *12*, 244.

[6] Tai, L. et al. *P Natl A Sci USA* **2021**, *118*, e2112703118.

[7] Li, F. *Ann Rev Virol* **2016**, *3*, 237 – 261.

[8] Jackson, C. B. et al. *Nat Rev Mol Cell Bio* **2022**, *23*, 3 – 20.

[9] Chen, L. H. et al. *J Biol Chem* **1992**, *267*, 17716 – 17721.

[10] Blackmond, D. G. *CSH Perspect Biol* **2019**, *11*, a032540.

[11] Krishnan, P. et al. *BMC Evol Biol* **2018**, *18*, 7.

[12] Hernández, S. B.; Cava, F. *Environ Microbiol* **2016**, *18*, 1673 – 1685.

[13] Ashiuchi, M. et al. *J Biochem* **1998**, *123*, 1156 – 1163.

[14] Ashiuchi, M. *Microb Biotechnol* **2013**, *6*, 664 – 674.

[15] Hashimoto, A.; Oka, T.; Nishikawa, T. *Eur J Neurosci* **1995**, *7*, 1657 – 1663.

[16] Wolosker, H.; Balu, D. T. *Transl Psychiatry* **2020**, *10*, 184.

[17] Wolosker, H. *Sci STKE* **2006**, *356*, pe41.

[18] Usiello, A. et al. *Int J Mol Sci* **2020**, *21*, 8718.

[19] Kalyankar, G. D.; Snell, E. E. *Biochemistry-US* **1962**, *1*, 594 – 600.

[20] Yamada, S. et al. *J Org Chem* **1983**, *1*, 843 – 846.

[21] Marcello, J. A.; Martell, A. E. *J Am Chem Soc* **1982**, *104*, 3441 – 3447.

[22] Farago, M. E.; Matthews, T. *J Chem Soc A* **1969**, 609 – 612.

[23] Toth, K.; Richard, J. P. *J Am Chem Soc* **2007**, *129*, 3013 – 3021.

[24] Watababe, A. et al. **1999**, *274*, 4189 – 4194.

[25] Kurokawa, Y. et al. *J Biochem* **1998**, *124*, 1163 – 1169.

[26] Fenn, T. D. et al. *Biochemistry-US* **2003**, *42*, 5775 – 5783.

[27] Fenn, T. D. et al. *Biochemistry-US* **2005**, *44*, 5317 – 5327.

[28] Wang, E.; Walsh, C. *Biochemistry-US* **1978**, *17*, 1313 – 1321.

[29] Faraci, W. S.; Walsh, C. T. *Biochemistry-US* **1989**, *28*, 431 – 437.

[30] Bernardo-García, N. et al. *Curr Org Chem* **2016**, *20*, 1222 – 1231.

[31] Gogami, Y. et al. *Phytochemistry* **2009**, *70*, 380 – 387.

[32] Yoshimura, T. *Biosci Biotech Bioch* **2022**, *86*, 1161 – 1172.

[33] Shaw, J. P.; Petsko, G. A.; Ringe, D. *Biochemistry-US* **1997**, *36*, 1329 – 1342.

[34] Stamper, C. G. F.; Morollo, A. A.; Ringe, D. *Biochemistry-US* **1998**, *37*, 10438 – 10445.

[35] Morollo, A. A.; Petsko, G. A.; Ringe, D. *Biochemistry-US* **1999**, *38*, 3293 – 3301.

[36] Reardon, D.; Farber, G. K. *FASEB J* **2018**, *9*, 497 – 503.

[37] Espaillat, A. et al. *Acta Crystallogr D* **2014**, *D70*, 79 – 90.

[38] Agmon, N. *Chem Phys Lett* **1995**, *244*, 456 – 462.

[39] Pilone, M. S. *Cell Mol Life Sci* **2000**, *57*, 1732 – 1747.

[40] Hayashi, J. et al. *Appl Environ Microb* **2017**, *83*, e00491-17.

[41] Azam, M. A.; Jayaram, U. *J Enzym Inhib Med Ch* **2016**, *31*, 517 – 526.

[42] Watanabe, A. et al. *J Biol Chem* **2002**, *277*, 19166 – 19172.

[43] Ondrechen, M. J.; Briggs, J. M.; McCammon, J. A. *J Am Chem Soc* **2001**, *123*, 2830 – 2834.

[44] Watanabe, A. et al. *J Biochem* **1999**, *126*, 781 – 786.

[45] Major, D. T.; Gao, J. *J Am Chem Soc* **2006**, *128*, 16345 – 16357.

[46] Faraci, W. S.; Walsh, C. T. *Biochemistry-US* **1988**, *27*, 3267 – 3276.

[47] Warshel, A. et al. *Chem Rev* **2006**, *106*, 3210 – 3235.

[48] Yow, G.-Y. et al. *J Mol Catal B-Enzym* **2003**, *23*, 311 – 319.

[49] Zheng, J. et al. *Science* **2006**, *313*, 1951 – 1955.

[50] Fischer, S. A.; Gunlycke, D. *J Phys Chem B* **2019**, *123*, 5536 – 5544.

[51] Julin, D. A. et al. *Biochemistry-US* **1989**, *28*, 3815 – 3821.

[52] Sun, S.; Toney, M. D. *Biochemistry-US* **1999**, *38*, 4058 – 4065.

[53] Yoshimura, T.; Esaki, N. *J Biosci Bioeng* **2003**, *96*, 103 – 109.

[54] Spies, M. A.; Toney, M. D. *Biochemistry-US* **2003**, *42*, 5099 – 5107.

[55] Espaillat, A. et al. *CSBJ* **2021**, *19*, 1119 – 1126.

[56] Krebs, H. A. *Hoppe Seylers Z Physiol Chem* **1933**, *217*, 191 – 227.

[57] Krebs, H. A. *Biochem J* **1935**, *29*, 1620 – 1644.

[58] Bernheim, F.; Bernheim, M. L. C.; Webster, M. D. *J Biol Chem* **1935**, *110*, 165 – 172.

[59] Conlon, H. D. et al. *Biotechnol Bioeng* **1995**, *46*, 510 – 513.

[60] Liu, H. et al. *Sensor Actuat B-Chem* **2020**, *304*, 127333.

[61] Jiang, H. et al. *Diabetologia* **2021**, *64*, 693 – 706.

[62] Nagata, Y. et al. *Biochem J* **1989**, *257*, 291 – 292.

[63] Errico, F. et al. *Gene* **2006**, *374*, 50 – 57.

[64] Ono, K. et al. *J Neural Transm* **2009**, *116*, 1335 – 1347.

[65] LaRue, T. A.; Spencer, J. F. T. *Can J.Microbiol* **1967**, *13*, 777 – 788.

[66] Mattevi, A. *Biophys Chem* **1998**, *70*, 217 – 222.

[67] Mizutani, H. et al. *J Biochem* **1996**, *120*, 14 – 17.

[68] Tishkov, V. I.; Khoronenkova, S. V. *Biochemistry-Moscow* **2005**, *70*, 40 – 54.

[69] Matrevi, A. et al. *P Natl A Sci USA* **1996**, *93*, 7496 – 7501.

[70] Umhau, S. et al. *P Natl A Sci USA* **2000**, *97*, 12463 – 12468.

[71] Pollegioni, L. et al. *Protein Sci* **2003**, *12*, 1018 – 1029.

[72] Miyano, M. et al. *J Biochem* **1991**, *109*, 171 – 177.

[73] Müller, F. *Chemistry and Biochemistry of Flavoenzymes*, Volume 3.; CRC Press Inc.: Boca Raton, 1992.

[74] Marcotte, P.; Walsh, C. *Biochemistry-US* **1978**, *17*, 2864 – 2868.

[75] Ronchi, S. et al. *J Biol Chem* **1980**, *255*, 6044 – 6046.

[76] Molla, G.; Melis, R.; Pollegioni, L. *Biotechnol Adv* **2017**, *35*, 657 – 668.

[77] Pawelek, P. D. et al. *EMBO J* **2000**, *19*, 4204 – 4215.

[78] Mattevi, A. *Trends Biochem Sci* **2006**, *31*, 276 – 283.

[79] Saam, J. et al. *J Biol Chem* **2010**, *285*, 24439 – 24446.

[80] Todone, F. et al. *Biochemistry-US* **1997**, *2960*, 5853 – 5860.

[81] Kiss, D. J.; Ferenczy, G. G. *Org Biomol Chem* **2019**, *17*, 7973 – 7984.

[82] Pollegioni, L. et al. *J Biol Chem* **1993**, *268*, 13850 – 13857.

[83] Dengler, U. et al. *J Mol Biol* **1997**, *267*, 640 – 660.

[84] Domenech, J.; Ferrer, J. *BBA-Gen Subjects* **2006**, *1760*, 1667– 1674.

[85] Domenech, J.; Nieto, J. M.; Ferrer, J. *J Mol Catal B-Enzym* **2009**, *61*, 168 – 173.

[86] Yum, D.-Y. et al. *J Bacteriol* **1998**, *180*, 5984 – 5988.

[87] Casadaban, M. J. *J Mol Biol* **1976**, *104*, 541 – 555.

[88] Nuñez, M. F. et al. *Biochem J* **2001**, *354*, 707 – 715.

[89] Yun, H. et al. *Biotechnol Progr* **2005**, *21*, 366 – 371.

[90] Kochhar, S. et al. *J Biol Chem* **1992**, *267*, 8499 – 8513.

[91] Vinals, C. et al. *Proteins* **1995**, *21*, 307 – 318.

[92] Razeto, A. et al. *J Mol Biol* **2002**, *318*, 109 – 119.

[93] Meng, H. et al. *Sci Rep-UK* **2016**, *6*, 24887.

[94] Zheng, Z. et al. *PLoS One* **2011**, *6*, e19030.

[95] Zhou, L.; Zuo, Z.; Chow, M. S. S. *J Clin Pharmacol* **2005**, *45*, 1345 – 1359.

[96] Urban, F. J.; Moore, B. S. *J Heterocyclic Chem* **1992**, *29*, 431 – 438.

[97] Maurus, P. B.; Kaeding, C. C. *Oper Techn Sport Med* **2004**, *12*, 158 – 160.

[98] Hasırcı, V. et al. *J Biotechnol* **2001**, *86*, 135 – 150.

[99] Ginjupalli, K. et al. *Polym Degrad Stabil* **2017**, *144*, 520 – 535.

[100] Ammala, A. *Int J Cosmetic Sci* **2013**, *35*, 113 – 124.

[101] Choe, H. et al. *PLoS One* **2014**, *9*, e103111.

[102] Berg, J. M. et al. *Stryer Biochemie*, 8. Auflage.; Springer Verlag GmbH Deutschland: Heidelberg, 2018.

[103] Jurkaš, V. et al. *Eng Microbiol* **2022**, *2*, 100008.

[104] Matelska, D. et al. *BMC Evol Biol* **2018**, *18*, 199.

[105] Holton, S. J. et al. *J Struct Biol* **2013**, *181*, 179 – 184.

[106] Ptitsyn, O. B. *FEBS Lett* **1978**, *93*, 1 – 4.

[107] Clarke, A. R. et al. *Nature* **1986**, *324*, 699 – 702.

[108] Relyea, H. A.; van der Donk, W. A. *Bioorg Chem* **2005**, *33*, 171 – 189.

[109] Antonyuk, S. V. et al. *Acta Crystallogr F* **2009**, *65*, 1209 – 1213.

[110] Wilkinson, J. H. *Isoenzymes*, 2nd ed.; Chapman and Hall Ltd: London, 1970.

[111] Sund, H. *Pyridine Nucleotide Dependent Dehydrogenases*, Softcover; Springer Verlag Berlin: Heidelberg, 1970.

[112] Clarke, A. R. et al. *Biochemistry-US* **1988**, *27*, 1617 – 1622.

[113] Prejanò, M. et al. *ChemPhysChem* **2019**, *20*, 2881 – 2886.

[114] Yoshihara, A. et al. *Appl Microbiol Biot* **2023**, *107*, 233 – 245.

[115] Mitschke, L. et al. *J Biol Chem* **2010**, *285*, 31559 – 31570.

[116] Schenk, G.; Duggleby, R. G.; Nixon, P. F. *Int J Biochem Cell B* **1998**, *30*, 1297 – 1318.

[117] Abdoul-Zabar, J. et al. *Adv Synth Catal* **2013**, *355*, 116 – 128.

[118] Abdoul Zabar, J. et al. *Adv Synth Catal* **2015**, *357*, 1715 – 1720.

[119] Yi, D. et al. *Chem Commun* **2015**, *51*, 480 – 483.

[120] Saravanan, T. et al. *Green Chem* **2017**, *19*, 481 – 489.

[121] Saravanan, T. et al. *Angew Chem Int Edit* **2017**, *56*, 5358 – 5362.

[122] Zhou, C. et al. *ChemBioChem* **2017**, *18*, 455 – 459.

[123] Lorillière, M. et al. *ACS Catal* **2019**, *9*, 4754 – 4763.

[124] Schörken, U.; Sprenger, G. A. *Biochim Biophys Acta* **1998**, *1385*, 229 – 243.

[125] Hsu, L.-J. et al. *Protein Eng Des Sel* **2016**, *29*, 513 – 521.

[126] Fernández, I. F.; Hecquet, L.; Fessner, W.-D. *Adv Synth Catal* **2022**, *364*, 612 – 621.

[127] Planas, F.; McLeish, M. J.; Himo, F. *Beilstein J Org Chem* **2019**, *15*, 145 – 159.

[128] Nauton, L. et al. *Biochemistry-US* **2016**, *55*, 2144 – 2152.

[129] Kochetov, G. A.; Solovjeva, O. N. *Biochim Biophys Acta* **2014**, *1844*, 1608 – 1618.

[130] Jordan, F. *Nat Prod Rep* **2003**, *20*, 184 – 201.

[131] Krissinel, E.; Henrick, K. *J Mol Biol* **2007**, *372*, 774 – 797.

[132] Dai, S. et al. *Nature* **2019**, *573*, 609 – 613.

[133] Asztalos, P. et al. *Biochemistry-US* **2007**, *46*, 12037 – 12052.

[134] Pasquini, M. et al. *BBA-Gen Subjects* **2017**, *1861*, 2132 – 2145.

[135] Hsu, N.-S. et al. *Angew Chem Int Edit* **2018**, *57*, 1802 – 1807.

[136] von Pappenheim, F. R. et al. *Nat Chem Biol* **2020**, *16*, 1237 – 1245.

[137] Mayer, S. F.; Kroutil, W.; Faber, K. *Chem Soc Rev* **2001**, *30*, 332 – 339.

[138] Bell, E. L. et al. *Nat Rev Methods Primers* **2021**, *1*, 46.

[139] Schrittwieser, J. H. et al. *Chem Rev* **2018**, *118*, 270 – 348.

[140] Díaz-Rodríguez, A. et al. *Catal Sci Technol* **2015**, *5*, 1443 – 1446.

[141] Ahmed, M.; Kelly, T.; Ghanem, A. *Tetrahedron* **2012**, *68*, 6781 – 6802.

[142] Kamal, A. et al. *Coordin Chem Rev* **2008**, *252*, 569 – 592.

[143] Engel, J.; Bornscheuer, U. T.; Kara, S. *Org Process Res Dev* **2021**, *25*, 411 – 420.

[144] Serba, C.; Winssinger, N. *Eur J Org Chem* **2013**, *2013*, 4195 – 4214.

[145] Walsh, C. T.; Moore, B. S. *Angew Chem Int Edit* **2019**, *58*, 6846 – 6879.

[146] Tang, S. L. Y.; Smith, R. L.; Poliakoff, M. *Green Chem* **2005**, *7*, 761 – 762.

[147] www.genome.gov/sequencingcostsdata (accessed 05.09.2023).

[148] Arnold, F. H. *Angew Chem Int Edit* **2018**, *57*, 4143 – 4148.

[149] Drienovská, I.; Roelfes, G. *Nat Catal* **2020**, *3*, 20 – 22.

[150] Gregorio, N. E.; Levine, M. Z.; Oza, J. P. *Methods Protoc* **2019**, *2*, 24.

[151] Chowdhury, R.; Maranas, C. D. *AIChE J* **2020**, *66*, e16847.

[152] Kamble, A.; Srinivasan, S.; Singh, H. *Mol Biotechnol* **2019**, *61*, 53 – 59.

[153] Dawson, W. M.; Rhys, G. G.; Woolfson, D. N. *Curr Opin Chem Biol* **2019**, *52*, 102 – 111.

[154] Galanie, S.; Entwistle, D.; Lalonde, J. *Nat Prod Rep* **2020**, *37*, 1122 – 1143.

[155] Joensson, H. N.; Andersson Svahn, H. *Angew Chem Int Edit* **2012**, *51*, 12176 – 12192.

[156] Benítez-Mateos, A. I. et al. *React Chem Eng* **2021**, *6*, 599 – 611.

[157] Federsel, H.-J.; Moody, T. S.; Taylor, S. J. C. *Molecules* **2021**, *26*, 2822.

[158] Hülsewede, D.; Meyer, L.-E.; von Langermann, J. *Chem-Eur J* **2019**, *25*, 4871 – 4884.

[159] Armstrong, F. A. et al. *Chem Rev* **2023**, *123*, 5421 – 5458.

[160] Siedentop, R.; Rosenthal, K. *Int J Mol Sci* **2022**, *23*, 3605.

[161] Huffman, M. A. et al. *Science* **2019**, *366*, 1255 – 1259.

[162] Patel, N. R. et al. *Org Lett* **2020**, *22*, 4659 – 4664.

[163] Fryszkowska, A.; Devine, P. N. *Curr Opin Chem Biol* **2020**, *55*, 151 – 160.

[164] McIntosh, J. A. et al. *ACS Cent Sci* **2021**, *7*, 1980 – 1985.

[165] Burns, M. et al. *Org Process Res Dev* **2017**, *21*, 871 – 877.

[166] You, C. et al. *Biotechnol Bioeng* **2017**, *114*, 1855 – 1864.

[167] Wallerstein, J. S. et al. *Biochim Biophys Acta* **1947**, *1*, 327 – 334.

[168] Jiao, K.; Sun, G.; Zhang, S. *Sci China Ser B* **1998**, *41*, 345 – 352.

[169] Votyakova, T. V; Reynolds, I. J. *Arch Biochem Biophys* **2004**, *431*, 138 – 144.

[170] Karamitros, C. S.; Lim, J.; Konrad, M. *Anal Biochem* **2013**, *445*, 20 – 23.

[171] https://www.sigmaaldrich.com/DE/de/product/sigma/90101 (accessed 06.09.2023).

[172] Barham, D.; Trinder, P. *Analyst* **1972**, *97*, 142 – 145.

[173] Vojinović, V. et al. *J Mol Catal B-Enzym* **2004**, *28*, 129 – 135.

[174] Eisenstaedt, E. *J Org Chem* **1938**, *03*, 153 – 165.

[175] Holt, A.; Palcic, M. M. *Nat Protoc* **2006**, *1*, 2498 – 2505.

[176] Yao, N.; Deng, Z.; Zhou, Y. *Int J Environ An Ch* **2020**, *102*, 3147 – 3160.

[177] Kabasakalian, P. et al. *Clin Chem* **1974**, *20*, 606 – 607.

[178] Berneman, A. et al. *J Microb Biochem Technol* **2010**, *2*, 139 – 146.

[179] Mertsch, A.; Poschenrieder, S.; Fessner, W.-D. *Adv Synth Catal* **2020**, *362*, 5485 – 5495.

[180] Willies, S. C.; White, J. L.; Turner, N. J. *Tetrahedron* **2012**, *68*, 7564 – 7567.

[181] Wild, J. et al. *Mol Gen Genet* **1985**, *198*, 315 – 322.

[182] Snell, E. E.; Rannefeld, A. N. *J Biol Chem* **1945**, *157*, 475 – 489.

[183] Sheludko, Y. V.; Fessner, W.-D. *Curr Opin Struc Biol* **2020**, *63*, 123 – 133.

[184] Burstein, E. A.; Vedenkina, N. S.; Ivkova, M. N. *Photochem Photobiol* **1973**, *18*, 263 – 279.

[185] Yokoigawa, K. et al. *J Mol Catal B-Enzym* **2001**, *12*, 27 – 35.

[186] Inagaki, K. et al. *Biochemistry-US* **1986**, *25*, 3268 – 3274.

[187] Al Temimi, A. H. K. et al. *Chem Commun* **2016**, *52*, 12849 – 12852.

[188] Einhorn, A.; Hollandt, F. *Justus Liebigs Ann Chem* **1898**, *301*, 95 – 115.

[189] Miranda, L. P.; Alewood, P. F. *P Natl A Sci USA* **1999**, *96*, 1181 – 1186.

[190] Carpino, L. A. et al. *Angew Chem Int Edit* **2002**, *41*, 441 – 445.

[191] Albericio, F. et al. *Org Prep Proced Int* **2001**, *33*, 203 – 313.

[192] de Chiara, C. et al. *Nat Chem Biol* **2020**, *16*, 686 – 694.

[193] Simonetta, M. P.; Vanoni, M. A.; Casalin, P. *Biochim Biophys Acta* **1987**, *914*, 136 – 142.

[194] Gabler, M.; Hensel, M.; Fischer, L. *Enzyme Microb Technol.* **2000**, *27*, 605 – 611.

[195] Pollegioni, L. et al. *Biochem J* **1995**, *310*, 577 – 583.

[196] Kuan, I.-C. et al. *Biochem Eng J* **2010**, *49*, 408 – 413.

[197] Lee, Y.-H.; Chu, W.-S. *Process Biochem* **1998**, *33*, 461 – 467.

[198] Lee, Y.-H.; Chu, W.-S. *Lett Appl Microbiol* **1996**, *23*, 283 – 286.

[199] Wang, S.-J. et al. *Biotechnol Lett* **2008**, *30*, 1415 – 1422.

[200] Pilone, M. S.; Pollegioni, L.; Buto, S. *Biotechnol Appl Bioc* **1992**, *16*, 252 – 262.

[201] Boselli, A. et al. *Biochimie* **2007**, *89*, 360 – 368.

[202] Pollegioni, L.; Falbo, A.; Pilone, M. S. *Biochim Biophys Acta* **1992**, *1120*, 11 – 16.

[203] Liu, D. et al. *Microbiology* **2015**, *84*, 202 – 209.

[204] Muhammad, M. et al. *Pol J Microbiol* **2019**, *68*, 331 – 341.

[205] Sun, X. et al. *PLoS One* **2015**, *10*, e0133516.

[206] Meziane-Cherif, D. et al. *mBio* **2015**, *6*, e00806-15.

[207] Abe, K. et al. *J Biochem* **2006**, *139*, 235 – 244.

[208] Shibata, K. et al. *Comp Biochem Phys B* **2003**, *134*, 307 – 314.

[209] Wu, H.-M. et al. *PLoS One* **2012**, *7*, e48301.

[210] Kuan, Y.-C. et al. *Process Biochem* **2011**, *46*, 1914 – 1920.

[211] Adachi, M. et al. *Amino Acids* **2019**, *51*, 331 – 343.

[212] Lim, Y.-H. et al. *J Bacteriol* **1993**, *175*, 4213 – 4217.

[213] Tanaka, H. et al. *J Biol Chem* **2011**, *286*, 27548 – 27558.

[214] Uhl, M. K. et al. *PLoS One* **2015**, *10*, e0124056.

[215] von Borzyskowski, L. S. et al. *Nature* **2019**, *575*, 500 – 504.

[216] Chow, M. A. et al. *Biochemistry-US* **2004**, *43*, 12780 – 12787.

[217] Fukumoto, M. et al. *Biosci Biotech Bioch* **2012**, *76*, 1275 – 1284.

[218] Holland, C. K. et al. *Plant J* **2018**, *94*, 304 – 314.

[219] Chang, C.-Y. et al. *Angew Chem Int Edit* **2014**, *53*, 1943 – 1948.

[220] Rudden, M. et al. *Sci Rep-UK* **2020**, *10*, 12500.

[221] Rossi, F. et al. *J Biol Chem* **2004**, *279*, 50214 – 50220.

[222] Blahut, M. et al. *J Biol Chem* **2019**, *294*, 12444 – 12458.

[223] Stammer, C. H. et al. *J Med Chem* **1970**, *13*, 1013 – 1015.

[224] Panizutti, R. et al. *P Natl A Sci USA* **2001**, *98*, 5294 – 5299.

[225] Saier, M. H. J.; Jenkins, W. T. *J Biol Chem* **1967**, *242*, 91 – 100.

[226] Svilenov, H. L. et al. *Mol Pharm* **2020**, *17*, 2638 – 2647.

[227] Sonia, F.; Loredano, P.; Mirella, P. S. *Enzyme Microb Technol* **2001**, *29*, 407 – 412.

[228] Schwede, T. et al. *Nucleic Acids Res* **2003**, *31*, 3381 – 3385.

[229] Asano, Y.; Endo, K. *Appl Microbiol Biot* **1988**, *29*, 523 – 527.

[230] Simonetta, M. P. et al. *Eur J Biochem* **1989**, *180*, 199 – 204.

[231] Chebotareva, N. A.; Roman, S. G.; Kurganov, B. I. *Biophys Rev* **2016**, *8*, 397 – 407.

[232] Dolinsky, T. J. et al. *Nucleic Acids Res* **2004**, *32*, W665 – W667.

[233] L'Enfant, M. et al. *Adv Synth Catal* **2019**, *361*, 2550 – 2558.

[234] Pollegioni, L. et al. *Biotechnol Progr* **2004**, *20*, 467 – 473.

[235] Wang, W. *Int J Pharm* **1999**, *185*, 129 – 188.

[236] Maa, Y.-F.; Hsu, C. C. *Biotechnol Bioeng* **1997**, *54*, 503 – 512.

[237] Thomas, C. R.; Geer, D. *Biotechnol Lett* **2011**, *33*, 443 – 456.

[238] Butcher, D. et al. *Rapid Commun Mass Sp* **2019**, *33*, 399 – 404.

[239] Bes, M. . T. et al. *J Mol Catal A-Chem* **1995**, *98*, 161 – 169.

[240] Findrik, Z.; Valentović, I.; Vasić-Rački, Đ. *Appl Biochem Biotech* **2014**, *172*, 3092 – 3105.

[241] Findrik, Z.; Vrsalović Presečki, A.; Vasić-Rački, Đ. *Biochem Eng J* **2012**, *60*, 91 – 98.

[242] U.S. Geological Survey, September 2006, Collection of Water Samples (ver. 2.0): U.S. Geological Survey Techniques of Water-Ressources Investigations, book 9, chap. A4, accessed July 2022, from http://pubs.water.usgs.gov/twri9A.

[243] Bok, F.; Moog, H. C.; Brendler, V. *Front Nucl Eng* **2023**, *2*, 1158109.

[244] Trost, E.-M.; Fischer, L. *J Mol Catal B-Enzym* **2002**, *19 – 20*, 189 – 195.

[245] Switala, J.; Loewen, P. C. *Arch Biochem Biophys* **2002**, *401*, 145 – 154.

[246] Hofmeister, F. *Arch Exp Pathol Phar* **1888**, *24*, 247 – 260.

[247] Cacace, M. G.; Landau, E. M.; Ramsden, J. J. *Q Rev Biophys* **1997**, *30*, 241 – 277.

[248] Gilliland, G. L. *J Cryst Growth* **1988**, *90*, 51 – 59.

[249] Clarke, A. R. et al. *BBA-Protein Struct M* **1986**, *870*, 112 – 126.

[250] Fushinobu, S. et al. *J Biol Chem* **1996**, *271*, 25611 – 25616.

[251] Buchner, R.; Hefter, G. T.; May, P. M. *J Phys Chem A* **1999**, *103*, 1 – 9.

[252] Kurganov, B. I. et al. *Biochemistry-US* **2000**, *39*, 13144 – 13152.

[253] Dib, I. et al. *Biotechnol Bioeng* **2006**, *94*, 645 – 654.

[254] Guisán, J. M. *Enzyme Microb Technol* **1988**, *10*, 375 – 382.

[255] Fernandez-Lafuente, R. et al. *Enzyme Microb Technol* **1993**, *15*, 546 – 550.

[256] Fernández-Lafuente, R.; Rodriguez, V.; Guisán, J. M. *Enzyme Microb Technol* **1998**, *23*, 28 – 33.

[257] Pedroche, J. et al. *Enzyme Microb Technol* **2007**, *40*, 1160 – 1166.

[258] Mateo, C. et al. *Enzyme Microb Technol* **2007**, *40*, 1451 – 1463.

[259] Harris, C. M. et al. *J Biol Chem* **1999**, *274*, 36233 – 36240.

[260] Boselli, A. et al. *Eur J Biochem* **2002**, *269*, 4762 – 4771.

[261] Pilone, M. S.; Pollegioni, L. *Biocatal Biotransfor* **2002**, *20*, 145 – 159.

[262] Piubelli, L. et al. *Protein Eng* **2003**, *16*, 1063 – 1069.

[263] Bommer, M.; Ward, J. M. *Anal Biochem* **2016**, *493*, 8 – 10.

[264] Winkler, L. W. *Ber Dtsch Chem Ges* **1888**, *21*, 2843 – 2854.

[265] Carpenter, J. I. *Limnol Oceanogr* **1965**, *10*, 135 – 140.

[266] Pai, S.-C., Gong, G.-C., Liu, K.-K. *Mar Chem* **1993**, *41*, 343 – 351.

[267] Horecker, B. L.; Smyrniotis, P. Z. *J Am Chem Soc* **1953**, *75*, 1009 – 1010.

[268] Fiedler, E. et al. *J Biol Chem* **2001**, *276*, 16051 – 16058.

[269] Wohlgemuth, R. *J Mol Catal B-Enzym* **2009**, *61*, 23 – 29.

[270] Yi, D. et al. *ChemBioChem* **2012**, *13*, 2290 – 2300.

[271] Casajus, H. et al. *ChemCatChem* **2020**, *12*, 5772 – 5779.

[272] Hedrick, J. L.; Sallach, H. J. *J Biol Chem* **1961**, *236*, 1867 – 1871.

[273] Hedrick, L.; Sallach, H. J. *Arch Biochem Biophys* **1964**, *105*, 261 – 269.

[274] Meshalkina, L. E. et al. *Biochemistry-Moscow* **2009**, *74*, 362 – 371.

[275] Solovjeva, O. N. et al. *Bioscience Rep* **2020**, *40*, BSR20180246.

[276] Villegas-Torres, M. F. et al. *Enzyme Microb Technol* **2015**, *81*, 23 – 30.

[277] Bawn, M. et al. *Enzyme Microb Technol* **2018**, *116*, 16 – 22.

[278] Lorillière, M. et al. *Green Chem* **2016**, *19*, 425 – 435.

[279] Lorillière, M. et al. *ChemCatChem* **2019**, *12*, 812 – 817.

[280] Lee, S. et al. *J Mol Catal B-Enzym* **1999**, *6*, 369 – 377.

[281] Tatsunami, R.; Yoshioka, T. *J Agr Food Chem* **2006**, *54*, 590 – 596.

[282] Ingram, C. U. *Biocatalytic Synthesis of Amino-Alcohols Using a de novo Designed Transketolase-β-Alanine: Pyruvate Transaminase Pathway in Escherichia coli* Department of Biochemical Engineering, University College London: London, 2006.

[283] Yu, H. et al. *FEBS J* **2020**, *287*, 1758 – 1776.

[284] Ranoux, A. et al. *ChemBioChem* **2012**, *13*, 1921 – 1931.

[285] Hecquet, L. et al. *J Mol Catal B-Enzym* **2001**, *11*, 771 – 776.

[286] Galman, J. L. et al. *Chem Commun* **2010**, *46*, 7608 – 7610.

[287] Ranoux, A.; Hanefeld, U. *Top Catal* **2013**, *56*, 750 – 764.

[288] Payongsri, P. et al. *Enzyme Microb Technol* **2015**, *71*, 45 – 52.

[289] Affaticati, P. E. et al. *Sci Rep-UK* **2016**, *6*, 35716.

[290] Dickens, F.; Williamson, D. H. *Nature* **1958**, *181*, 1790 – 1790.

[291] Sprenger, G. A. et al. *Eur J Biochem* **1995**, *230*, 525 – 532.

[292] Villafranca, J. J.; Axelrod, B. *J Biol Chem* **1971**, *246*, 3126 – 3131.

[293] Demuynck, C. et al. *Tetrahedron Lett* **1991**, *32*, 5085 – 5088.

[294] Ziegler, T.; Straub, A.; Effenberger, F. *Angew Chem Int Edit* **1988**, *27*, 716 – 717.

[295] Myles, D. C.; Andrulis III, P. J.; Whitesides, G. M. *Tetrahedron Lett* **1991**, *32*, 4835 – 4838.

[296] Morris, K. G. et al. *Tetrahedron-Asymmetr* **1996**, *7*, 2185 – 2188.

[297] Hecquet, L.; Bolte, J.; Demuynck, C. *Tetrahedron* **1994**, *50*, 8677 – 8684.

[298] Cárdenas-Fernández, M. et al. *Org Biomol Chem* **2021**, *19*, 6493 – 6500.

[299] Zimmermann, F. *Chemo-enzymatische Synthesen mit der Transketolase aus Escherichia coli*, Fakultät für Mathematik, Informatik und Naturwissenschaften, Rheinisch-Westfälische Technische Hochschule Aachen: Aachen, 1999.

[300] Kobori, Y.; Myles, D. C.; Whitesides, G. M. *J Org Chem* **1992**, *57*, 5899 – 5907.

[301] Milas, N. A. et al. *J Am Chem Soc* **1950**, *81*, 4730 – 4733.

[302] Nicolao, K. C. et al. *Chem-Eur J* **2004**, *10*, 5581 – 5606.

[303] Friedova, N. et al. *Basic Clin Phamacol* **2020**, *127*, 429 – 433.

[304] Fringuelli, F. et al. *Synth Commun* **1943**, *19*, 1939 – 1943.

[305] Clayden, J.; Greeves, N.; Warren, S. *Organische Chemie*, 2. Auflage.; Springer Verlag Berlin: Heidelberg, 2013.

[306] Ireland, R. E. et al. *J Org Chem* **1969**, *34*, 3717 – 3729.

[307] Gonzalez, C. et al. *Carbohyd Res* **2016**, *432*, 17 – 22.

[308] Brown, H. C.; Kim, S. C.; Krishnamurthy, S. *J Org Chem* **1980**, *45*, 1 – 12.

[309] Burkhardt, E. R.; Matos, K. *Chem Rev* **2006**, *106*, 2617 – 2650.

[310] Brown, H. C.; Mead, E. J.; Subba Rao, B. C. *J Am Chem Soc* **1955**, *77*, 6209 – 6213.

[311] Suzuki, Y.; Kaneno, D.; Tomoda, S. *J Phys Chem A* **2009**, *113*, 2578 – 2583.

[312] Bikiel, D. et al. *Inorg Chem* **2005**, *44*, 5286 – 5292.

[313] Krishnamurthy, S. *J Organomet Chem* **1978**, *156*, 171 – 181.

[314] Dahanukar, V. H.; Rychnovsky, S. D. *J Org Chem* **1996**, *61*, 8317 – 8320.

[315] Rychnovsky, S. D.; Bax, B. M. *Tetrahedron Lett* **2000**, *41*, 3593 – 3596.

[316] Faber, K. *Biotransformations in Organic Chemistry: A Textbook*, 7th ed.; Springer-Verlag Berlin: Heidelberg, 2018.

[317] Wittrup Larsen, M. et al. *ChemBioChem* **2010**, *11*, 796 – 801.

[318] Plattner, J. J. et al. *J Org Chem* **1974**, *39*, 303 – 311.

[319] Caro, Y.; Masaguer, C. F.; Ravina, E. *Tetrahedron-Asymmetr* **2001**, *12*, 1723 – 1726.

[320] Hernandez, K. et al. *ACS Catal* **2017**, *7*, 1707 – 1711.

[321] Katulić, M. C. et al. *Ind Eng Chem Res* **2021**, *60*, 13846 – 13858.

[322] Mu, Q. et al. *Metab Eng* **2021**, *67*, 321 – 329.

[323] https://www.bldpharm.com/products/56-45-1.html (accessed 09.09.2023).

[324] https://www.bldpharm.com/products/312-84-5.html (accessed 09.09.2023).

[325] https://www.bldpharm.com/products/302-84-1.html (accessed 09.09.2023).

[326] Ocal, N. et al. *Adv Synth Catal* **2022**, *356*, 78 – 87.

[327] Good, N. E. et al. *Biochemistry-US* **1966**, *5*, 467 – 477.

[328] Carballares, D.; Rocha-Martin, J.; Fernandez-Lafuente, R. *Catalysts* **2022**, *12*, 1009.

[329] Zhang, L. et al. *J Chromatogr B* **2003**, *793*, 159 – 165.

[330] Bailly, C. et al. *Bioorg Med Chem* **2020**, *28*, 115463.

[331] Gerlach, T.; Nugroho, D. L.; Rother, D. *ChemCatChem* **2021**, *13*, 2398 – 2406.

[332] Rale, M. et al. *Chem-Eur J* **2011**, *17*, 2623 – 2632.

[333] Yang, J. et al. *J Agr Food Chem* **2016**, *64*, 9497 – 9505.

[334] Velasquez, M.; Santamaria, A.; Batiot-Dupeyrat, C. *Appl Catal B-Environ* **2014**, *160 – 161*, 606 – 613.

[335] da Silva, G. P.; Mack, M.; Contiero, J. *Biotechnol Adv* **2009**, *27*, 30 – 39.

[336] Balat, M.; Balat, H. *Appl Energ* **2010**, *87*, 1815 – 1835.

[337] Liu, K. K.-C.; Pederson, R. L.; Wong, C.-H. *J Chem Soc Perk T 1* **1991**, *1*, 2669 – 2673.

[338] Marino, C.; Bordino, A. V. *Org Biomol Chem* **2022**, *20*, 934 – 962.

[339] Martin, O. R.; Korppi-Tommola, S.-L.; Szarek, W. A. *Can J Chem* **1982**, *60*, 1857 – 1862.

[340] Smith, A. et al. *Brit J Nutr* **2022**, *128*, 389 – 406.

[341] Bode, L. *Glycobiology* **2012**, *22*, 1147 – 1162.

[342] Wong, C.-H. et al. *Angew Chem Int Edit* **1995**, *34*, 412 – 432.

[343] Doores, K. J.; Gamblin, D. P.; Davis, B. G. *Chem-Eur J* **2006**, *12*, 656 – 665.

[344] Zhao, Z.; Liu, H. *J Org Chem* **2001**, *66*, 6810 – 6815.

[345] Barclay, T. et al. *Carbohyd Res* **2012**, *347*, 136 – 141.

[346] https://www.sigmaaldrich.com/DE/de/search/l-phenylalanine?focus=products&page=1&perpage=30&sort=relevance&term=l-phenylalanine&type=product (accessed 14.09.2023).

[347] https://www.sigmaaldrich.com/DE/de/search/phenylpyruvat?focus=products&page=1&perpage=30&sort=relevance&term=phenylpyruvat&type=product (accessed 11.09.2023).

[348] https://www.sigmaaldrich.com/DE/de/product/aldrich/198994 (accessed 11.09.2023).

[349] https://www.sigmaaldrich.com/DE/de/search/l-valin?focus=products&page=1&perpage=30&sort=relevance&term=l-valin&type=product (accessed 11.09.2023).

[350] https://www.sigmaaldrich.com/DE/de/search/ketoglutarat?focus=products&page=1&perpage=30&sort=relevance&term=ketoglutarat&type=product (accessed 22.12.2023).

[351] https://www.sigmaaldrich.com/DE/de/search/glutaminsäure?focus=products&page=1&perpage=30&sort=relevance&term=glutaminsäure&type=product (accessed 09.09.2023).

[352] Orrego, A. H. et al. *ChemCatChem* **2018**, *10*, 3002 – 3011.

[353] Lundqvist, T. et al. *Nature* **2007**, *447*, 817 – 822.

[354] Fisher, S. L. *Microb Biotechnol* **2008**, *1*, 345 – 360.

[355] Francotte, E.; Lindner, W. *Chirality in Drug Research*, Volume 33; Wiley-VCH: Weinheim, 2006.

[356] https://www.sigmaaldrich.com/DE/de/search/hydroxyphenyllactic-acid?focus=products&page=1&perpage=30&sort=relevance&term=hydroxyphenyllactic acid&type=product (accessed 14.09.2023).

[357] Schrittwieser, J. H.; Velikogne, S.; Kroutil, W. *Catalysts* **2018**, *8*, 205.

[358] Stuhr-Hansen, N.; Padrah, S.; Strømgaard, K. *Tetrahedron Lett* **2014**, *55*, 4149 – 4151.

[359] Hummel, W.; Gröger, H. *J Biotechnol* **2014**, *191*, 22 – 31.

[360] Fukuzumi, S.; Lee, Y.-M.; Nam, W. *J Inorg Biochem* **2019**, *199*, 110777.

[361] Saba, T. et al. *Catal Today* **2020**, *339*, 281 – 288.

[362] Schütte, H.; Hummel, W.; Kula, M. R. *Appl Microbiol Biot* **1984**, *19*, 167 – 176.

[363] Lu, H. et al. *Process Biochem* **2018**, *72*, 119 – 123.

[364] Wichmann, R. et al. *Ann NY Acad Sci* **1984**, *434*, 87 – 90.

[365] Bodlenner, A. et al. *Tetrahedron* **2009**, *65*, 7752 – 7755.

[366] Krix, G. et al. *J Biotechnol* **1997**, *53*, 29 – 39.

[367] Busto, E. et al. *ACS Catal* **2016**, *6*, 2393 – 2397.

[368] *Anfärbereagenzien Für Dünnschicht Und Papier-Chromatographie*; Merck AG:

Darmstadt, 1964.

[369] Karlsson, J.-O. et al. *Anal Biochem* **1994**, *219*, 144 – 146.

[370] Baker, P. J. et al. *Structure* **1995**, *3*, 693 – 705.

[371] Akita, H. et al. *Acta Crystallogr D* **2015**, *D71*, 1136 – 1146.

[372] Akita, H. et al. *Front Microbiol* **2018**, *9*, 1760.

[373] Niefind, K.; Hecht, H. J.; Schomburg, D. *J Mol Biol* **1995**, *251*, 256 – 281.

7 Appendix

7.1 SDS-Page

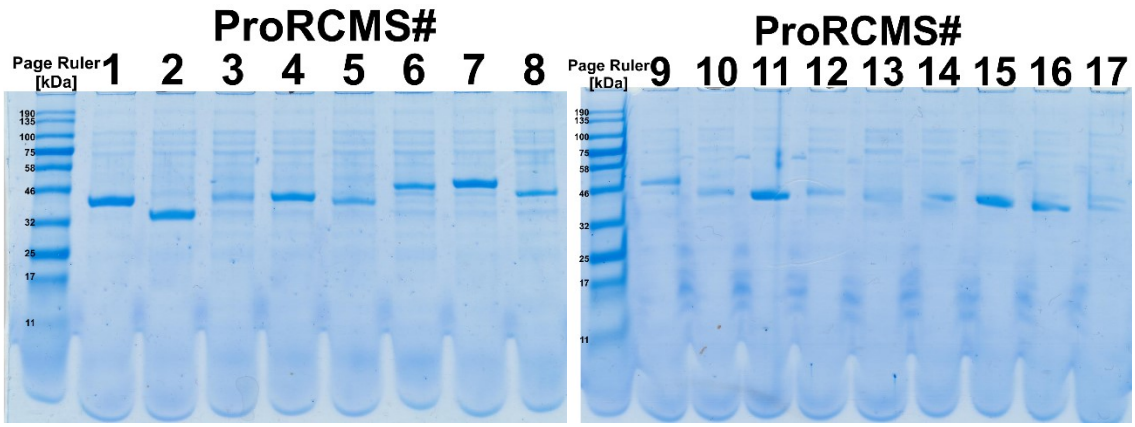


Figure 152: Prozomix ProRCMS racemase panel #1 - 17, cell-free extract, 3 µg (BCA assay) of soluble protein per lane.

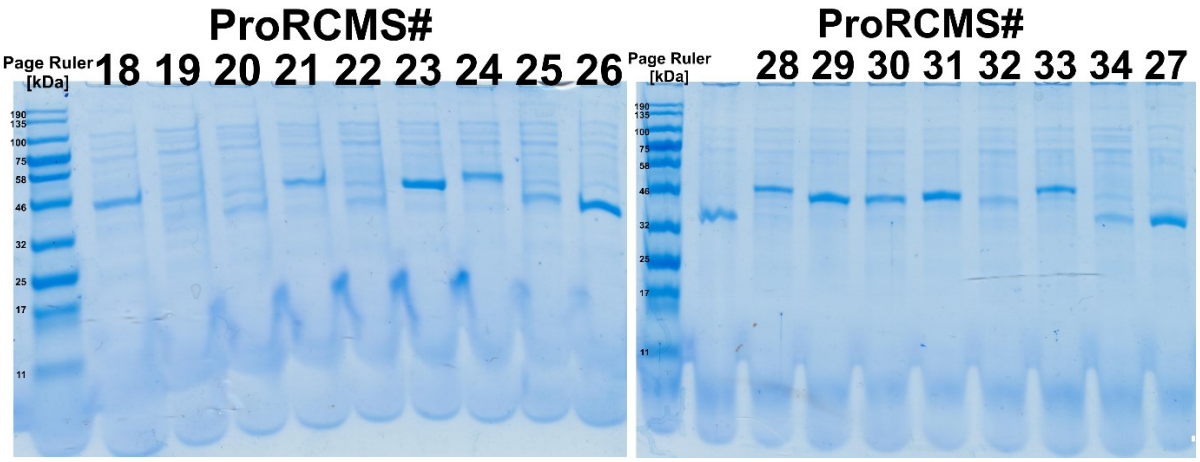


Figure 153: Prozomix ProRCMS racemase panel #18 - 34, cell-free extract, 3 µg (BCA assay) of soluble protein per lane.

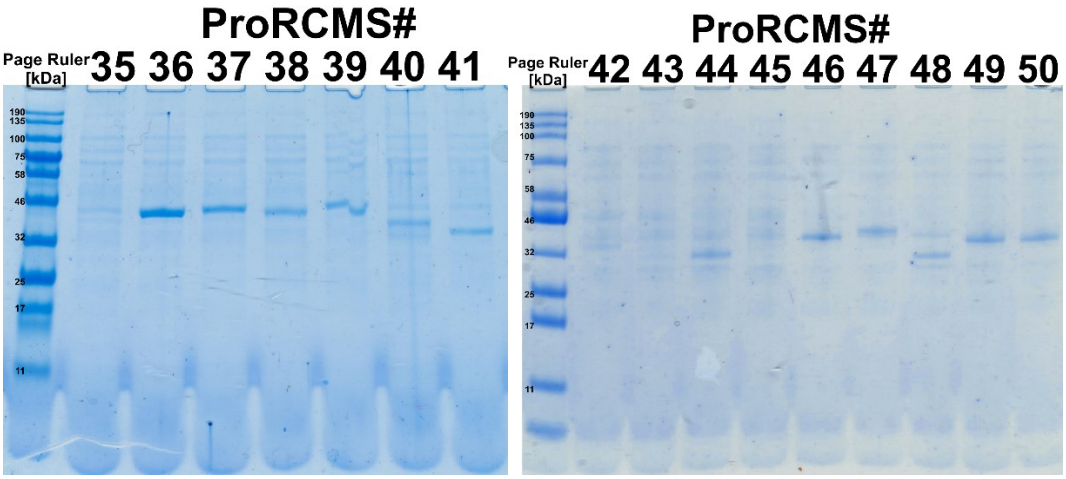


Figure 154: Prozomix ProRCMS racemase panel #35 - 50, cell-free extract, 3 µg (BCA assay) of soluble protein per lane.

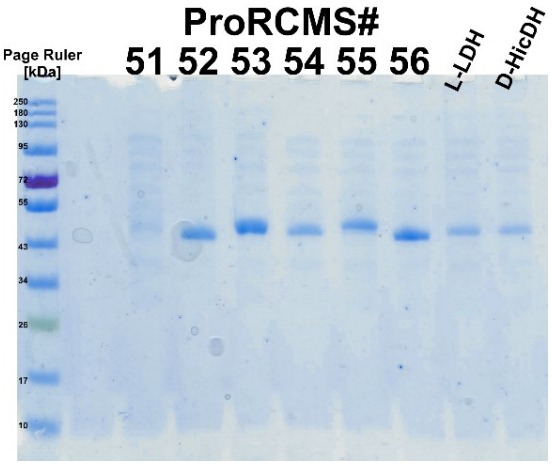


Figure 155: Prozomix ProRCMS racemase panel #51 – 56, cell-free extract; L-leucine dehydrogenase (L-LeuDH) from *Bacillus stearothermophilus* no His Tag, pdbid: 1LEH, [370] expressed in E. coli BL21 DE3, cell-free extract, Prozomix Ltd.; and *meso*-diaminopimelate dehydrogenase (mDAPDH) from *Ureibacillus thermosphaericus* Mutant (Q154L-D158G-T173I-R199M-H249N), pdbid: 3WYB, [371,372] expressed in E. coli BL21 DE3, cell-free extract, Prozomix Ltd.; 3 µg (BCA assay) of soluble protein per lane.

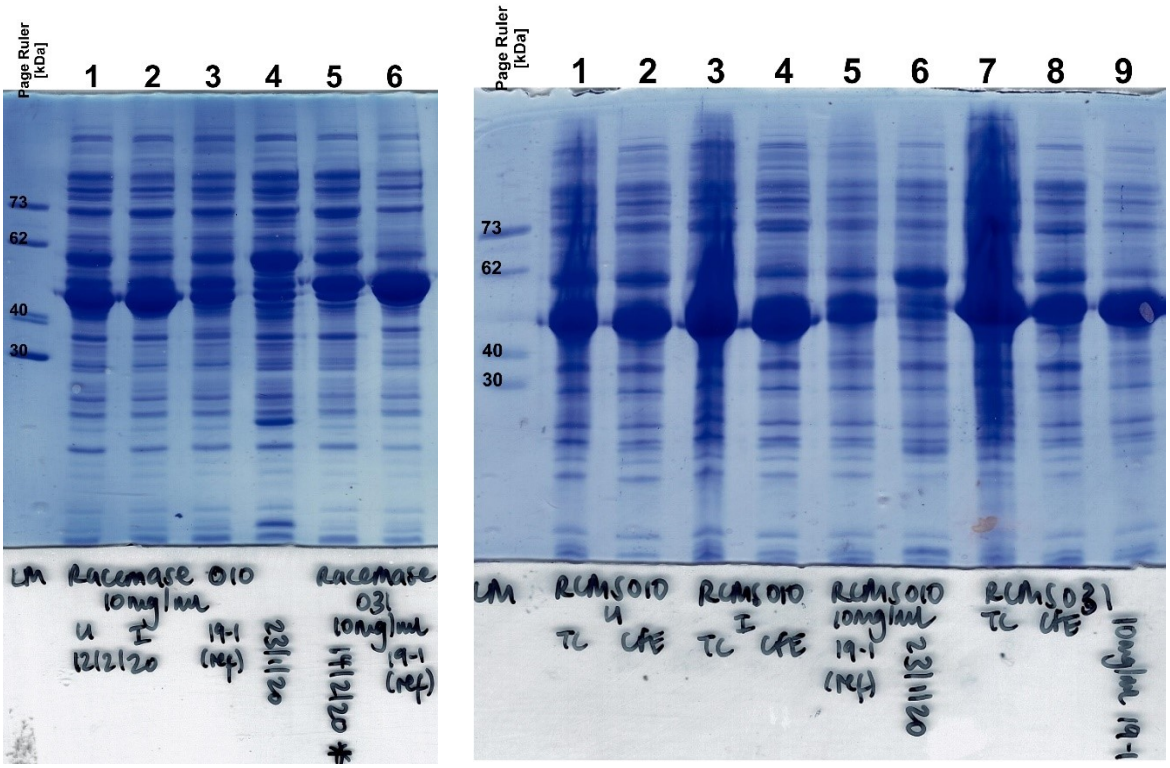


Figure 156: Prozomix ProRCMS racemase panel #10 u 31, large-scale production. Left: 5 µL of 10 mg/mL cell-free extract powder; 1: ProRCMS#10, batch a, 12.2.2020; 2: ProRCMS#10, batch b, 12.2.2020; 3: ProRCMS#10, 19.1.2020; 4: ProRCMS#10, 23.1.2020; 5: ProRCMS#31, 14.2.2020, sample freeze dried twice; 6: ProRCMS#31, 19.1.2020. Right: 1: ProRCMS#10 , batch a, 12.2.2020, 2 µL of total cell; 2: ProRCMS#10 , batch a, 12.2.2020, 2 µL of cell-free extract (post expression); 3: ProRCMS#10 , batch b, 12.2.2020, 2 µL of total cell; 4: ProRCMS#10 , batch b, 12.2.2020, 2 µL of cell-free extract (post expression); 5: ProRCMS#10, 19.1.2020, 5 µL of 10 mg/mL cell-free extract powder; 6: ProRCMS#10, 23.1.2020, 5 µL of 10 mg/mL cell-free extract powder; 7: ProRCMS#31, 12.2.2020, 2 µL of total cell; 8: ProRCMS#31 , 12.2.2020, 2 µL of cell-free extract (post expression) 9: ProRCMS#31, 19.1.2020, 5 µL of 10 mg/mL cell-free extract powder. Analyzed by Prozomix Ltd..

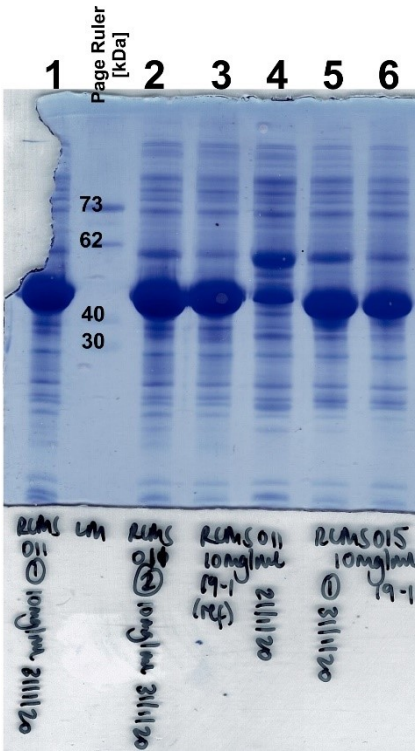


Figure 157: Prozomix ProRCMS racemase panel #11 u 15, large-scale production. 1: ProRCMS#11, batch a, 31.1.2020, 5 µL of 10 mg/mL cell-free extract powder; 2: ProRCMS#11, batch b, 31.1.2020, 5 µL of 10 mg/mL cell-free extract powder; 3: ProRCMS#11, 19.1.2020, 5 µL of 10 mg/mL cell-free extract powder; 4: ProRCMS#11, 21.1.2020, 5 µL of 10 mg/mL cell-free extract powder; 5: ProRCMS#15, 31.1.2020, 5 µL of 10 mg/mL cell-free extract powder, 6: ProRCMS#15, 19.1.2020, 5 µL of 10 mg/mL cell-free extract powder. Analyzed by Prozomix Ltd..

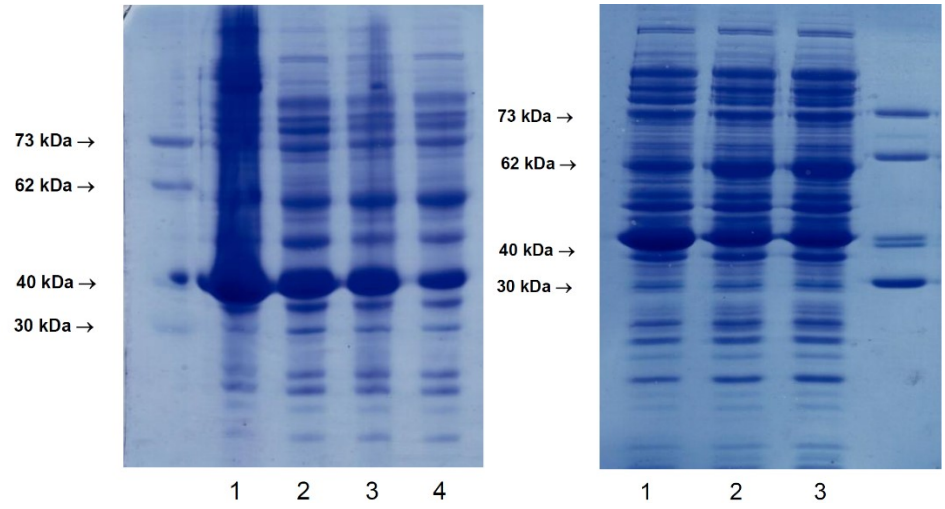


Figure 158: DAAO (*Rhodotorula gracilis*), expressed in E. coli BL21 DE3, Prozomix Ltd.. Left: 2 µL of total cell (1, 3), or cell-free extract (post expression, 2, 4).. Right: DAAO (*Rhodotorula gracilis*), expressed in E. coli BL21 DE3, cell-free extract, Prozomix Ltd.; 5 µL of 10 mg/mL cell-free extract powder (2, 4). Analyzed by Prozomix Ltd..

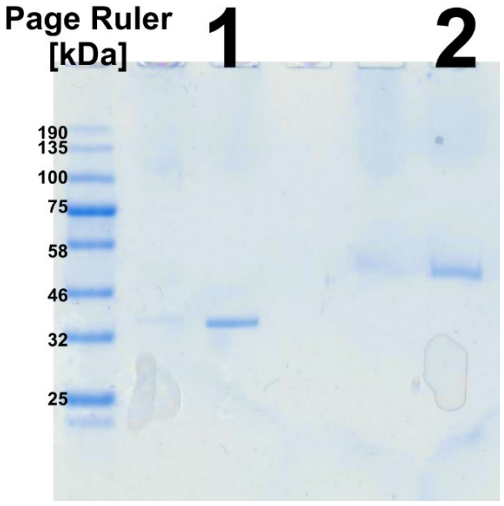


Figure 159: D-Amino acid oxidase (*Rhodotorula gracilis*), expressed in E. coli BL21 DE3, Prozomix Ltd.; 3 µg (BCA assay) of soluble protein per lane.

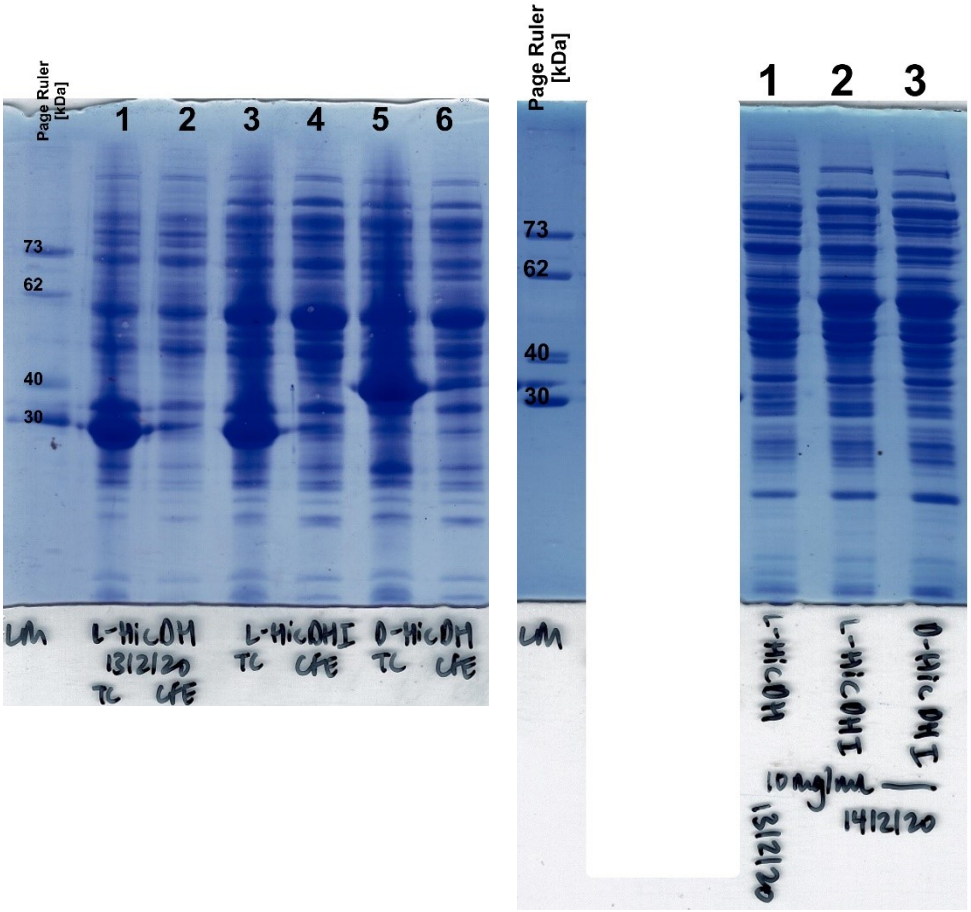


Figure 160: L-2-hydroxyisocaproate dehydrogenase (L-HicDH) from *Lactobacillus confusus*, pdbid: 1HYH, [373] expressed in E. coli BL21 DE3, Prozomix Ltd. and D-2-hydroxyisocaproate dehydrogenase (D-HicDH) from *Lactobacillus paracasei*, pdbid: 1DXY, [83] expressed in E. coli BL21 DE3, Prozomix Ltd. Left: 1: L-HicDH, 13.2.2020, 2 µL of total cell; 2: L-HicDH, 13.2.2020, 2 µL of cell-free extract (post expression); 3: L-HicDH, 14.2.2020, 2 µL of total cell; 4: L-HicDH, 14.2.2020, 2 µL of cell-free extract (post expression); 5: D-HicDH, 14.2.2020, 2 µL of total cell; 6: D-HicDH, 14.2.2020, 2 µL of cell-free extract (post expression). Right: 1: L-HicDH, 13.1.2020, 5 µL of 10 mg/mL cell-free extract powder; 2: L-HicDH, 14.1.2020, 5 µL of 10 mg/mL cell-free extract powder; 3: D-HicDH, 14.1.2020, 5 µL of 10 mg/mL cell-free extract powder. Analyzed by Prozomix Ltd..

7.2 Exemplary HPLC Spectra

7.2.1 Separation and Mass Spectrometric Detection of β-Aspartame

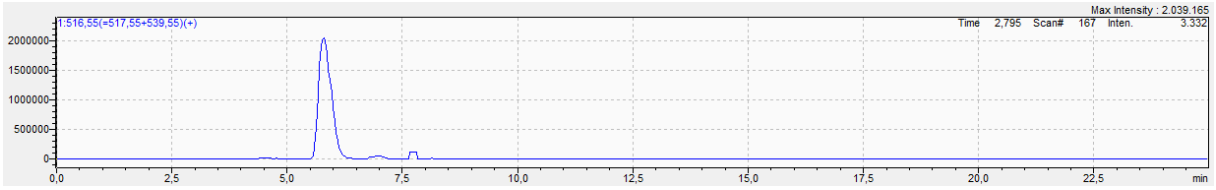


Figure 161: Exemplary HPLC/MS spectrum (Method chiral normal isocratic II with dual ion measurement detection at m/z = 517.55 and 538.55) of a Fmoc-derivatized sample, taken after incubation (18 h, 25 °C) of **3** (methyl L-aspartyl-β-phenylalaninate, 45 mM) in potassium phosphate buffer (50 mM, pH 7.5, 0.1 mM PLP), employing Prozomix racemase ProRCMS#2 (3.6 mg/mL, cell-free extract), showing the single peak of methyl N^{α}-(((*9H*-fluoren-9-yl)methoxy)carbonyl)-L-aspartyl-β-phenylalaninate.

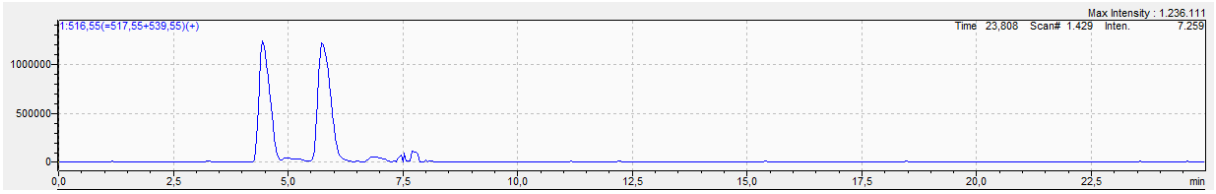


Figure 162: Exemplary HPLC/MS spectrum (Method chiral normal isocratic II with dual ion measurement detection at m/z = 517.55 and 538.55) of a Fmoc-derivatized sample, taken after racemisation (18 h, 25 °C) of **3** (methyl L-aspartyl-β-phenylalaninate, 45 mM) in potassium phosphate buffer (50 mM, pH 7.5, 0.1 mM PLP), employing Prozomix racemase ProRCMS#11 (3.6 mg/mL, cell-free extract), showing the two peaks of methyl methyl N^{α}-(((*9H*-fluoren-9-yl)methoxy)carbonyl)-D-aspartyl-β-phenylalaninate and methyl N^{α}-(((*9H*-fluoren-9-yl)methoxy)carbonyl)-L-aspartyl-β-phenylalaninate.

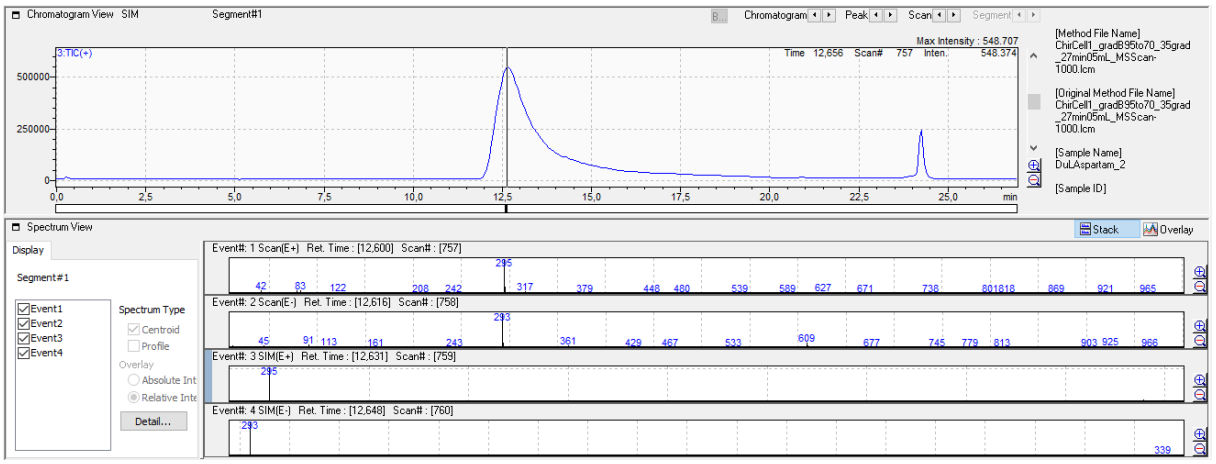


Figure 163: Exemplary HPLC/MS spectrum (Method chiral reverse gradient with MS detection, M^{+}+H = 295, M^{-}H = 293) of a sample, taken after incubation (18 h, 25 °C) of **3** (methyl L-aspartyl-β-phenylalaninate, 45 mM) in potassium phosphate buffer (50 mM, pH 7.5, 0.1 mM PLP), employing Prozomix racemase ProRCMS#2 (3.6 mg/mL, cell-free extract), showing the single peak of **3** (methyl L-aspartyl-β-phenylalaninate

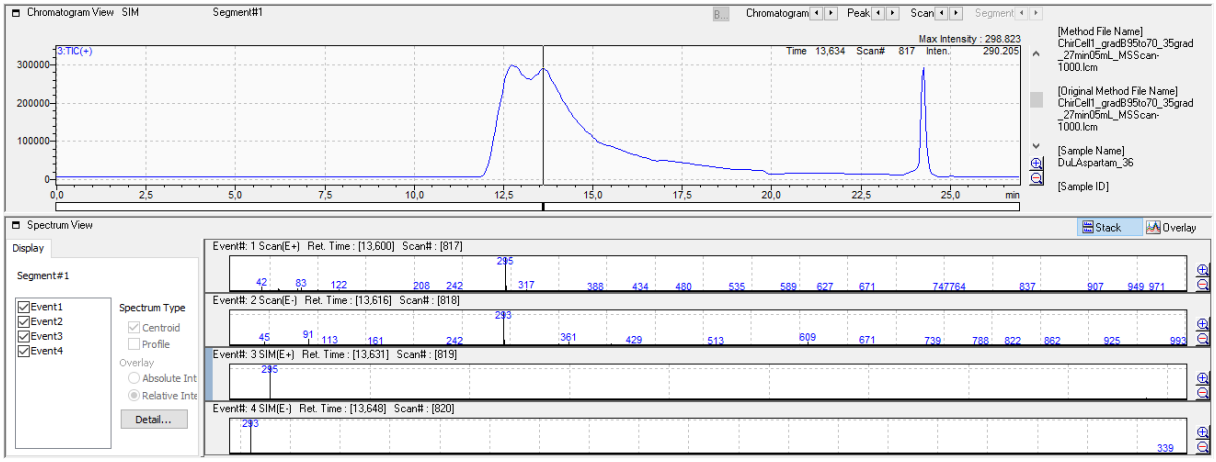


Figure 164: Exemplary HPLC/MS spectrum (Method chiral reverse gradient with MS detection, M^++H = 295, M^-H = 293) of a sample, taken after racemisation (18 h, 25 °C) of **3** (methyl L-aspartyl-β-phenylalaninate, 45 mM) in potassium phosphate buffer (50 mM, pH 7.5, 0.1 mM PLP), employing Prozomix Ltd. racemase ProRCMS#36 (3.6 mg/mL, cell-free extract), showing the two peaks of methyl D-aspartyl-β-phenylalaninate and **3** (methyl L-aspartyl-β-phenylalaninate).

7.3 GC/MS-Spectra and Calibration

7.3.1 Calibration for 2(*tert*-Butyldimethylsilyl)oxy-γ-butyrolactone

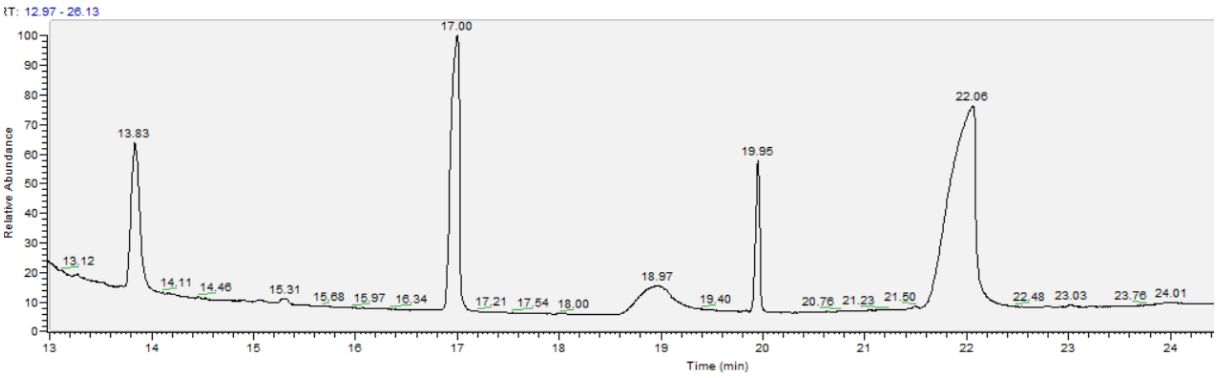


Figure 165:GC/MS (DB5 column) spectrum of a *tert*-butyldimethylsilyl chloride derivatized sample of a reaction mixture, initially containing ***rac*-16** (3-hydroxy-γ-butyrolactone, 97 mM*)*, lipase (2.7 mg/mL, immobilized *Candida antarctica* lipase B, beads, Sprin technologies), molecular sieves (5.1 mg/mL, powder, 4 Å) and vinyl acetate (205 mM, 2.1 eq.) in THF (dry), after reacting at 50 °C for 16 h, 35 min. Before derivatization, n-butanol is added as an internal standard. Excess derivatization agent is quenched with ethanol. 2(*tert*-butyldimethylsilyl)oxy-γ-butyrolactone: 22.06 min, 2-acetoxy-γ-butyrolactone: 19.95 min.

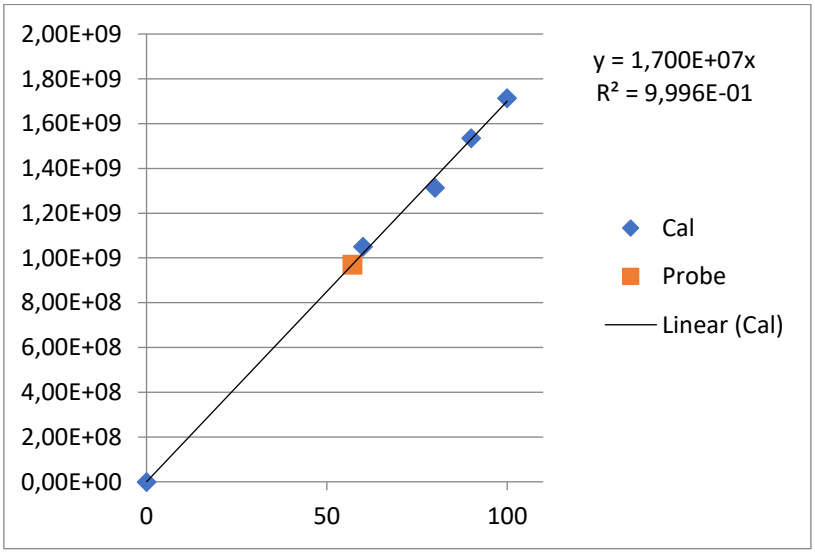


Figure 166: Calibration (Cal) as received from the analytic department. X-axis displaying the concentration [mM], y-axis displaying the peak area of the 2(*tert*-butyldimethylsilyl)oxy-γ-butyrolactone-peak at 22.06 min. The internal standard was not used. The sample (Probe) concentration was calculated to be 57.1 mM in the reaction.

7.3.2 Determination of the 2-Acetoxy-γ-butyrolactone Enantiomeric Excess

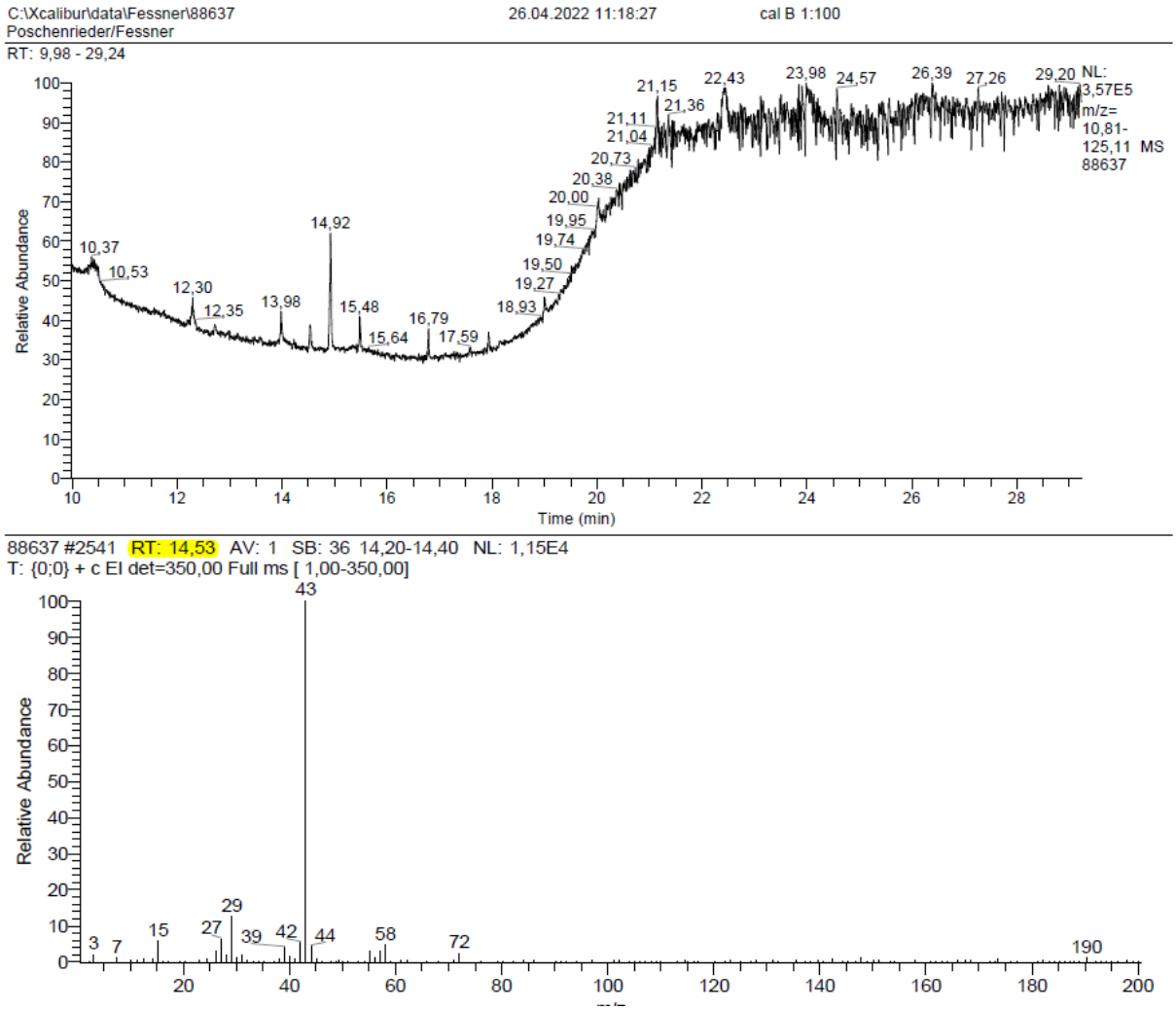


Figure 167: GC/MS spectra of the (*R*)-2-acetoxy-γ-butyrolactone standard, DexCB column, temperature program A.

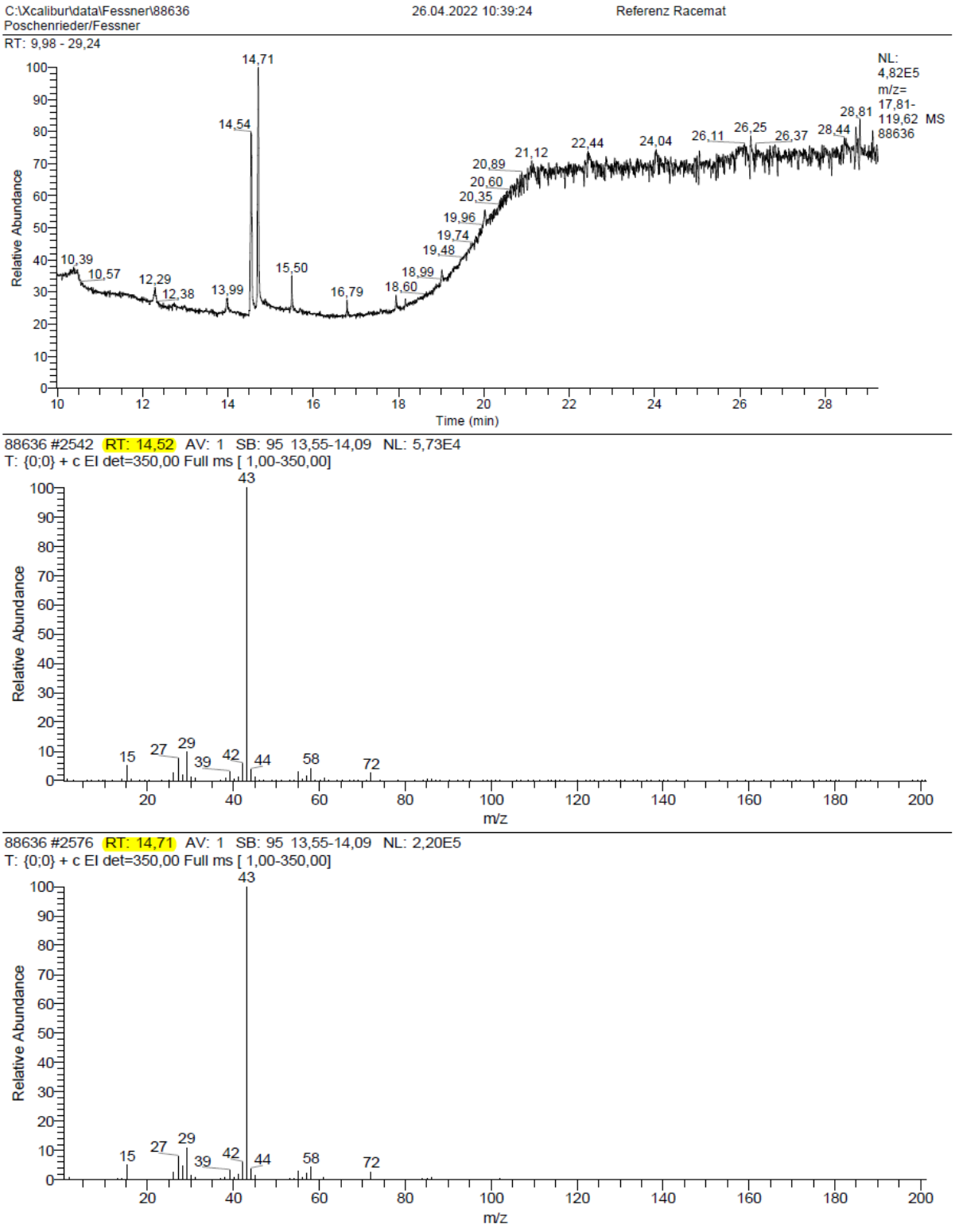


Figure 168: GC/MS spectra of the (*R*/*S*)-2-acetoxy-γ-butyrolactone standard DexCB column, temperature program A.

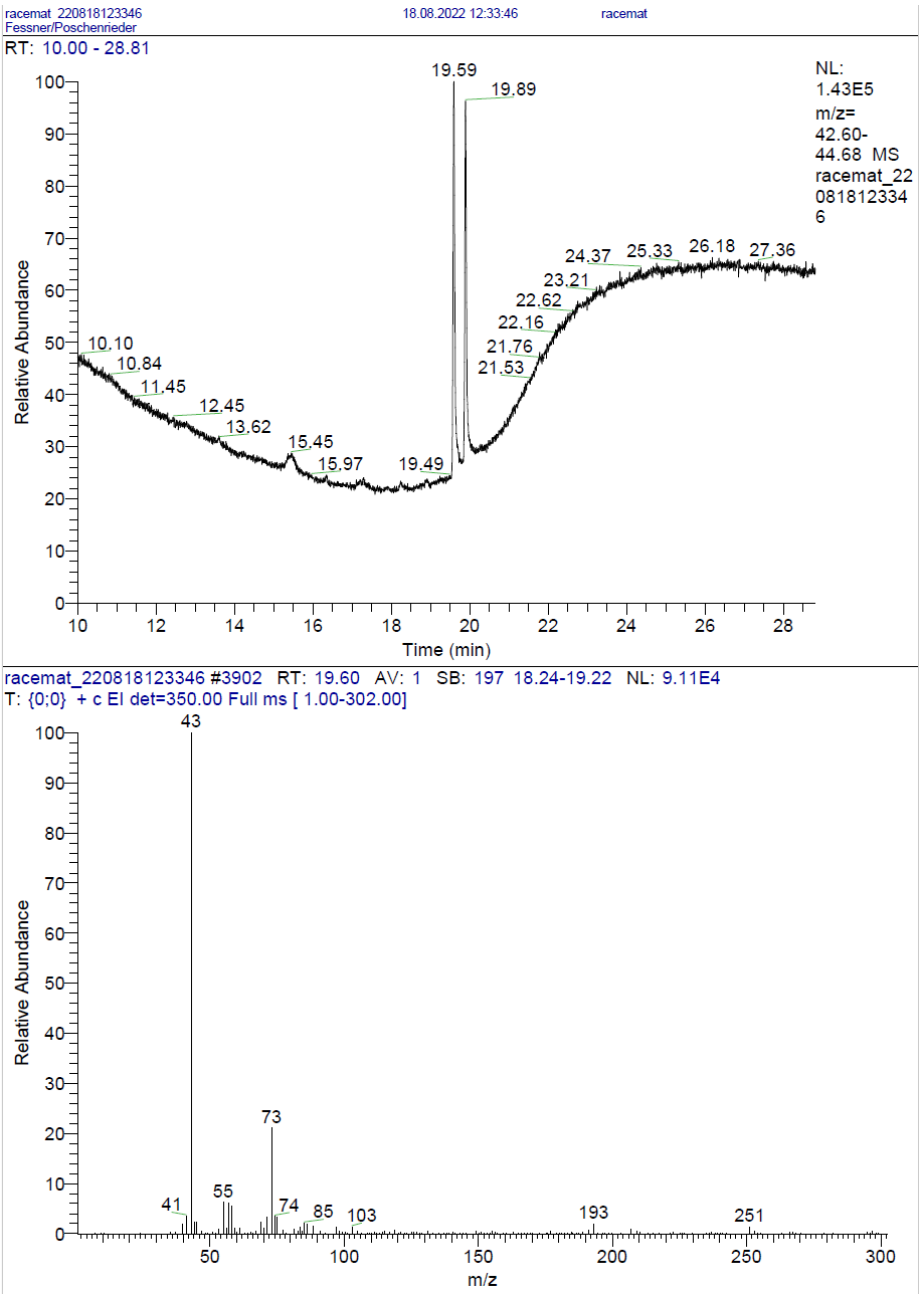


Figure 169: GC/MS spectra of the *rac*-2-acetoxy-γ-butyrolactone standard DexCB column, temperature program B.

PEAK LIST
6_50.RAW
RT: 18.77 - 21.11
Number of detected peaks: 2

Apex RT	Start RT	End RT	Area	%Area	Height	%Height
19.55	19.47	19.73	8518583.058	91.58	2822132.273	91.58
19.84	19.78	19.93	782943.688	8.42	259603.721	8.42

PEAK LIST
7_50.RAW
RT: 19.00 - 21.23
Number of detected peaks: 2

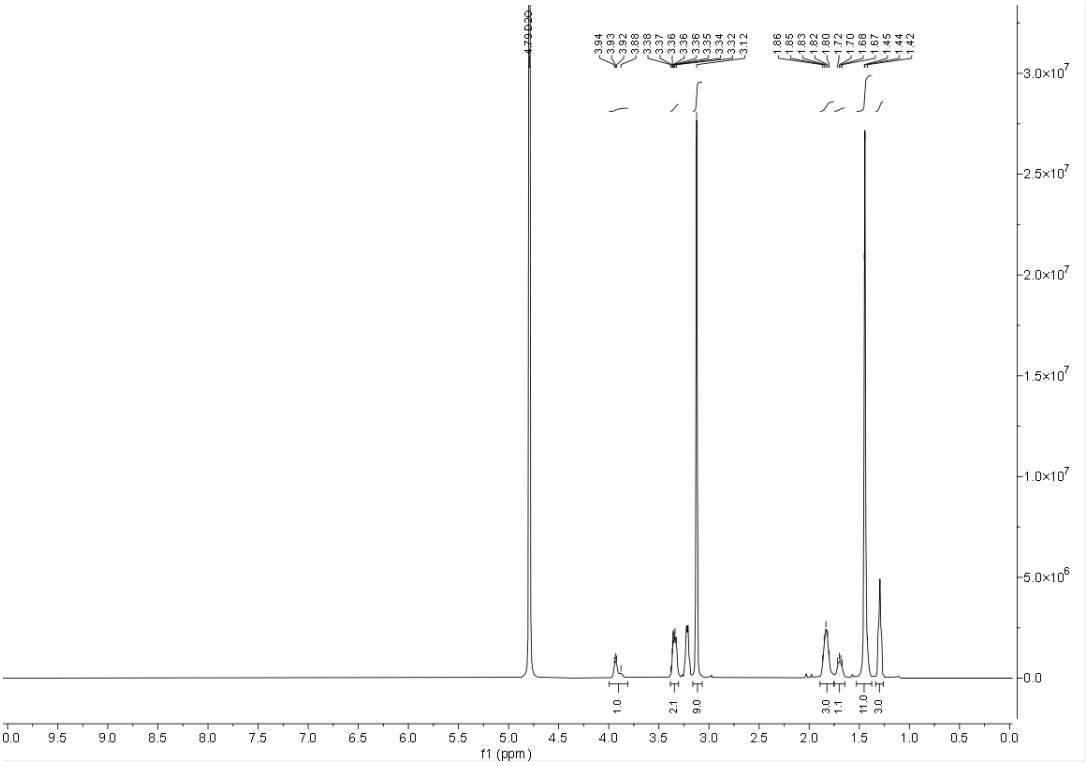
Apex RT	Start RT	End RT	Area	%Area	Height	%Height
19.62	19.56	19.83	2439330.651	92.43	884266.693	92.94
19.92	19.85	20.00	199917.753	7.57	67146.097	7.06

Figure 170: GC/MS (DexCB column) results as received from the analytics department. (*R*)-2-acetoxy-γ-butyrolactone: 19.6 min, (*S*)-2-acetoxy-γ-butyrolactone: 19.8 or, 19.9 min. Samples of the reaction mixture, initially containing ***rac*-16** (3-hydroxy-γ-butyrolactone, 97 mM*)*, lipase (2.7 mg/mL, immobilized *Candida antarctica* lipase B, beads, Sprin technologies), molecular sieves (5.1 mg/mL, powder, 4 Å) and vinyl acetate (205 mM, 2.1 eq.) in THF (dry), after reacting at 50 °C for 16 h, 35 min.

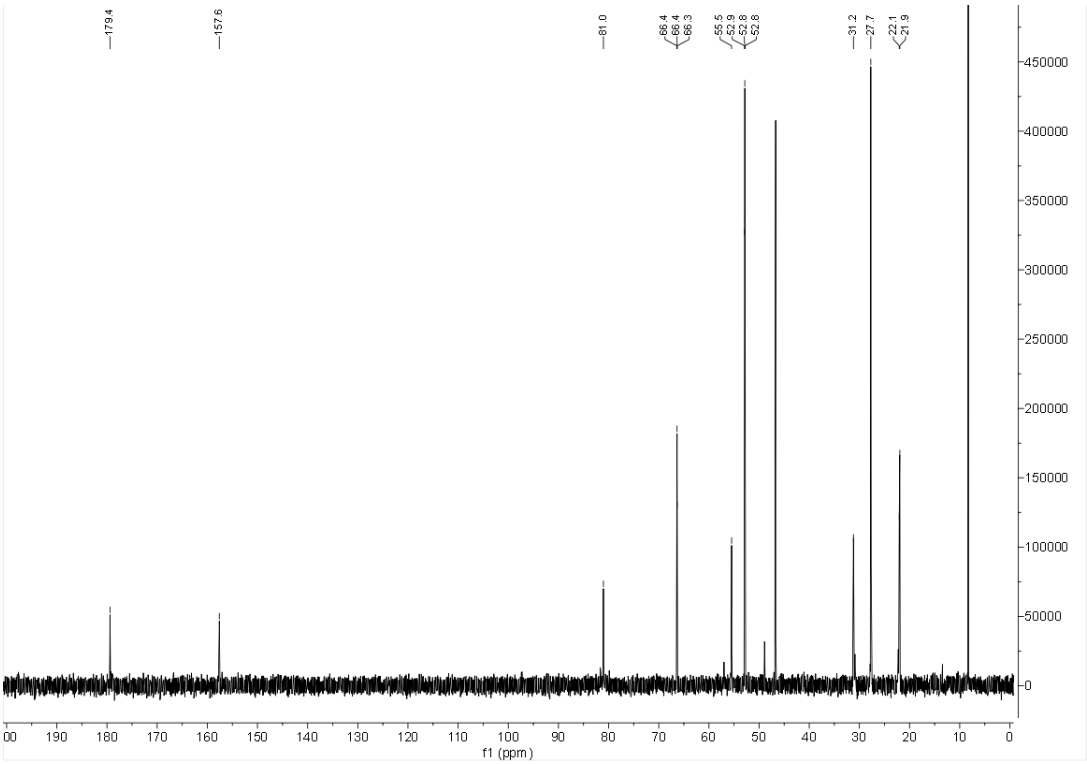
7.4 NMR-Spectra

N_α-(*tert*-Butyloxycarbonyl)N_ε,N_ε,N_ε-trimethyl-L-lysine iodide (**5**)

^{1}H NMR (500 MHz, D_2O)

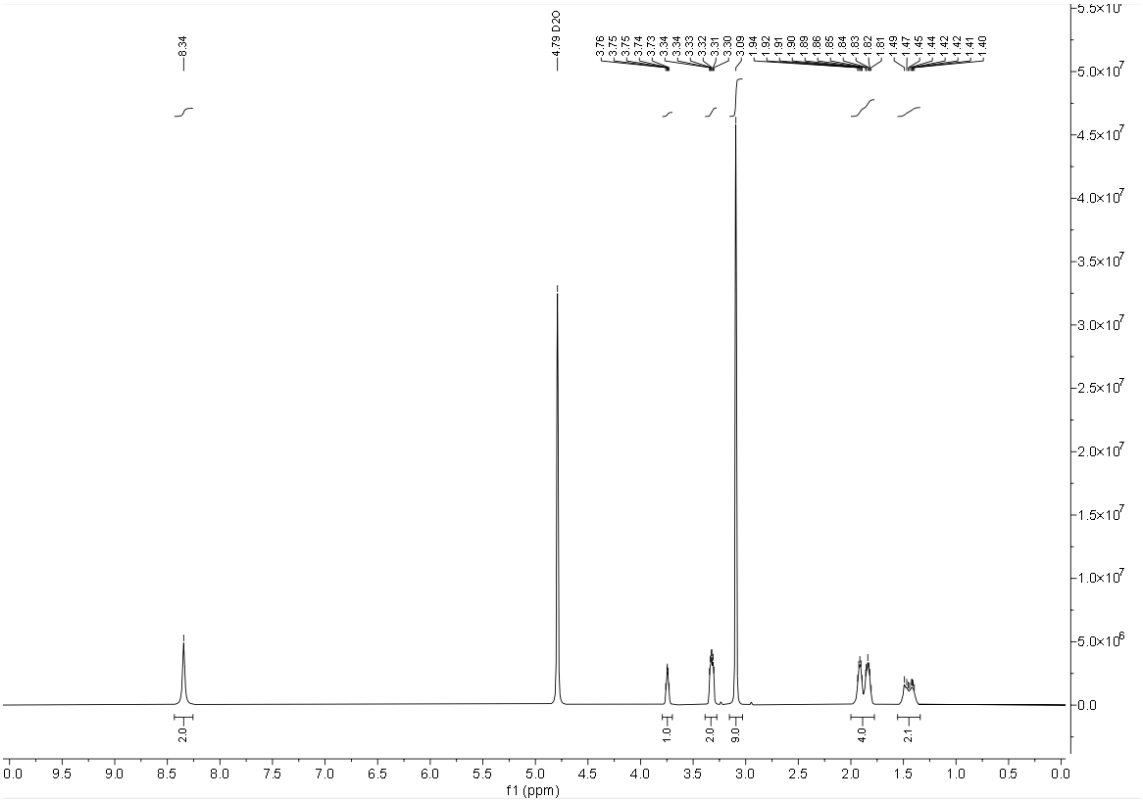


^{13}C NMR (126 MHz, D_2O)

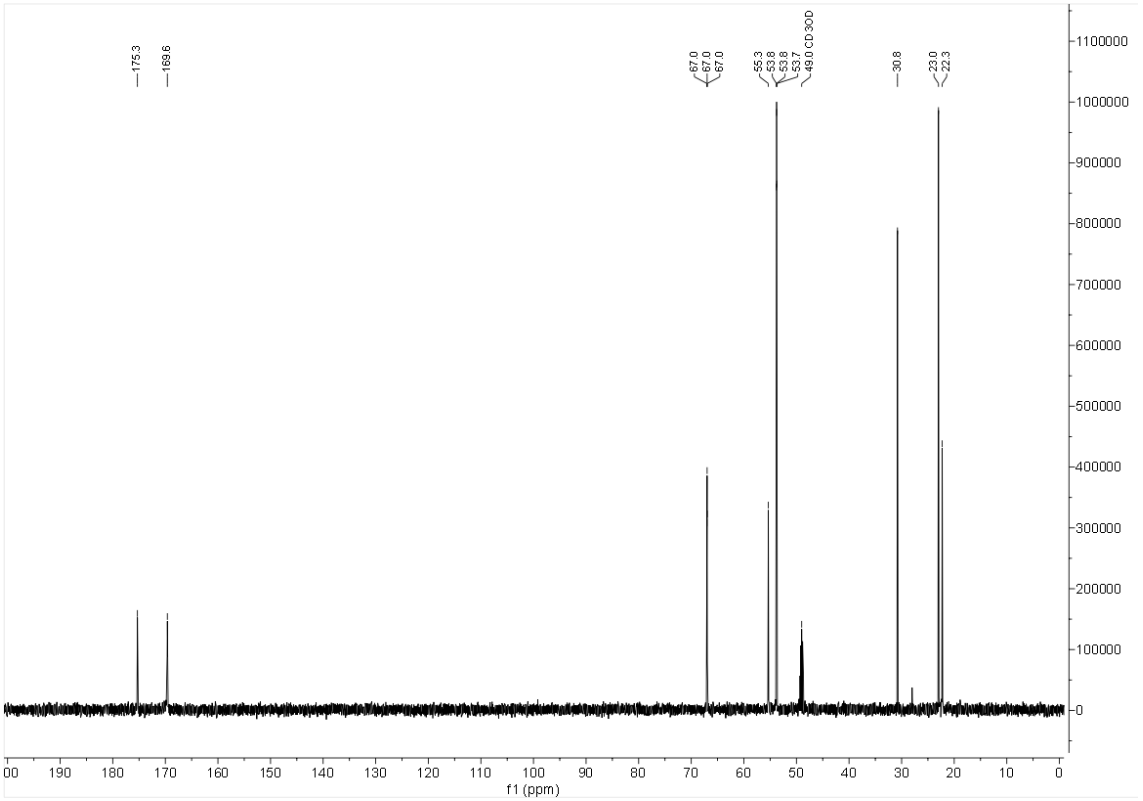


N_ε,N_ε,N_ε-Trimethyl-L-lysine dihydroformate (**1**)

^{1}H NMR: (500 MHz, D_2O/methanol-d^5)

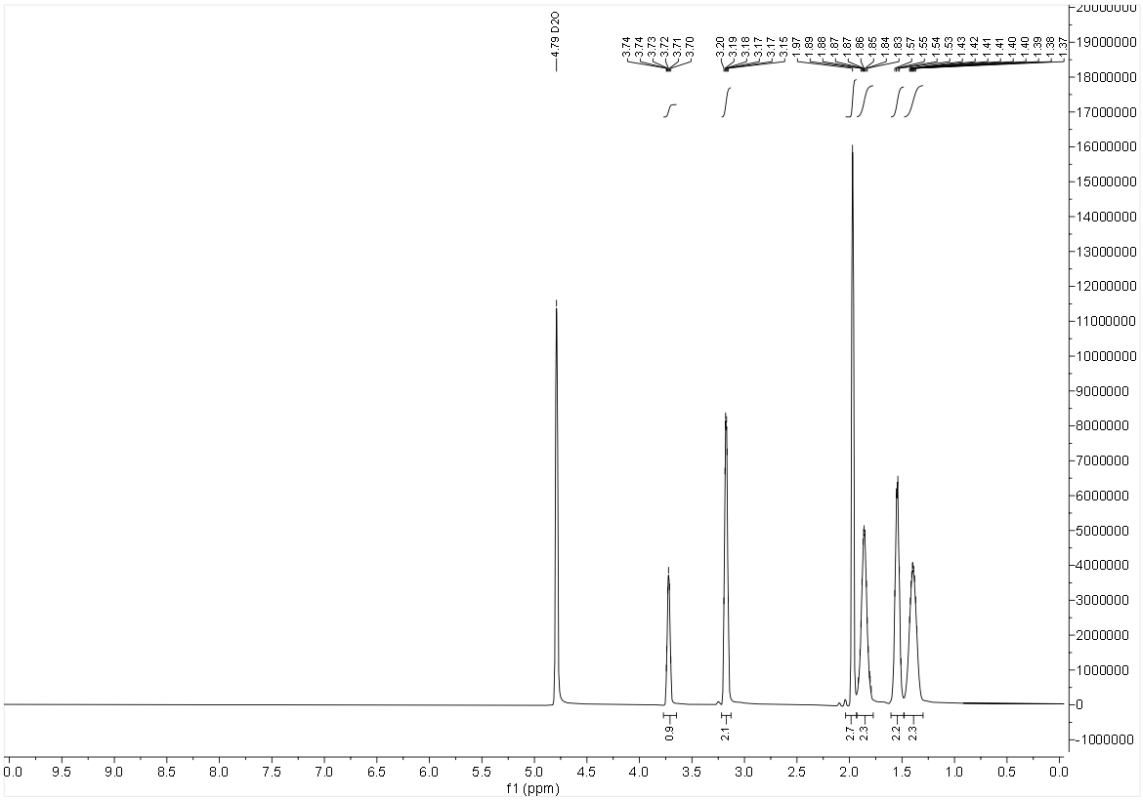


^{13}C NMR: (126 MHz, D_2O/methanol-d^5)

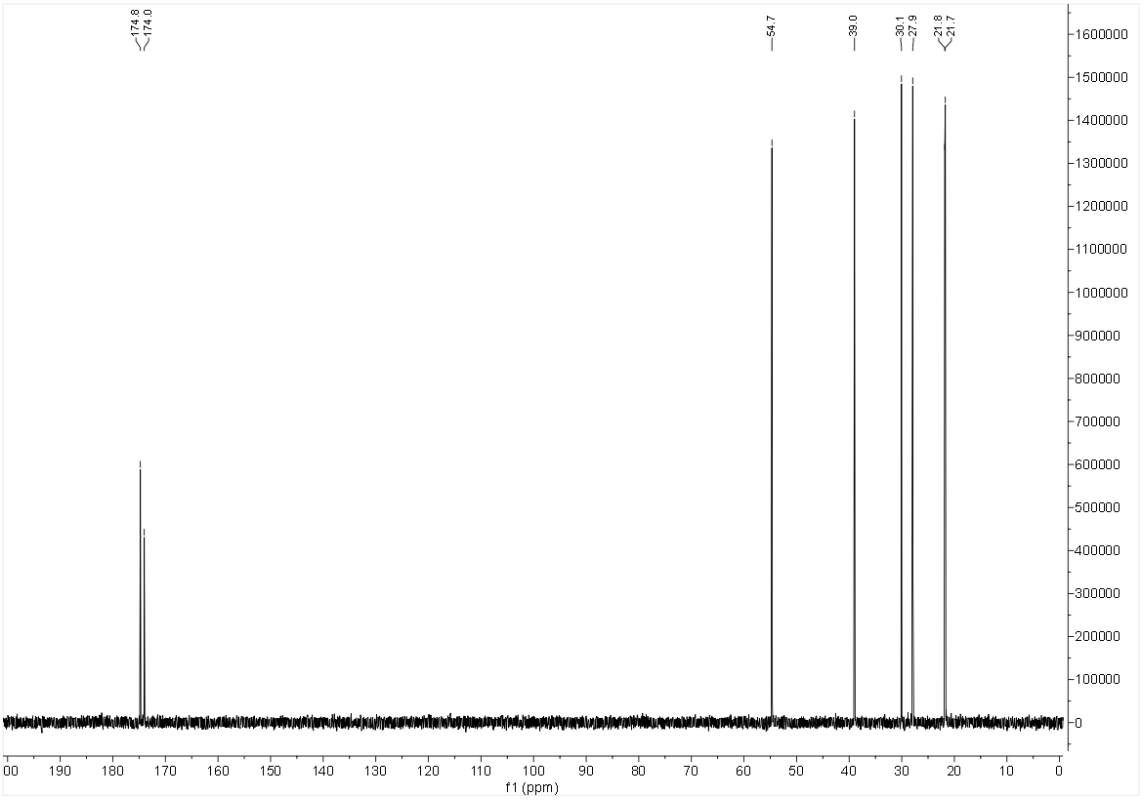


N$_\varepsilon$-Acetyl-D-lysine (**2**)

^{1}H NMR: (500 MHz, D_2O)

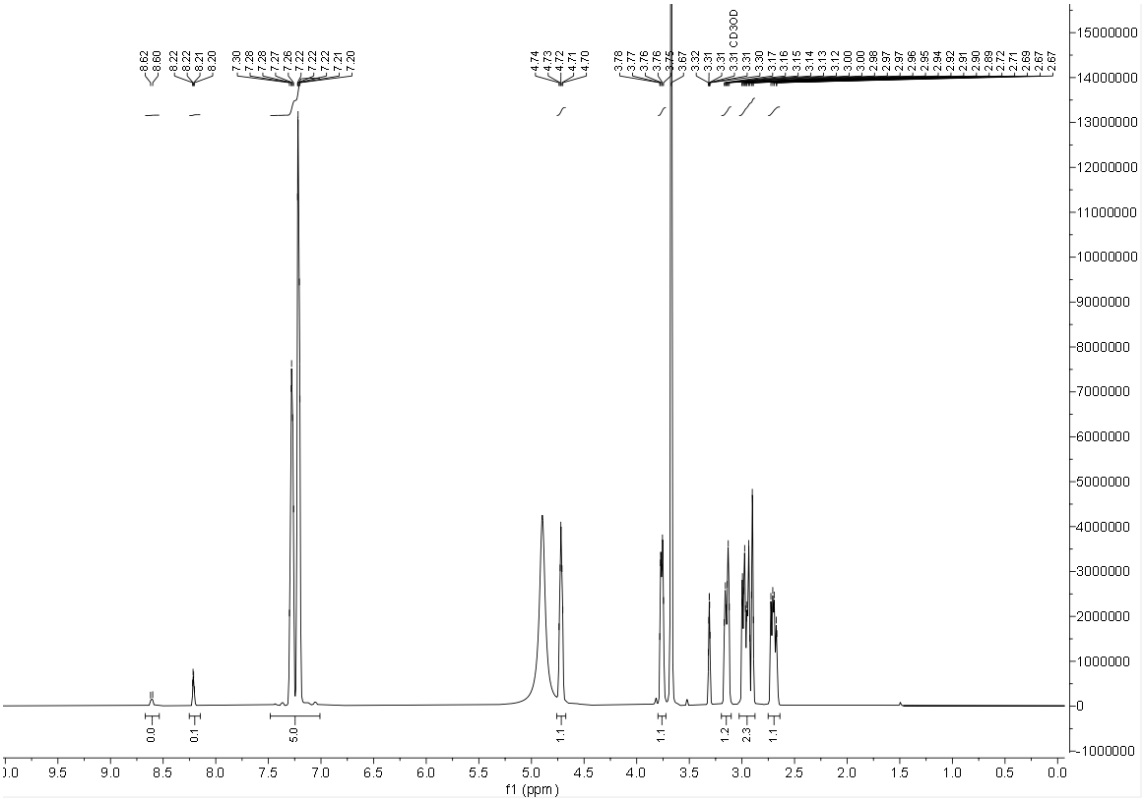


^{13}C NMR: (126 MHz, $CDCl_3$)

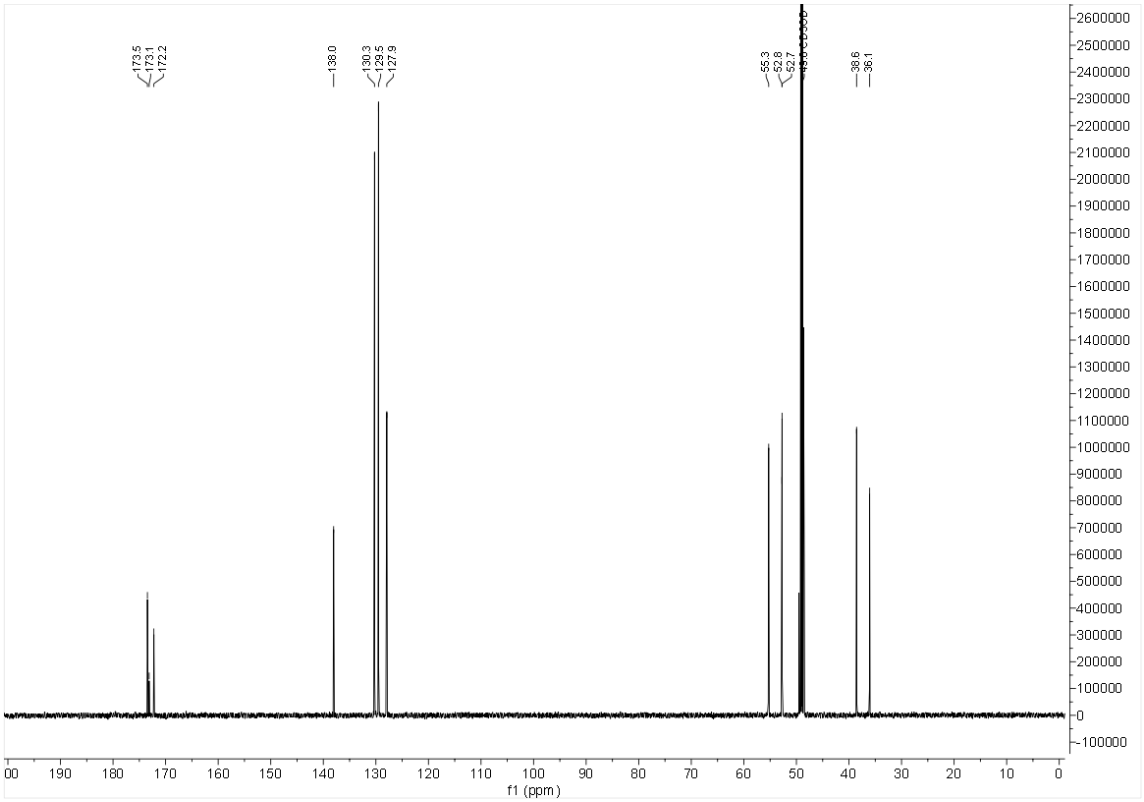


Methyl D-α-aspartyl-L-phenylalaninate (**3**)

^{1}H NMR: (500 MHz, methanol-d^{4})

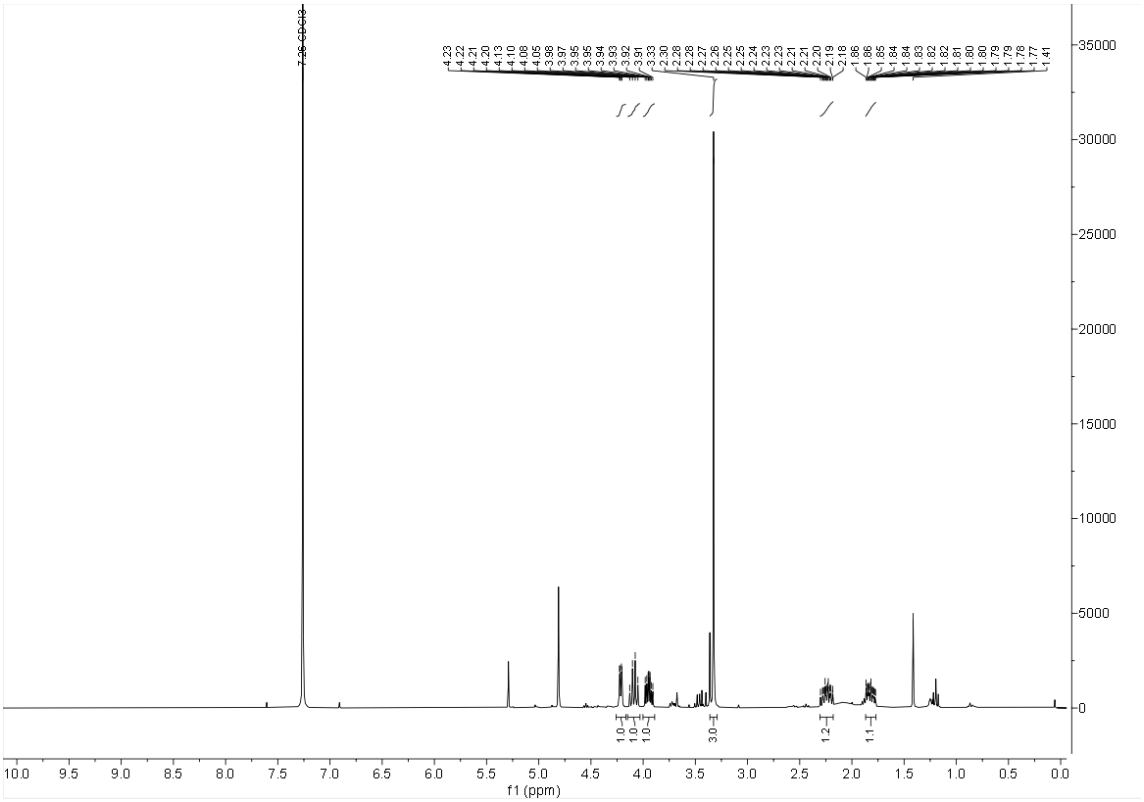


^{13}C NMR: (127 MHz, methanol-d^{4})

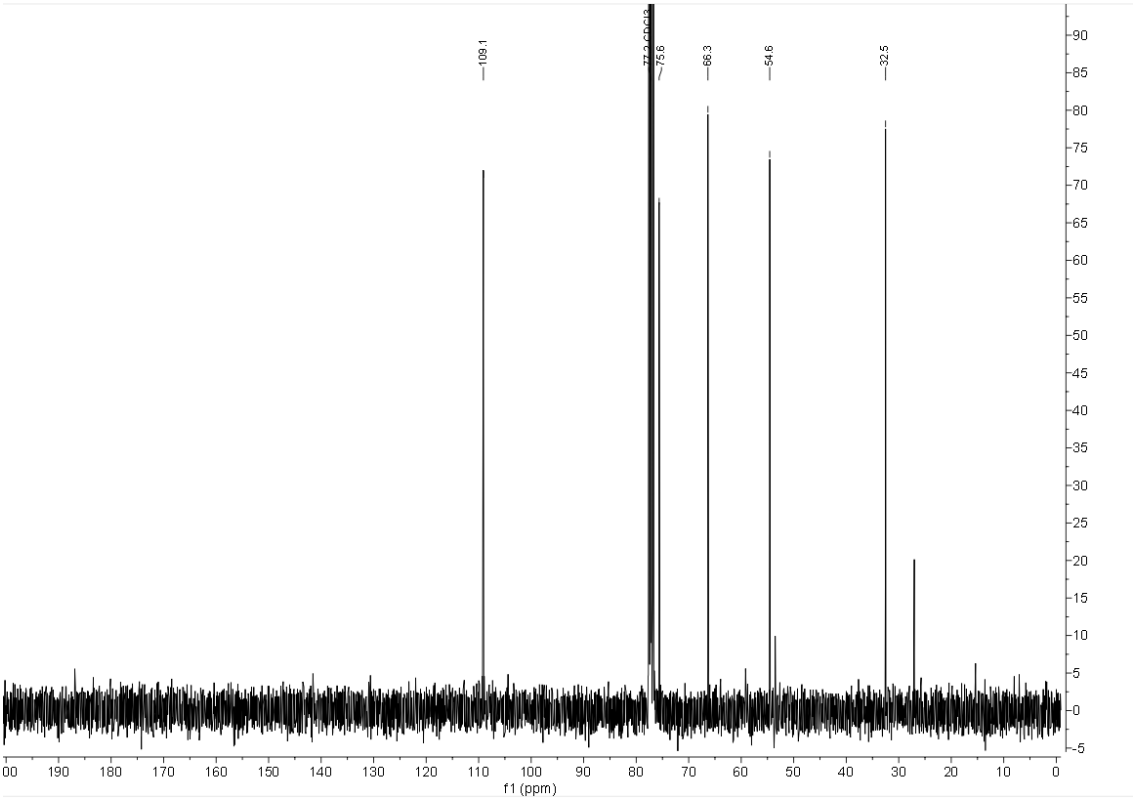


syn-3-Hydroxy-2-methoxytetrahydrofuran(***rac*-12**)

^{1}H NMR: (300 MHz, $CDCl_3$)

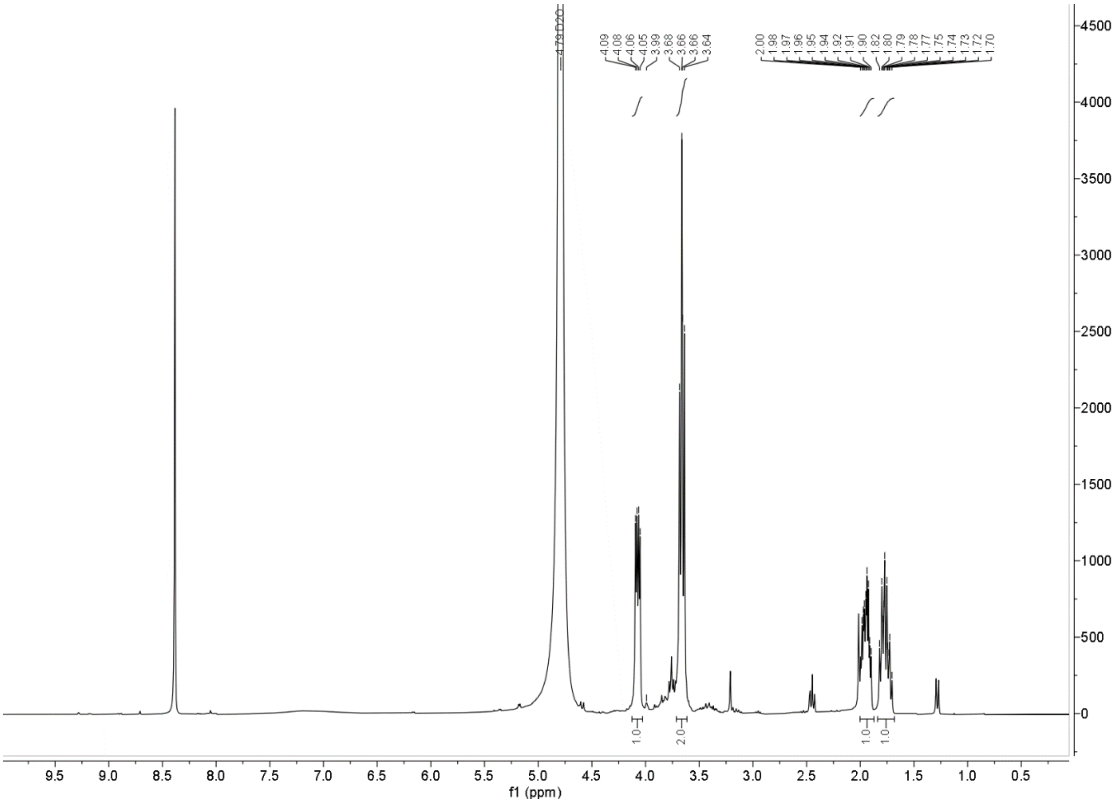


^{13}C NMR: (75 MHz, $CDCl_3$)

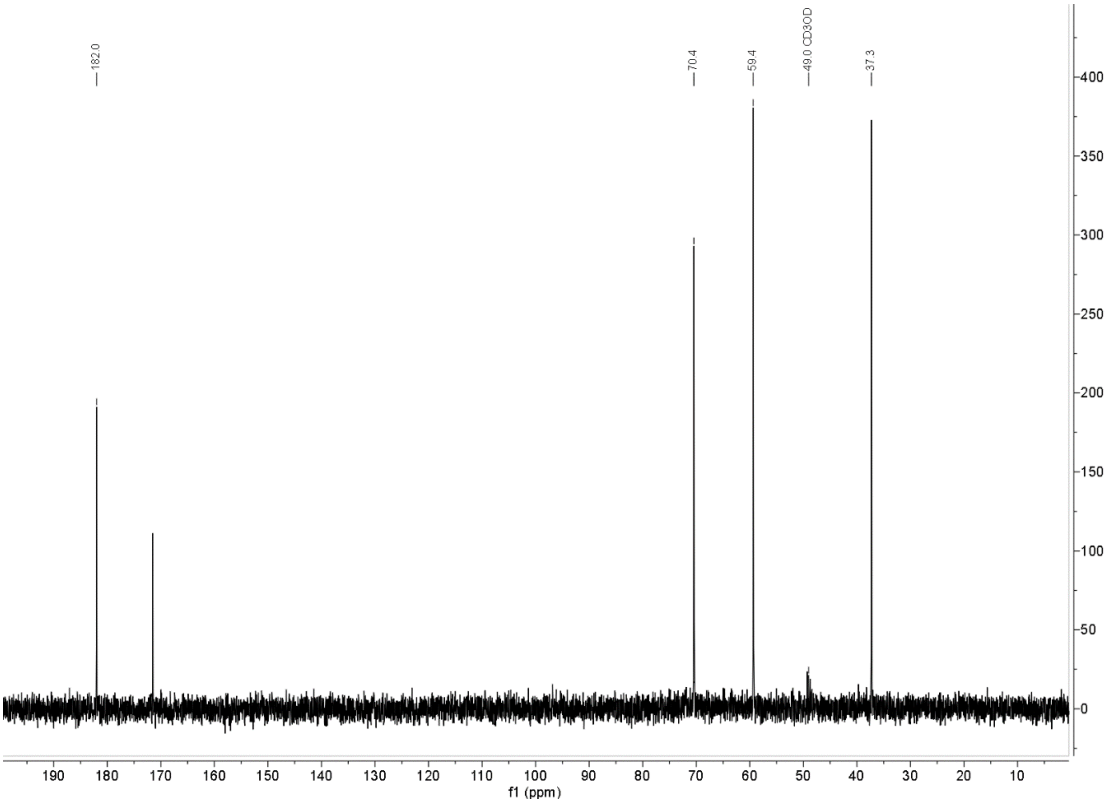


D-2,4-Dihydroxybutyric acid (**15**)

^{1}H NMR: (300 MHz, D_2O/methanol-d^4)

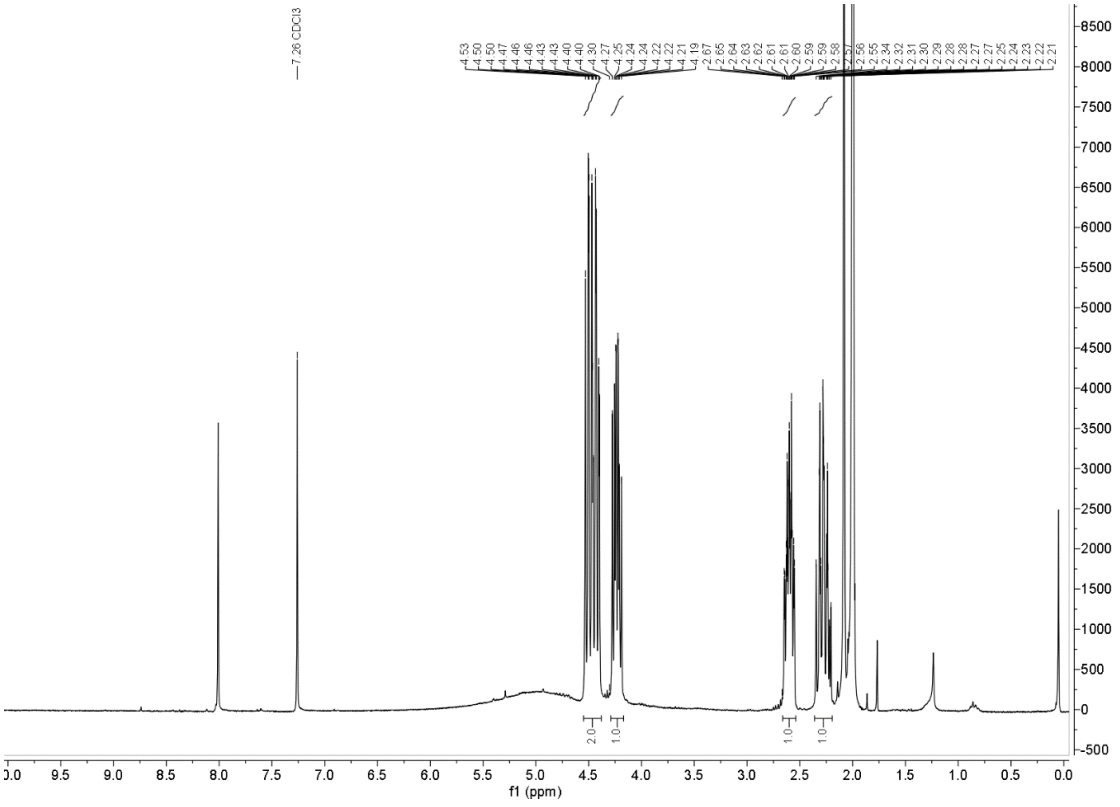


^{13}C NMR: (75 MHz, D_2O/methanol-d^4)

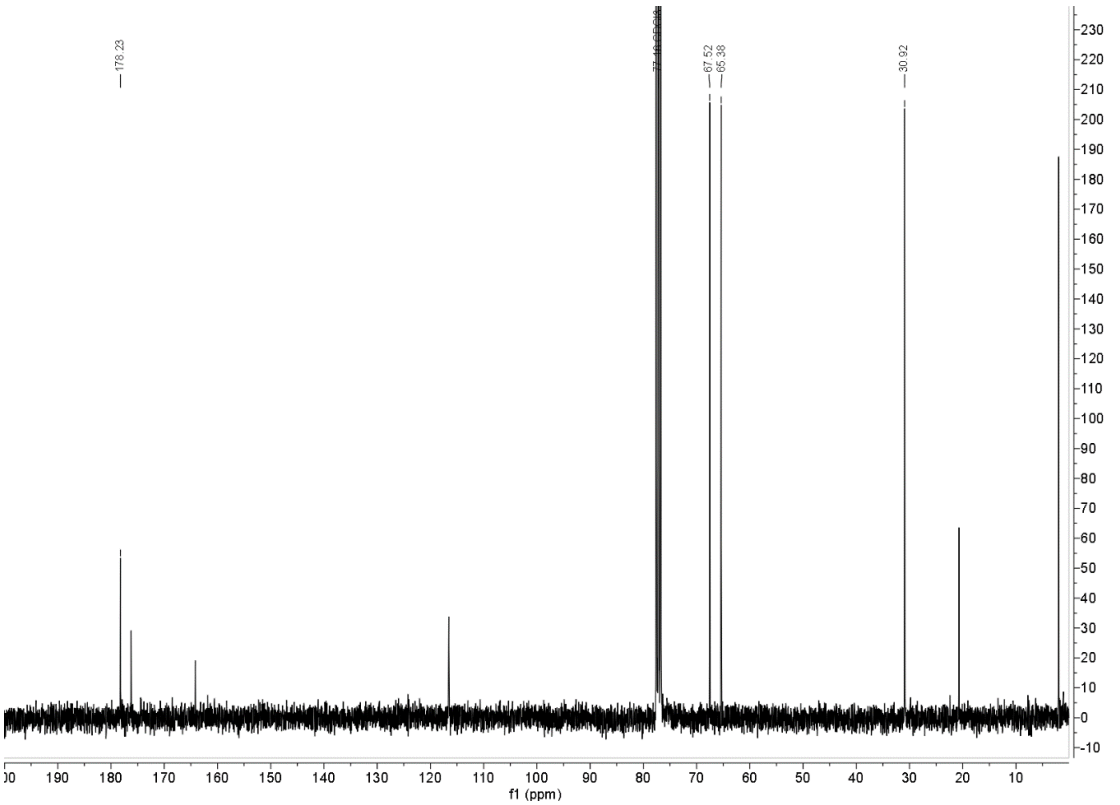


2(*R*)-Hydroxy-γ-butyrolactone (***R*-16**, I)

^{1}H NMR: (300 MHz, $CDCl_3$)

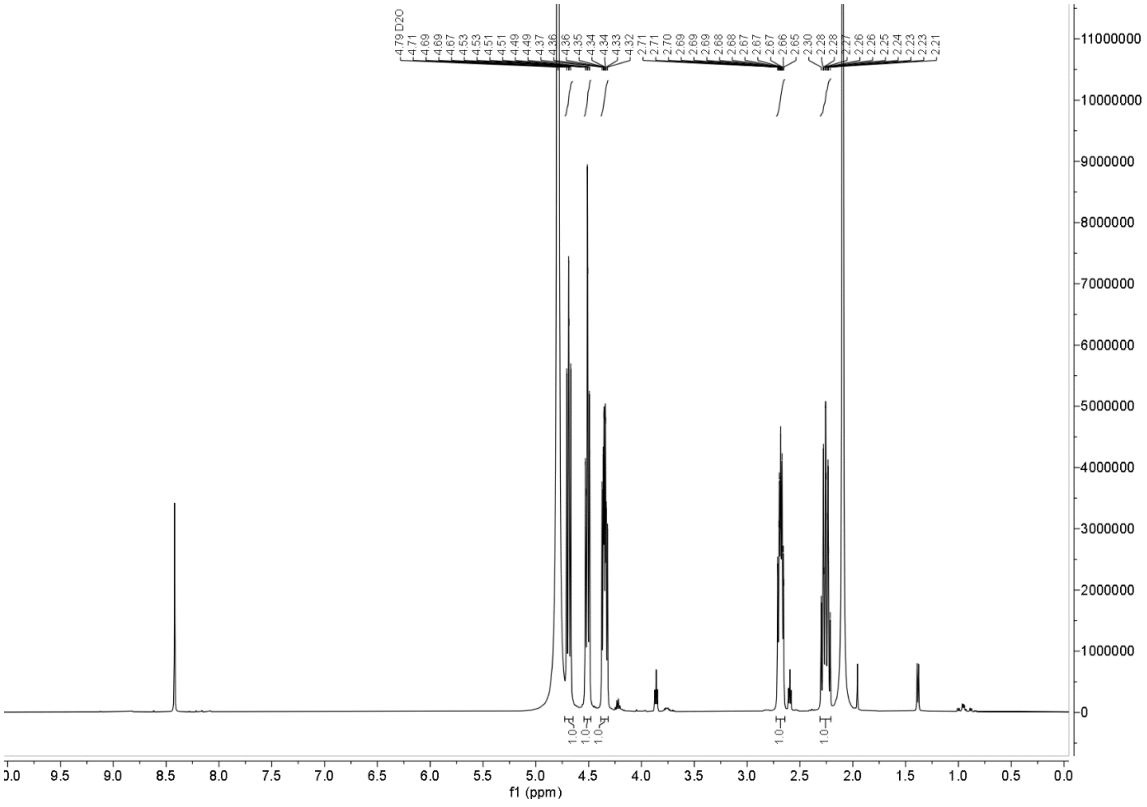


^{13}C NMR: (75 MHz, $CDCl_3$)

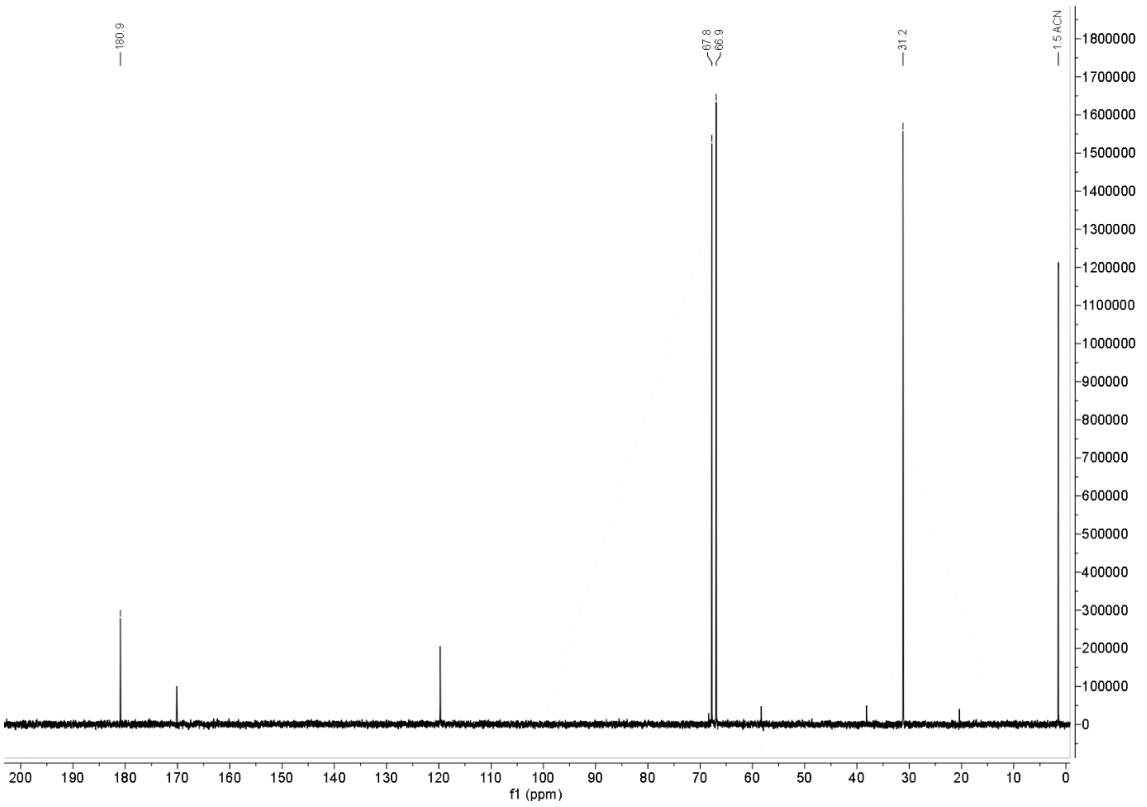


2(*R*)-Hydroxy-γ-butyrolactone (***R*-16**, II)

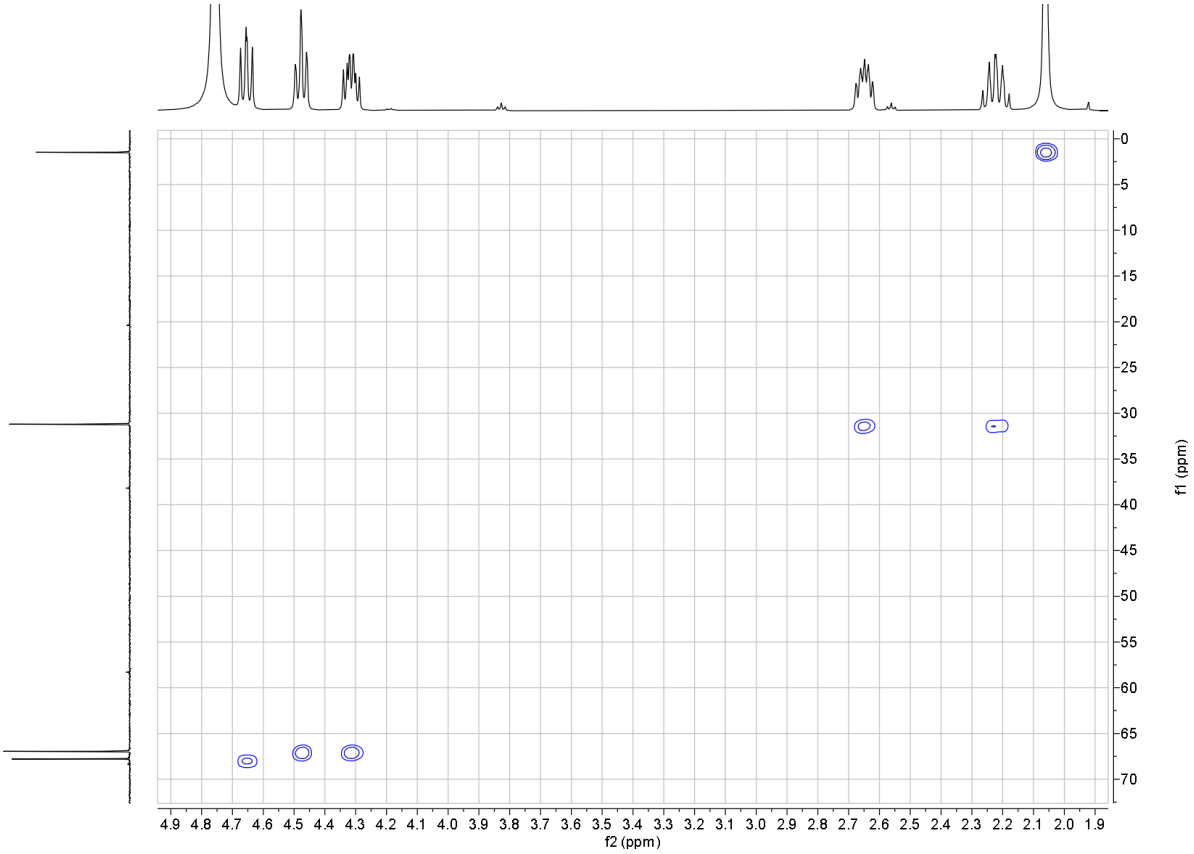
^{1}H NMR: (500 MHz, D_2O)



^{13}C NMR: (126 MHz, D_2O, acetonitrile-std: 1.47 ppm)

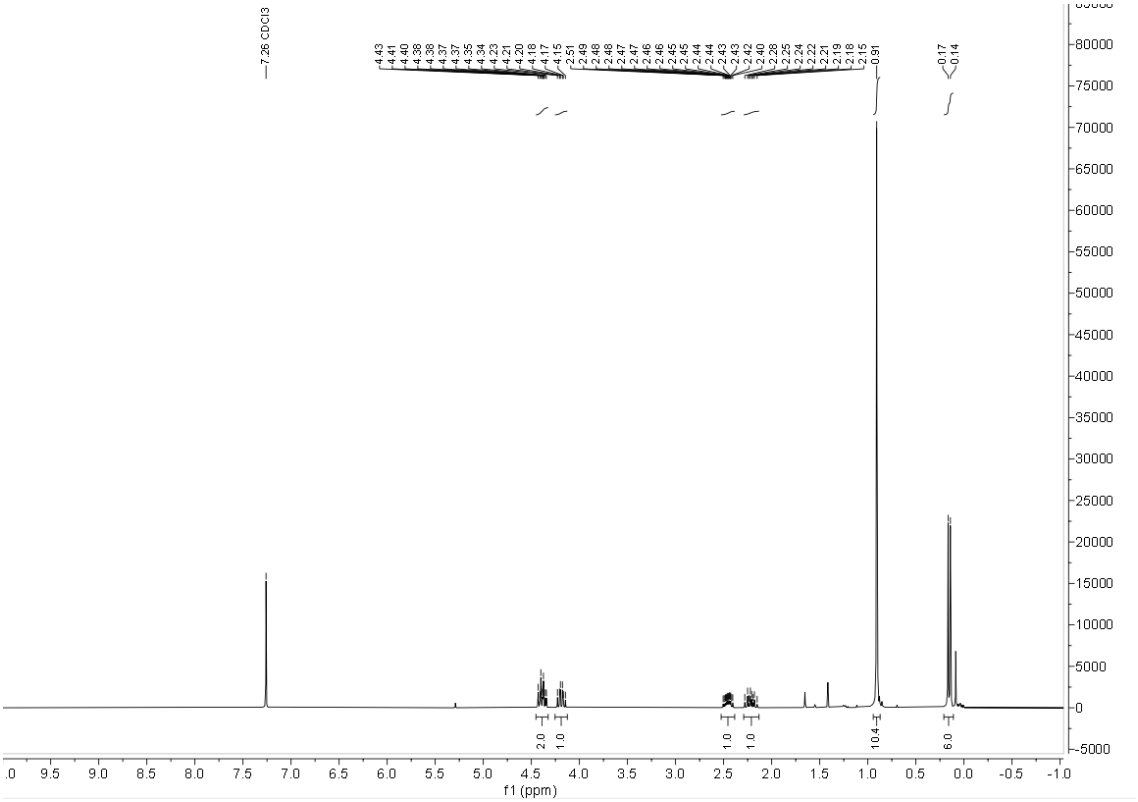


HSQC: (D_2O, acetonitrile-std: 2.06 (^1H-trace), 1.47 ppm (^{13}C-trace).

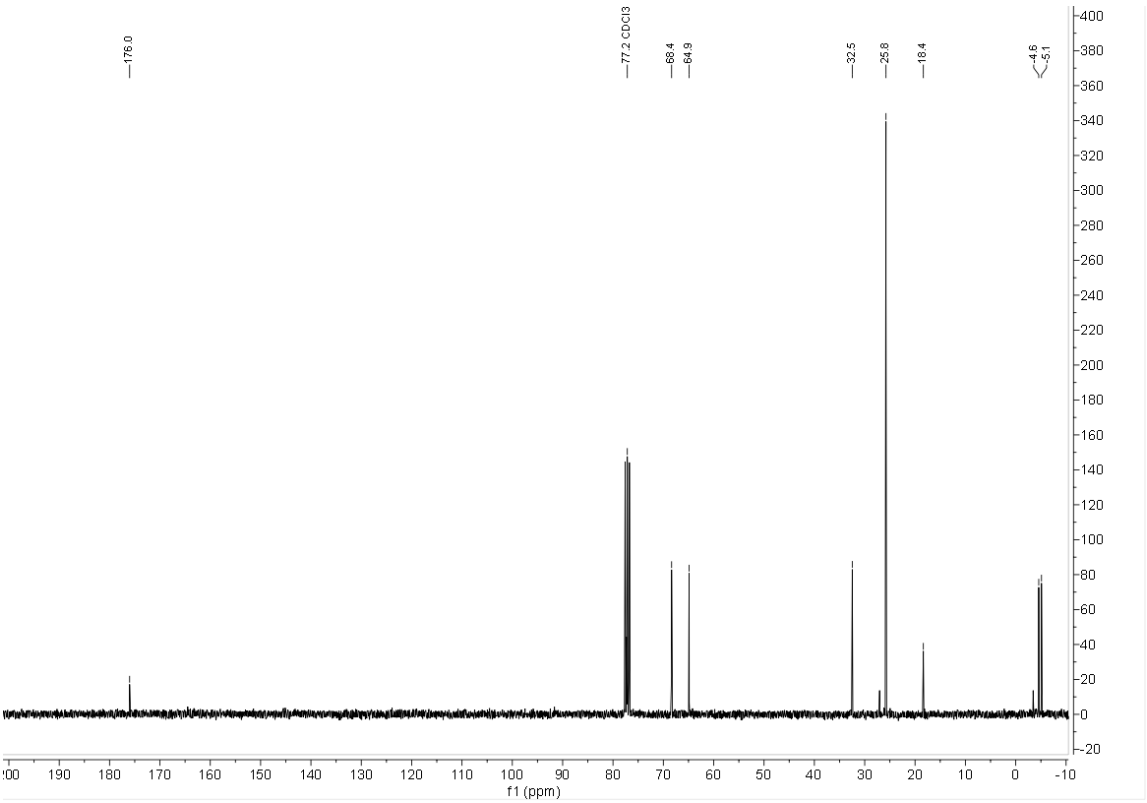


rac-2-(*tert*-Butyldimethylsilyl)-γ-butyrolactone (***rac*-17TBS**)

^{1}H NMR: (300 MHz, $CDCl_3$)

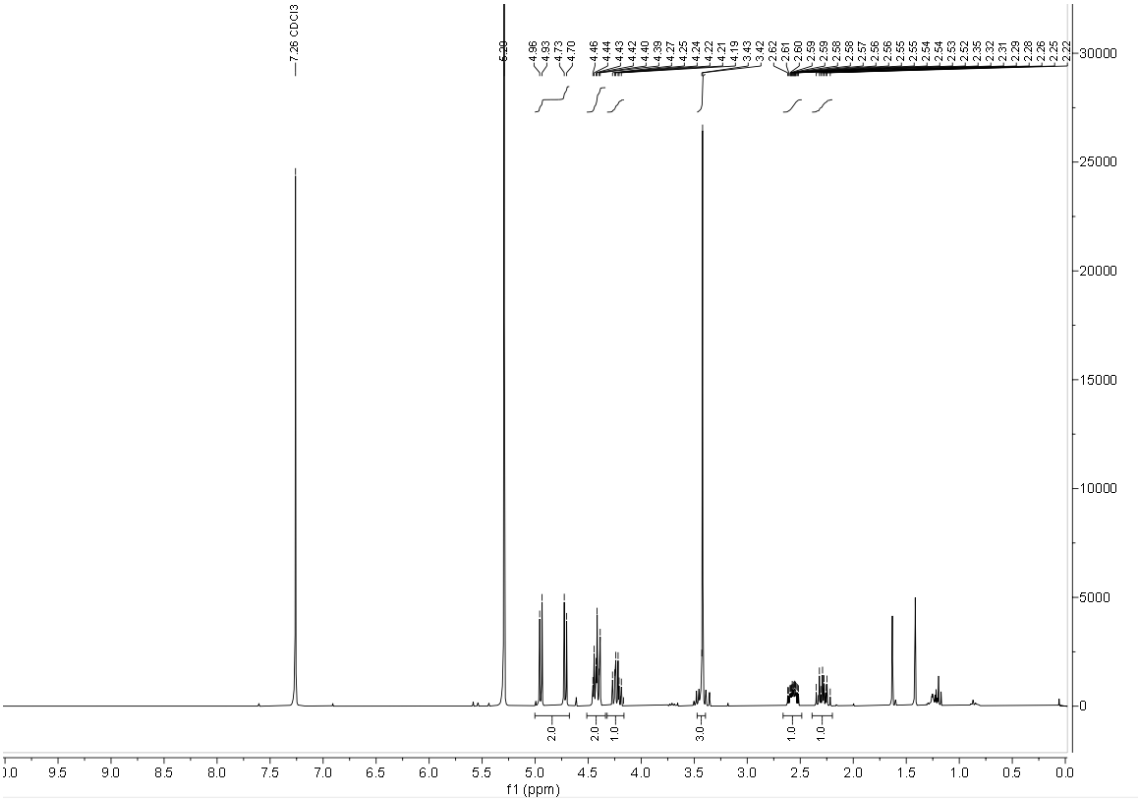


^{13}C NMR: (75 MHz, $CDCl_3$)

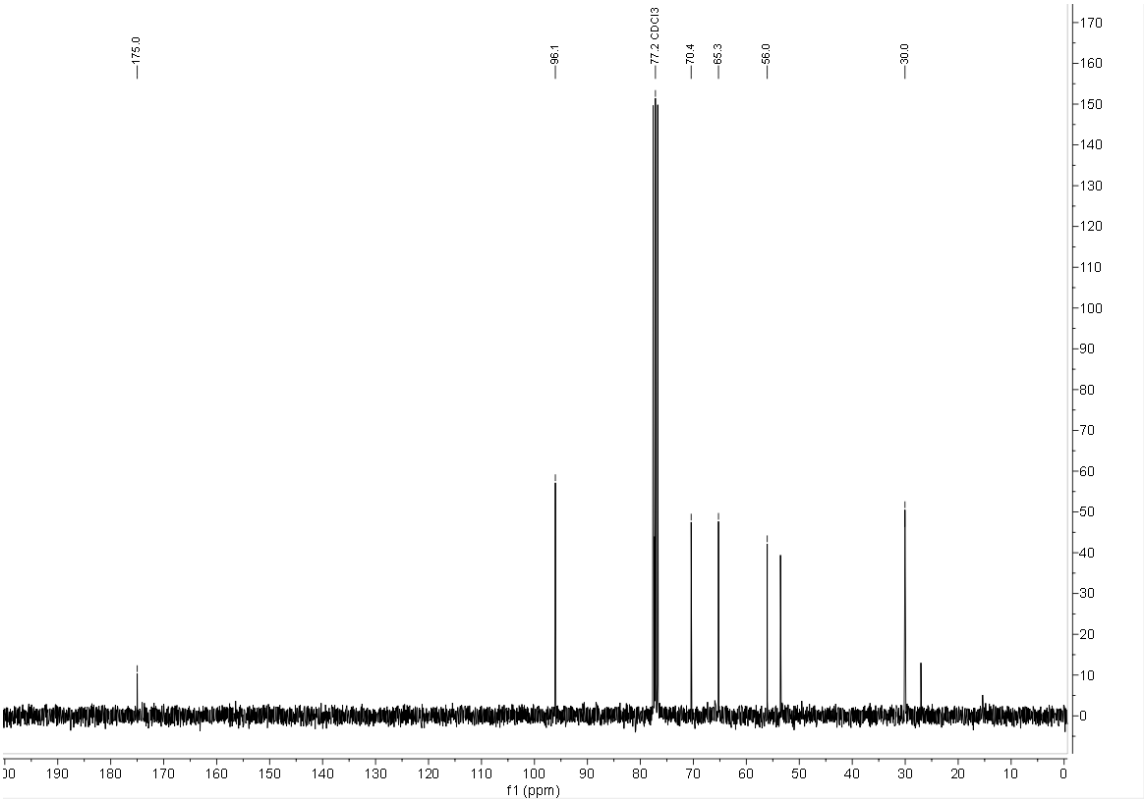


rac-2-(Methoxymethoxy)-γ-butyrolactone (***rac*-17MOM**)

^{1}H NMR: (300 MHz, $CDCl_3$)

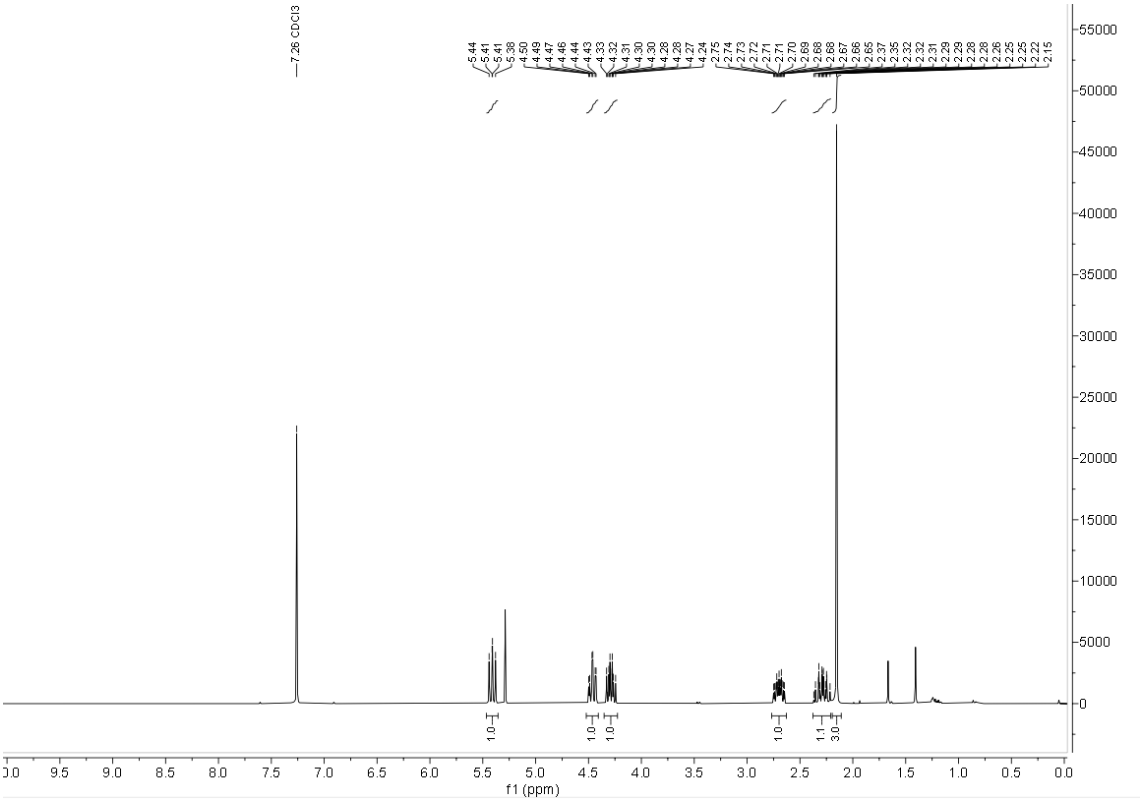


^{13}C NMR: (75 MHz, $CDCl_3$)

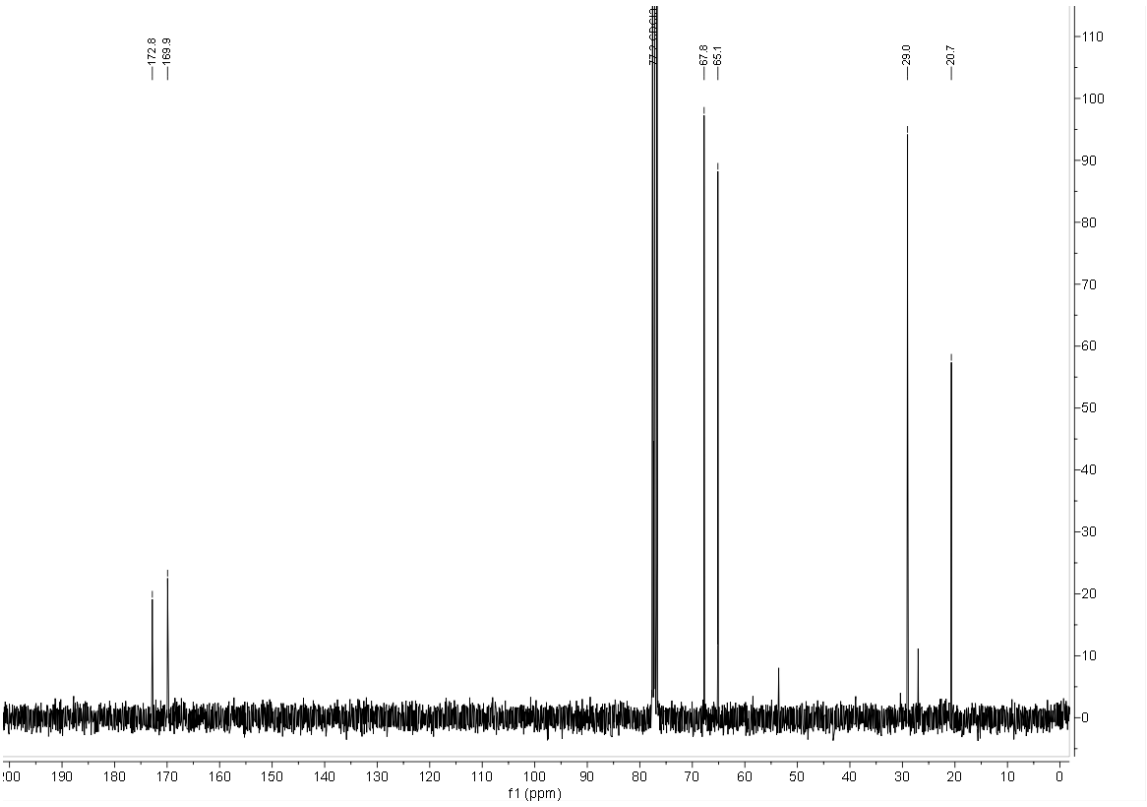


rac-2-Acetoxy-γ-butyrolactone (***rac*-17Ac**)

^{1}H NMR: (300 MHz, $CDCl_3$)

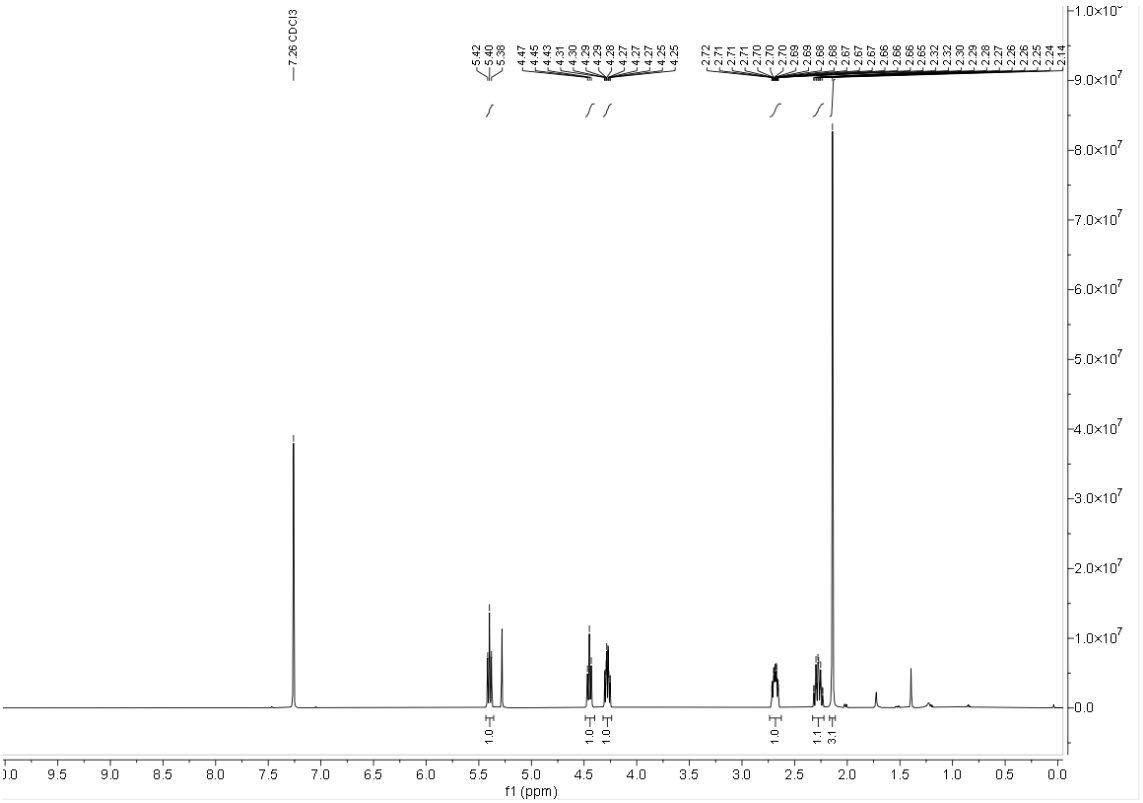


^{13}C NMR: (75 MHz, $CDCl_3$)

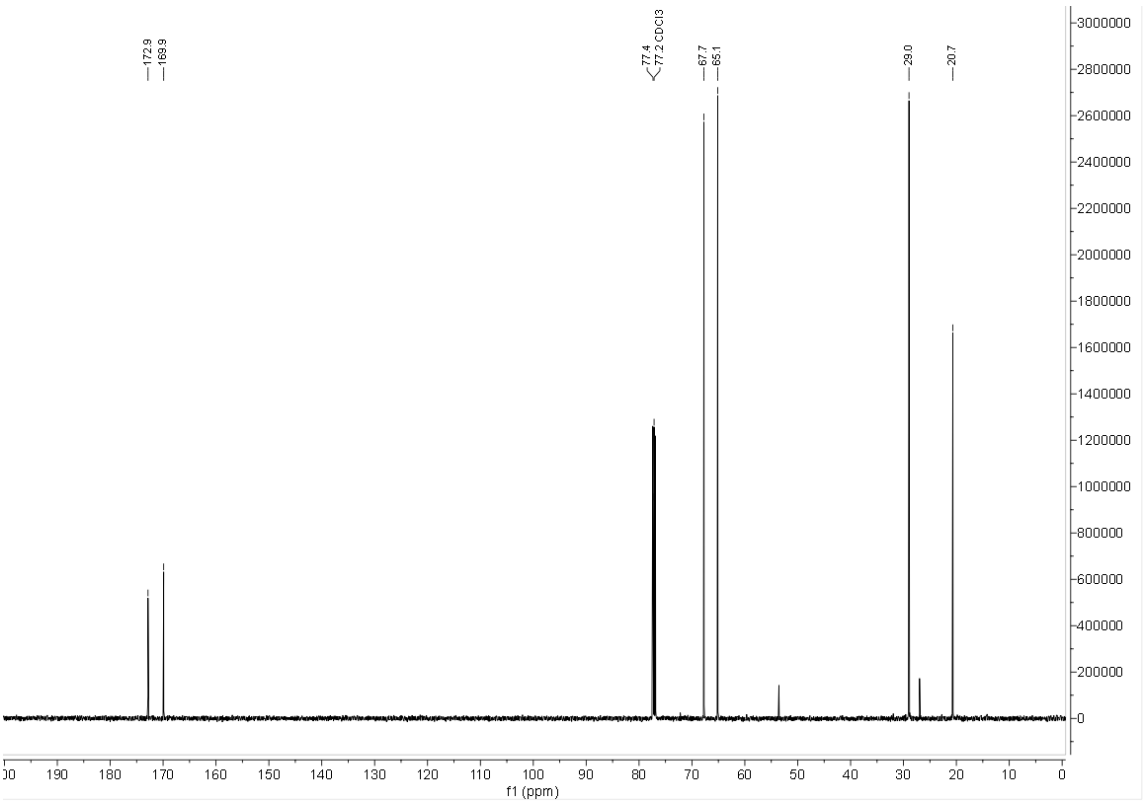


(*R*)-2-Acetoxy-γ-butyrolactone (***R*-17Ac**)

^{1}H NMR: (500 MHz, $CDCl_3$)

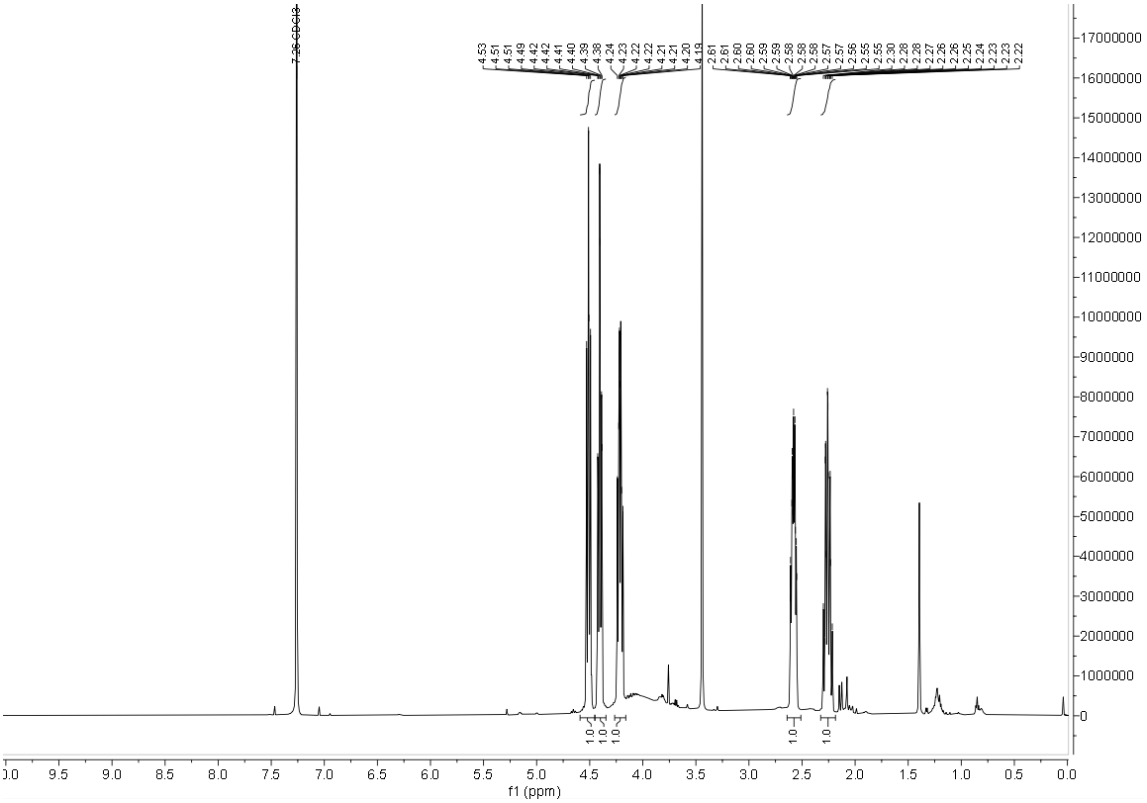


^{13}C NMR: (126 MHz, $CDCl_3$)

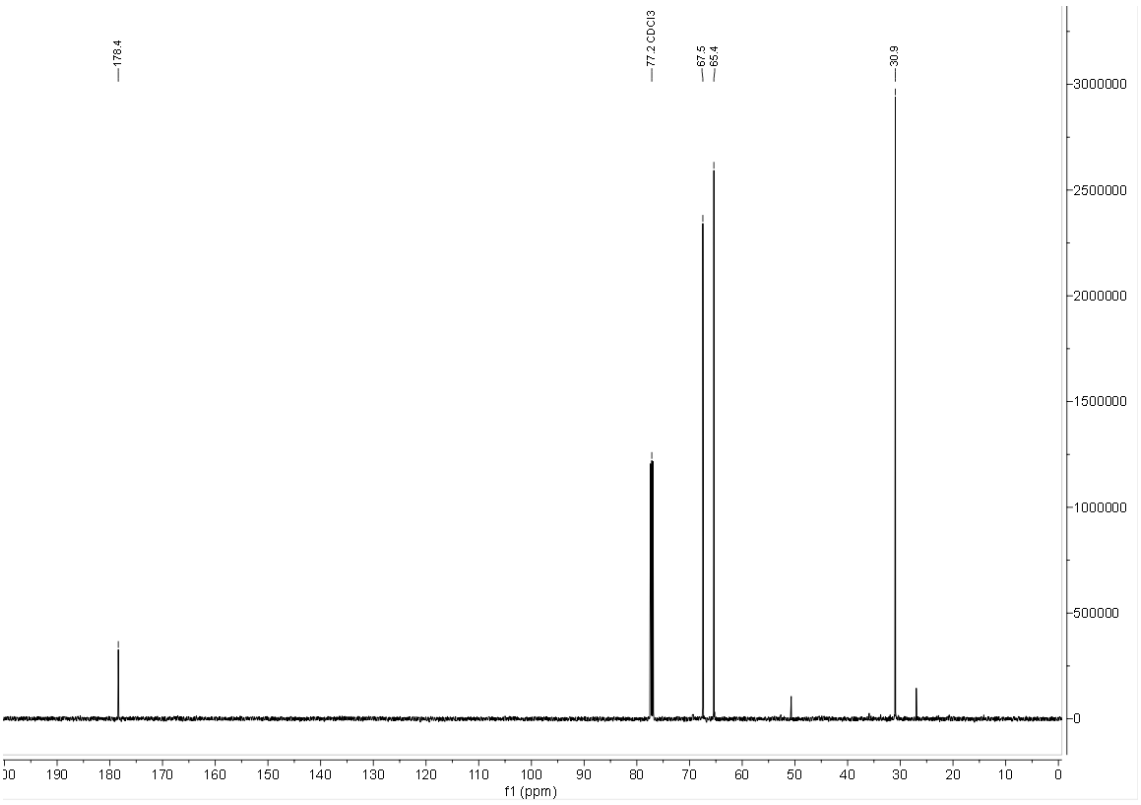


(*S*)-2-Hydroxy-γ-butyrolactone (***S*-16**)

^{1}H NMR: (500 MHz, $CDCl_3$)

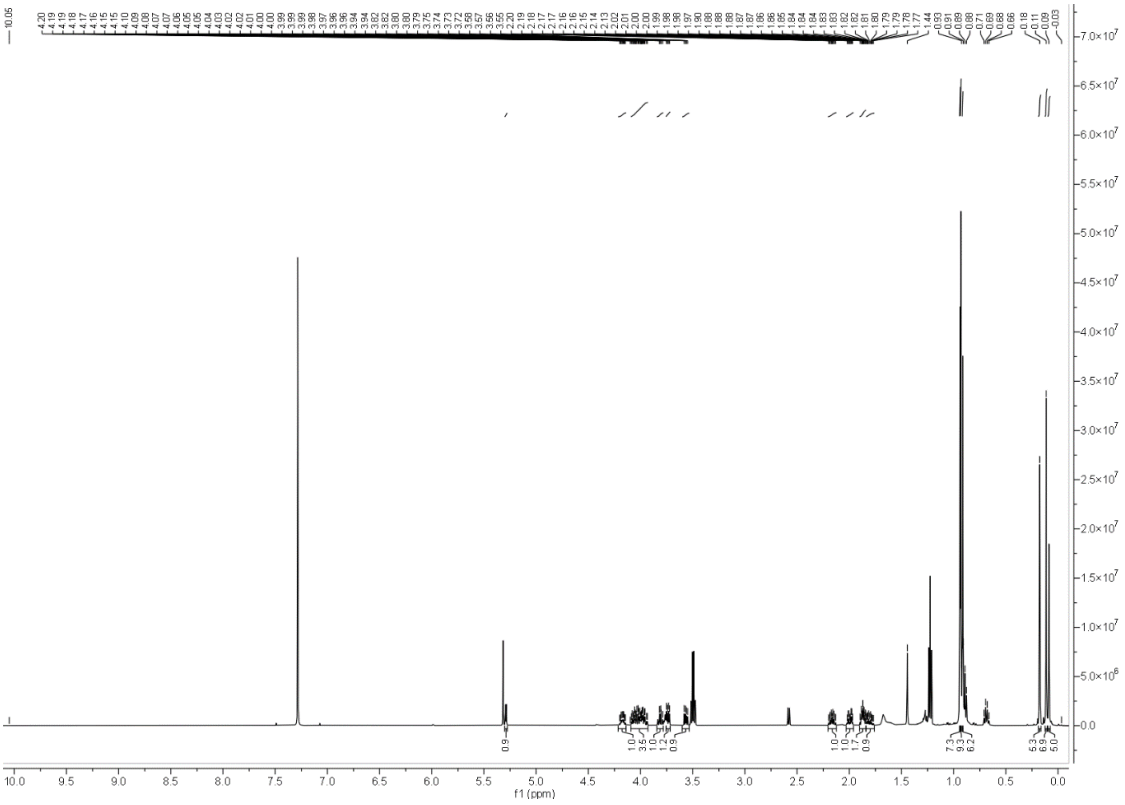


^{13}C NMR: (126 MHz, $CDCl_3$)

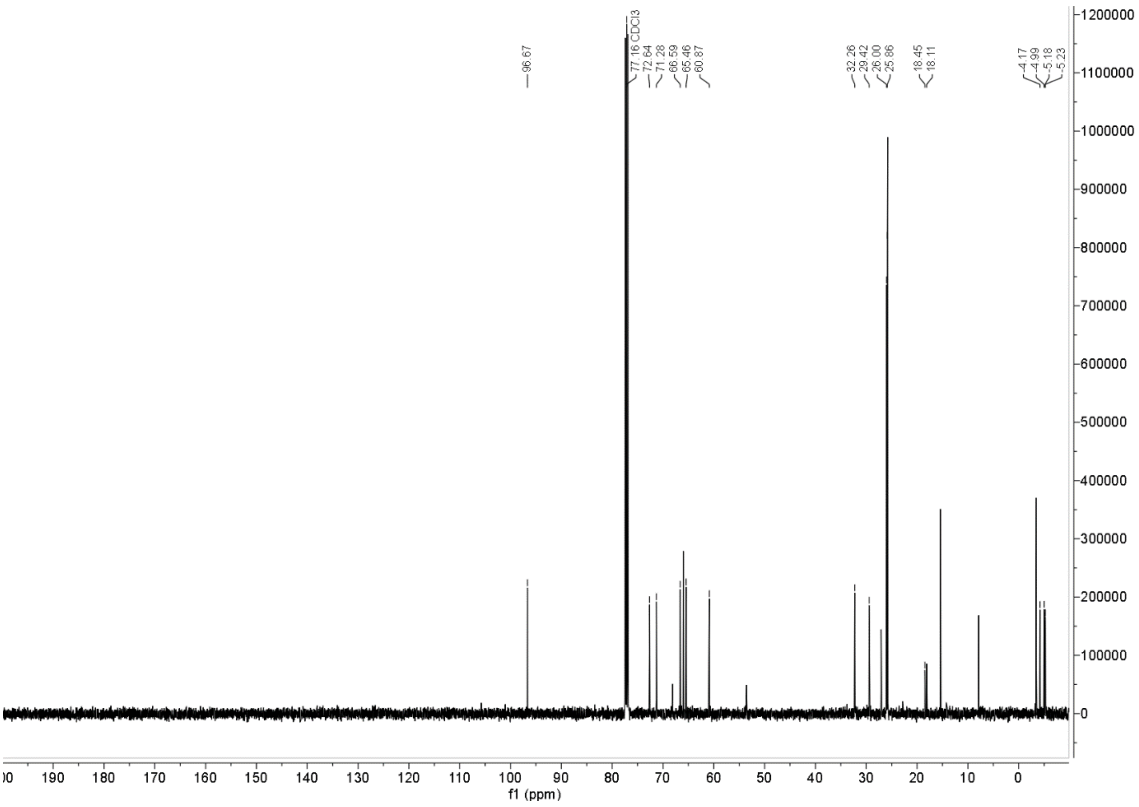


2-*O*-(*tert*-Butyldimethylsilyl)-3-deoxy-D/L-*glycero*-tetrose (***rac*-18TBS**, product A)

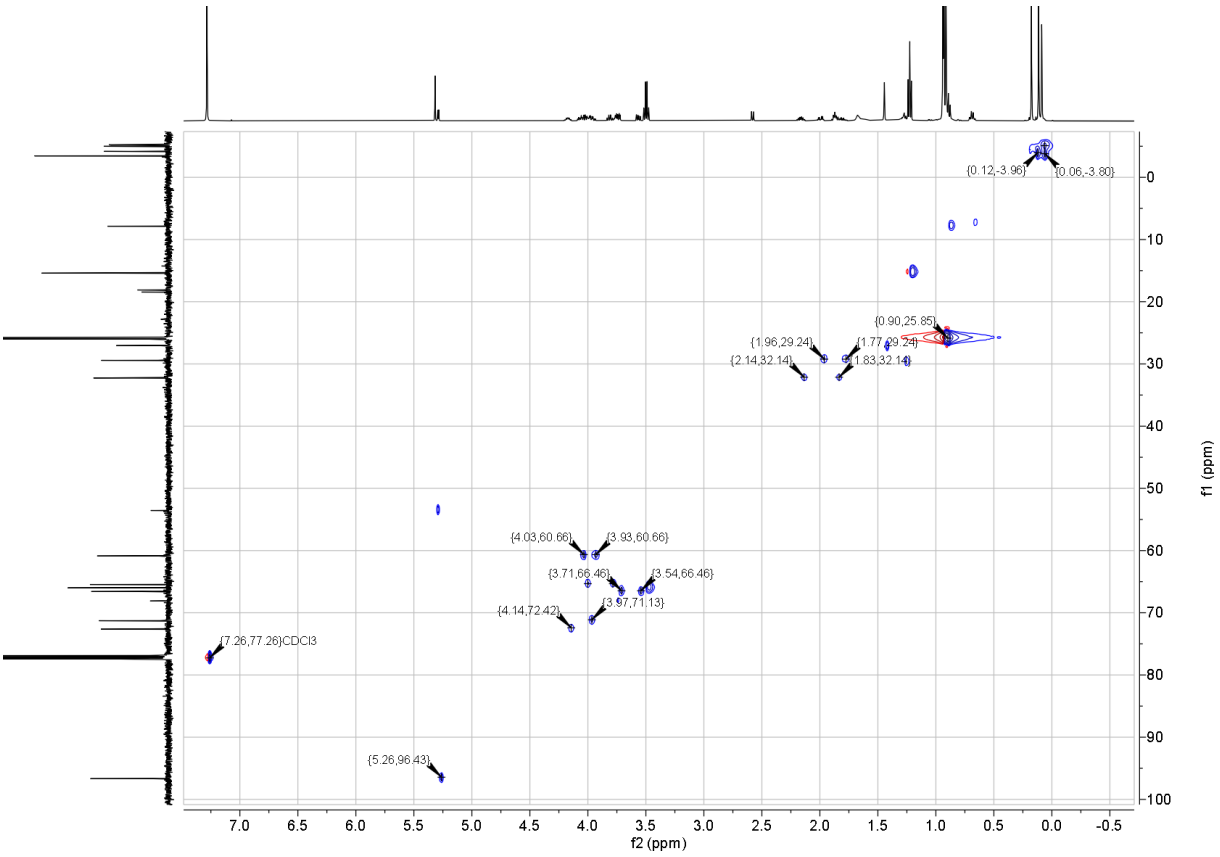
^{1}H NMR (500 MHz, $CDCl_3$):



^{13}C NMR (127 MHz, $CDCl_3$):

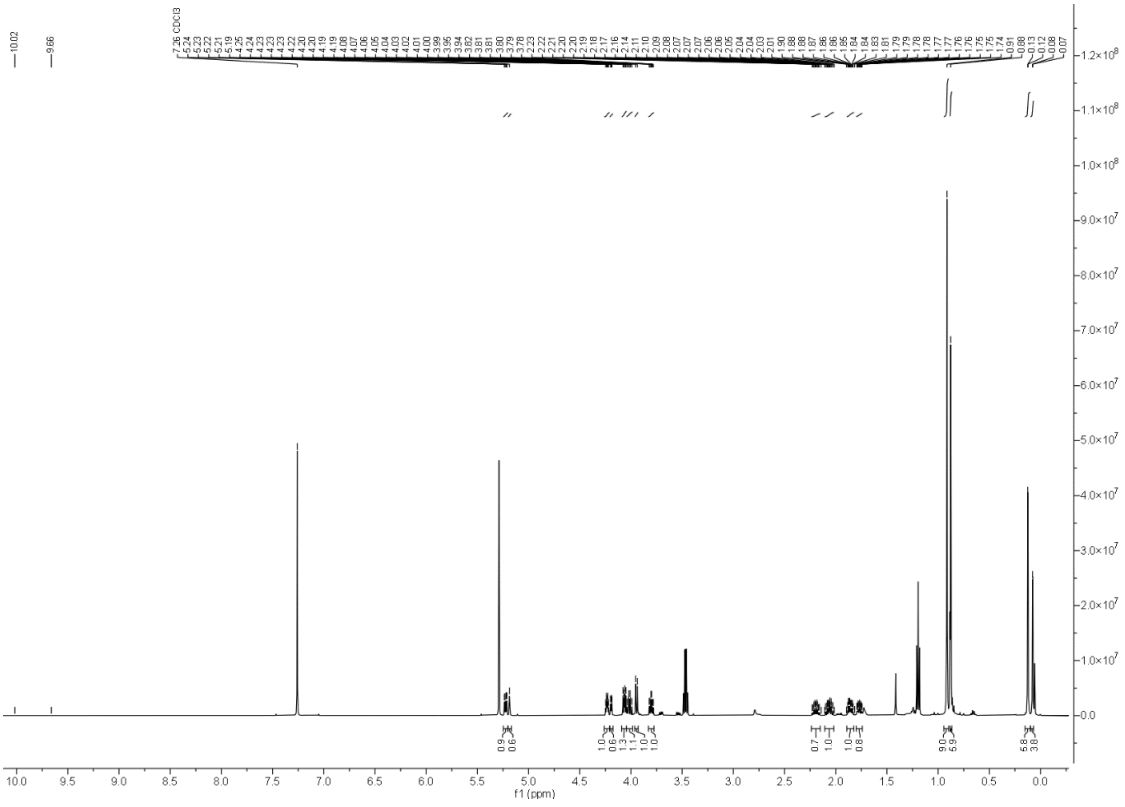


HSQC: ($CDCl_3$).

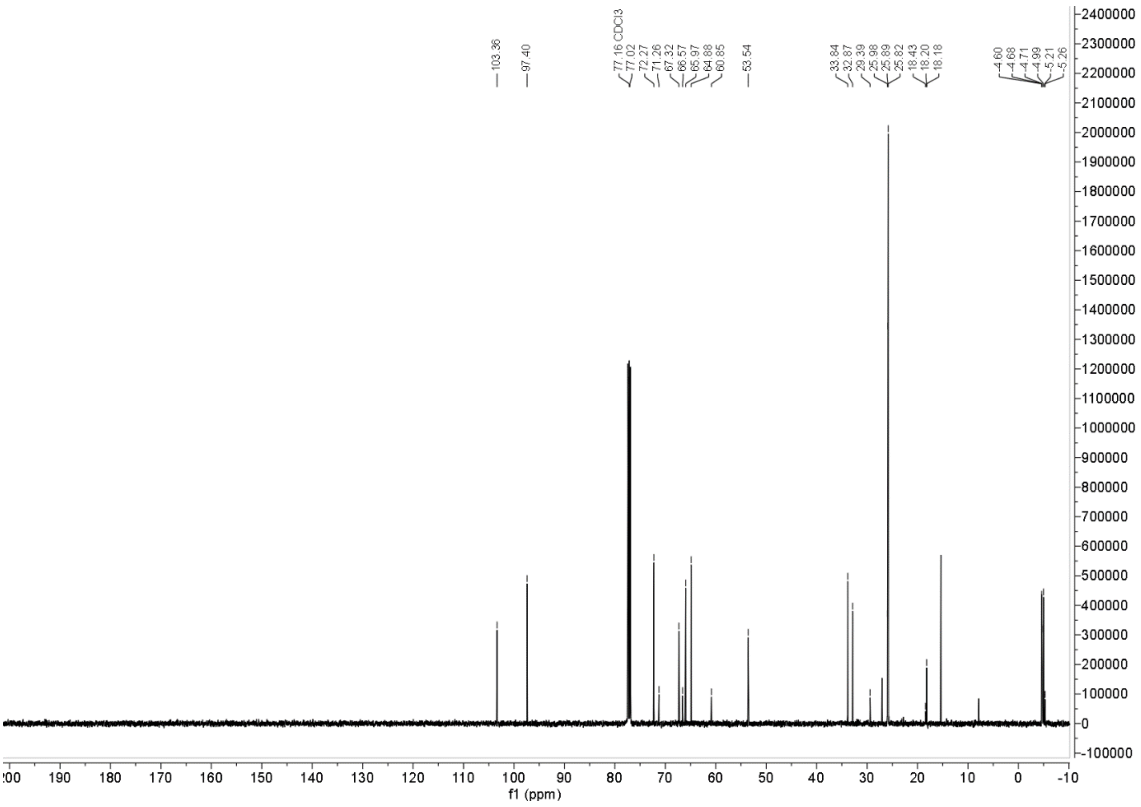

{0.12,-3.96}
{0.06,-3.80}
{0.90,25.85}
{1.96,29.24}
{1.77,29.24}
{2.14,32.14}
{1.83,32.14}
{4.03,60.66}
{3.93,60.66}
{3.71,66.46}
{3.54,66.46}
{3.97,71.13}
{4.14,72.42}
{7.26,77.26}CDCl3
{5.26,96.43}
f2 (ppm)
f1 (ppm)

2-*O*-(*tert*-Butyldimethylsilyl)-3-deoxy-D/L-*glycero*-tetrose (***rac*-18TBS**, product B)

^{1}H NMR (500 MHz, $CDCl_3$):

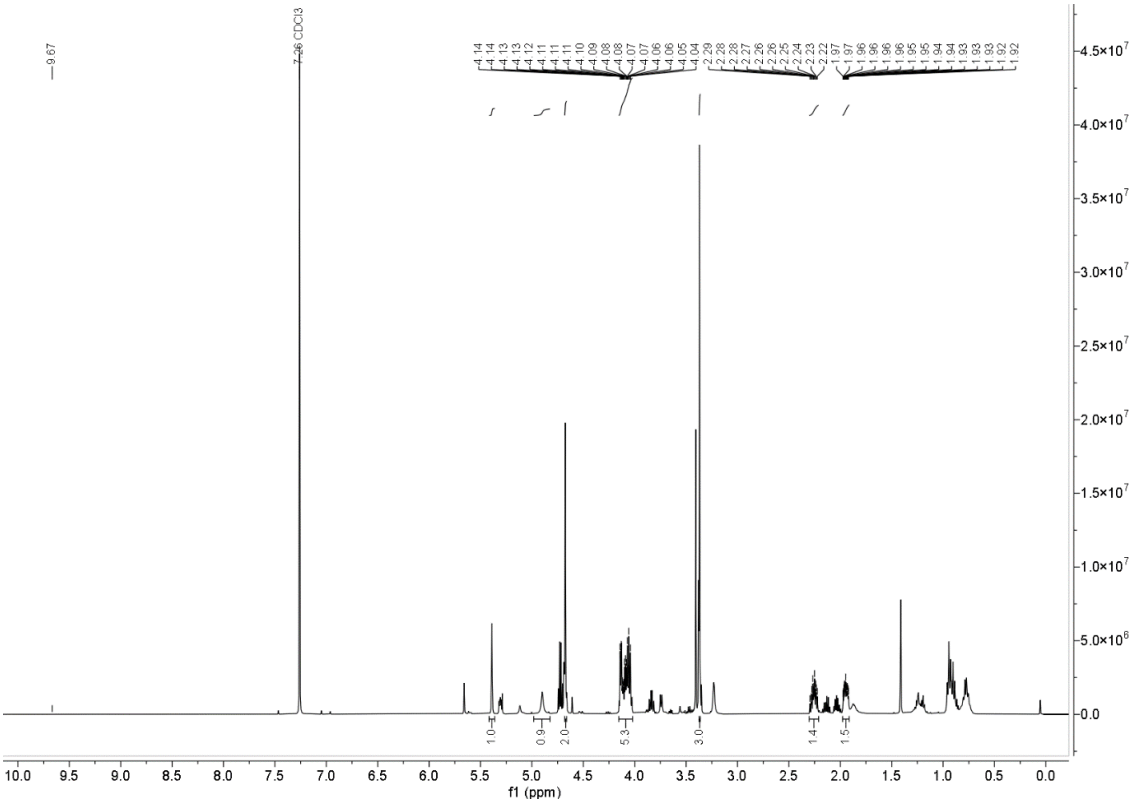


^{13}C NMR (127 MHz, $CDCl_3$):

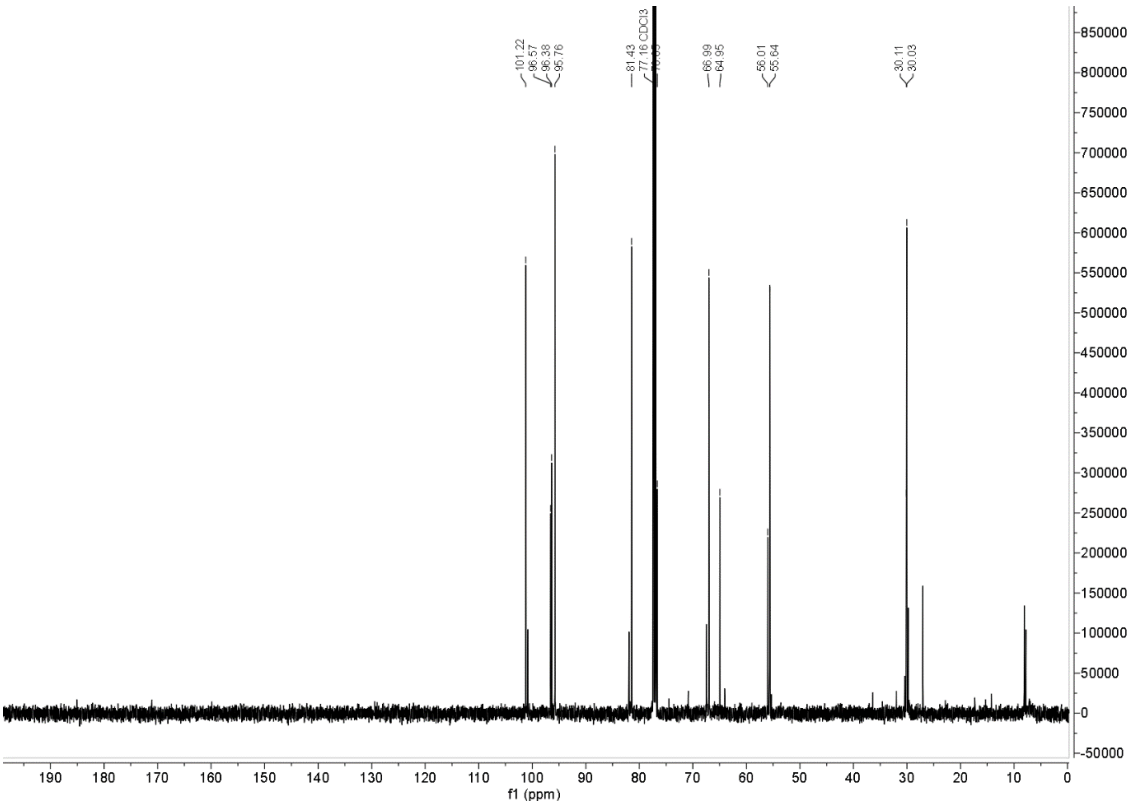


3-Deoxy-2-*O*-(methoxymethyl)-D/L-*glycero*-tetrose (***rac*-18MOM**)

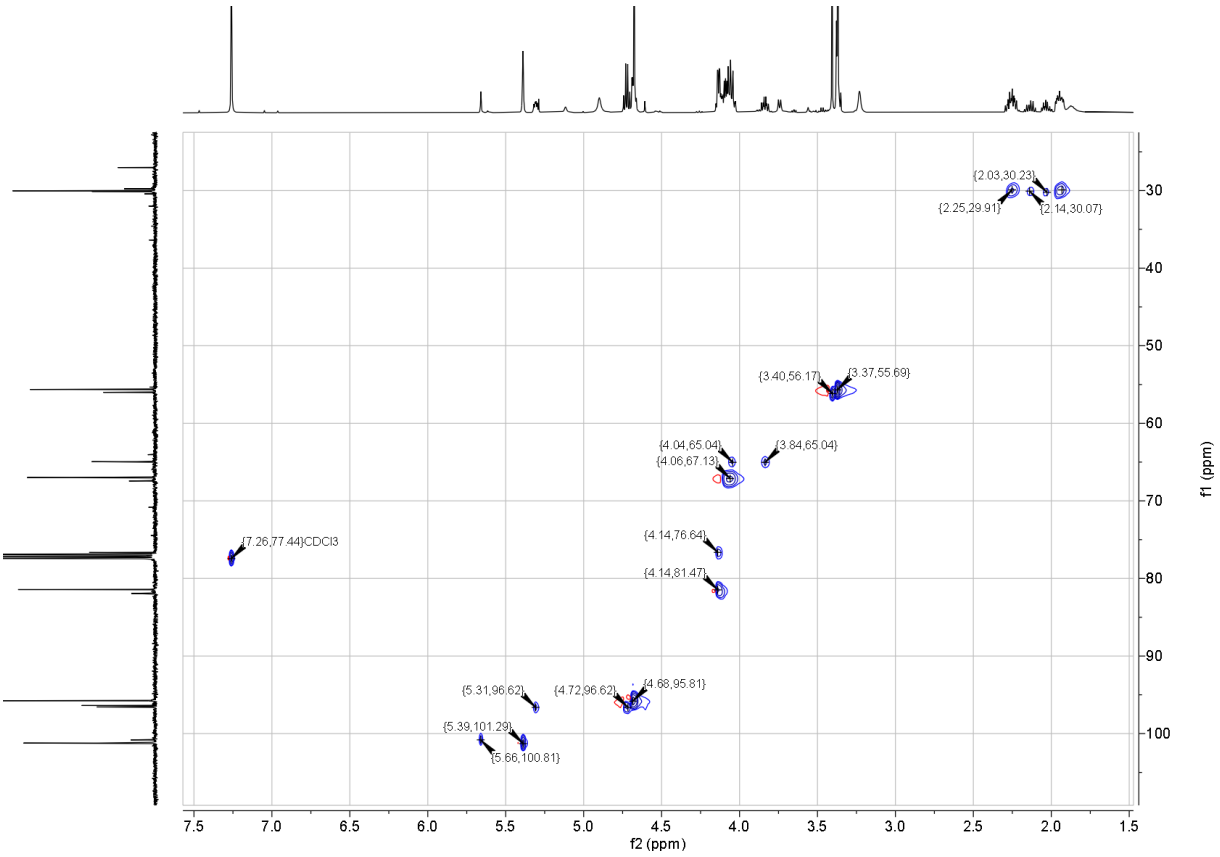
^{1}H NMR (500 MHz, $CDCl_3$):



^{13}C NMR (127 MHz, $CDCl_3$):

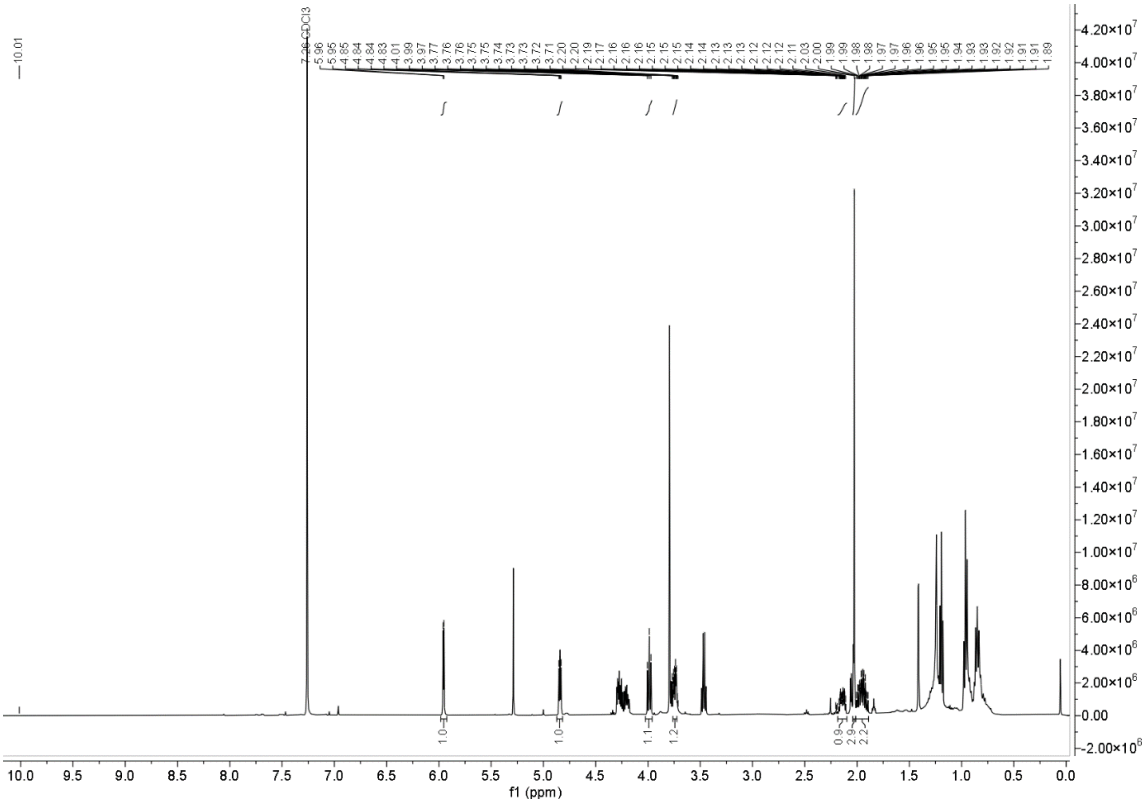


HSQC ($CDCl_3$):

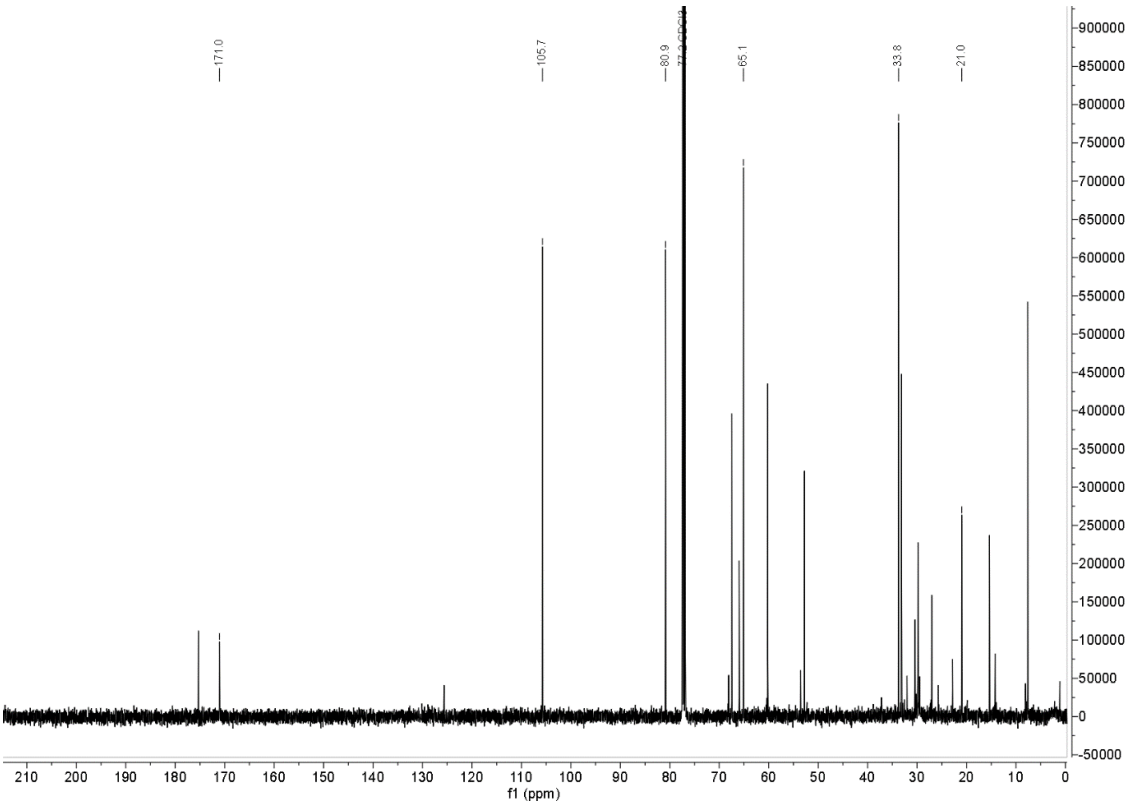

{2.03,30.23}
{2.25,29.91}
{2.14,30.07}
{3.40,56.17}
{3.37,55.69}
{4.04,65.04}
{4.06,67.13}
{3.84,65.04}
{7.26,77.44}CDCl3
{4.14,76.64}
{4.14,81.47}
{5.31,96.62}
{4.72,96.62}
{4.68,95.81}
{5.39,101.29}
{5.66,100.81}
f1 (ppm)
f2 (ppm)

2-*O*-Acetyl-3-deoxy-D/L-*glycero*-tetrose (***rac*-18Ac**)

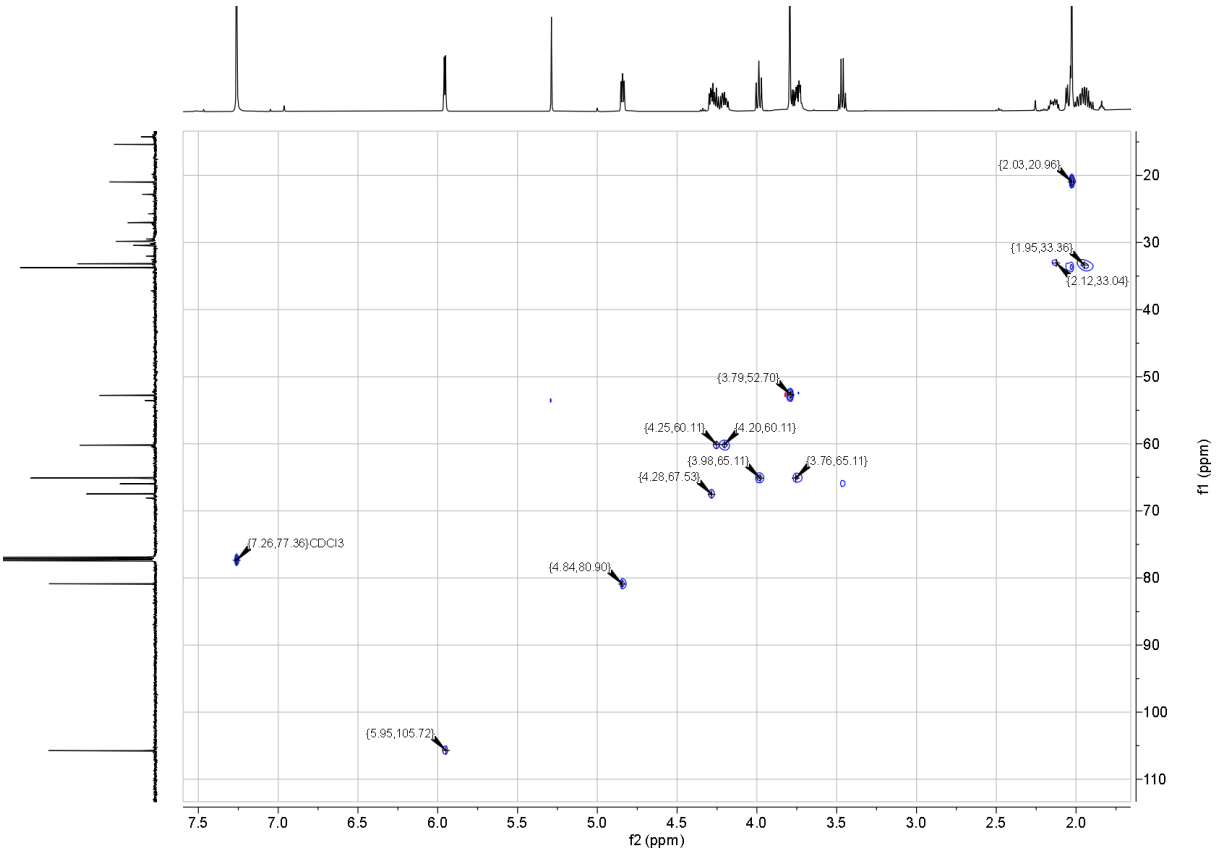
^{1}H NMR (500 MHz, $CDCl_3$):



^{13}C NMR (127 MHz, $CDCl_3$):

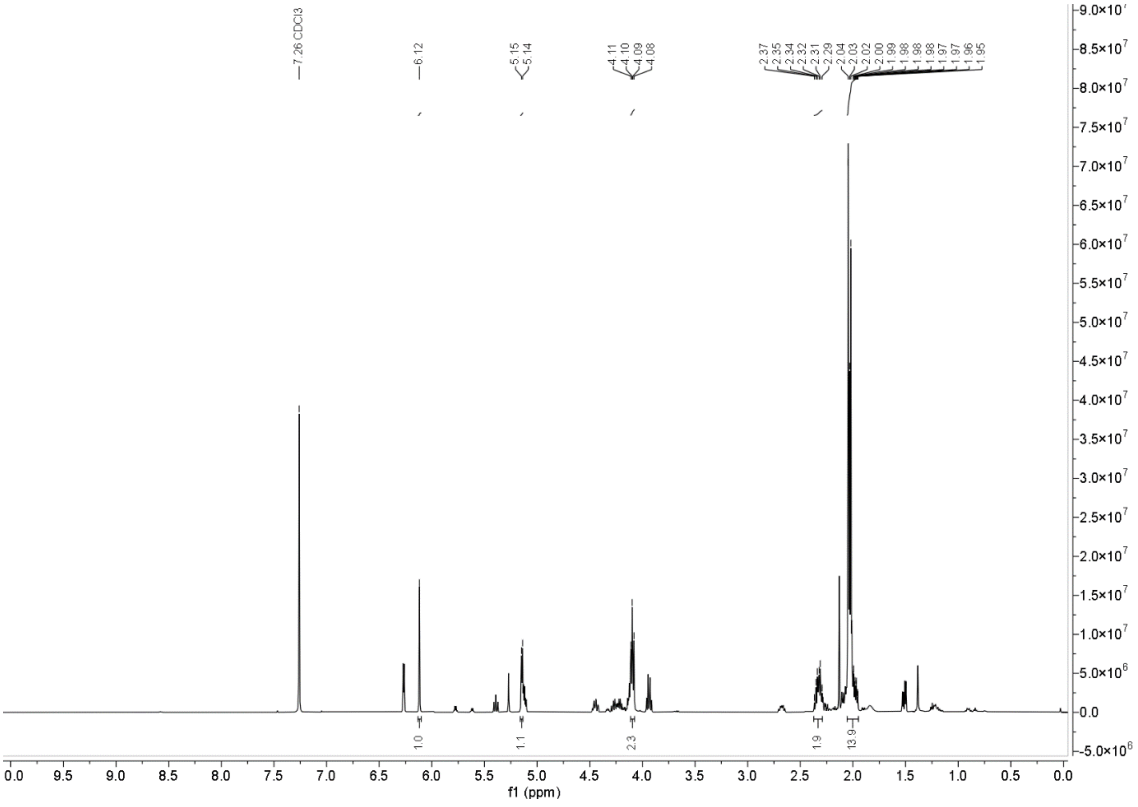


HSQC ($CDCl_3$):

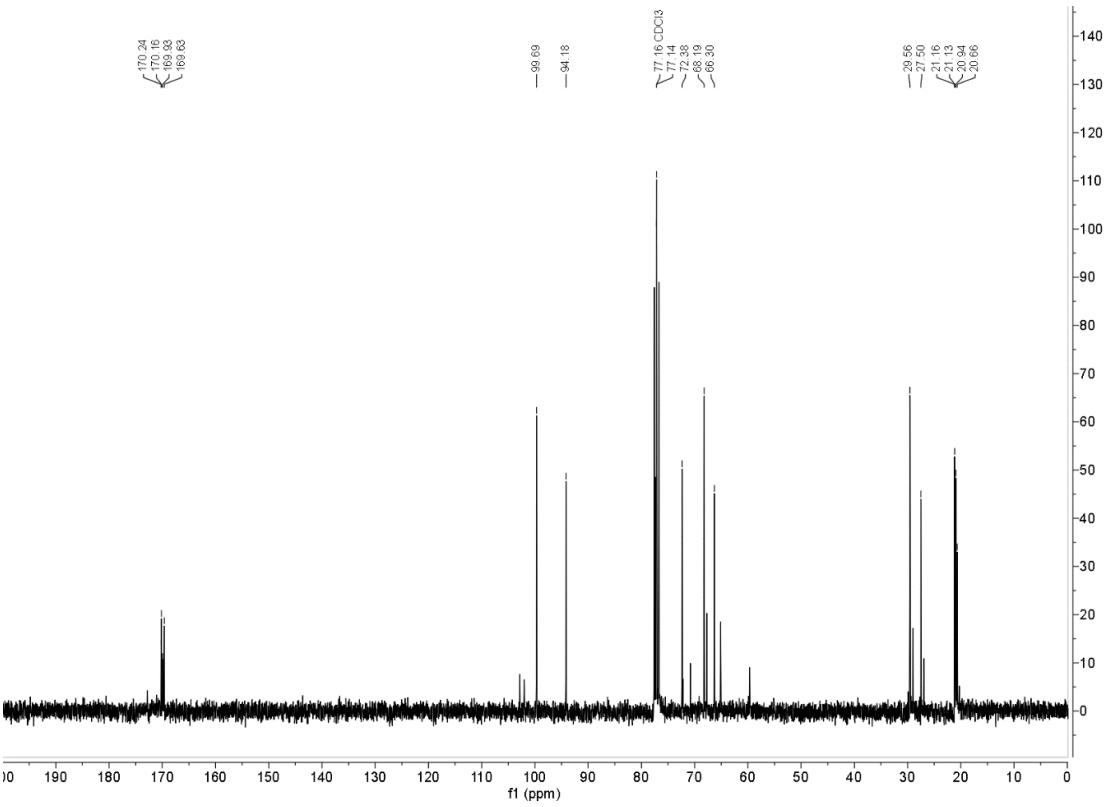

{2.03,20.96}
{1.95,33.36}
{2.12,33.04}
{3.79,52.70}
{4.25,60.11}
{4.20,60.11}
{3.98,65.11}
{3.76,65.11}
{4.28,67.53}
{7.26,77.36}CDCl3
{4.84,80.90}
{5.95,105.72}
f2 (ppm)
f1 (ppm)

2,3(*R*)-Diacetoxytetrahydrofuran (***R*-13**, product A):

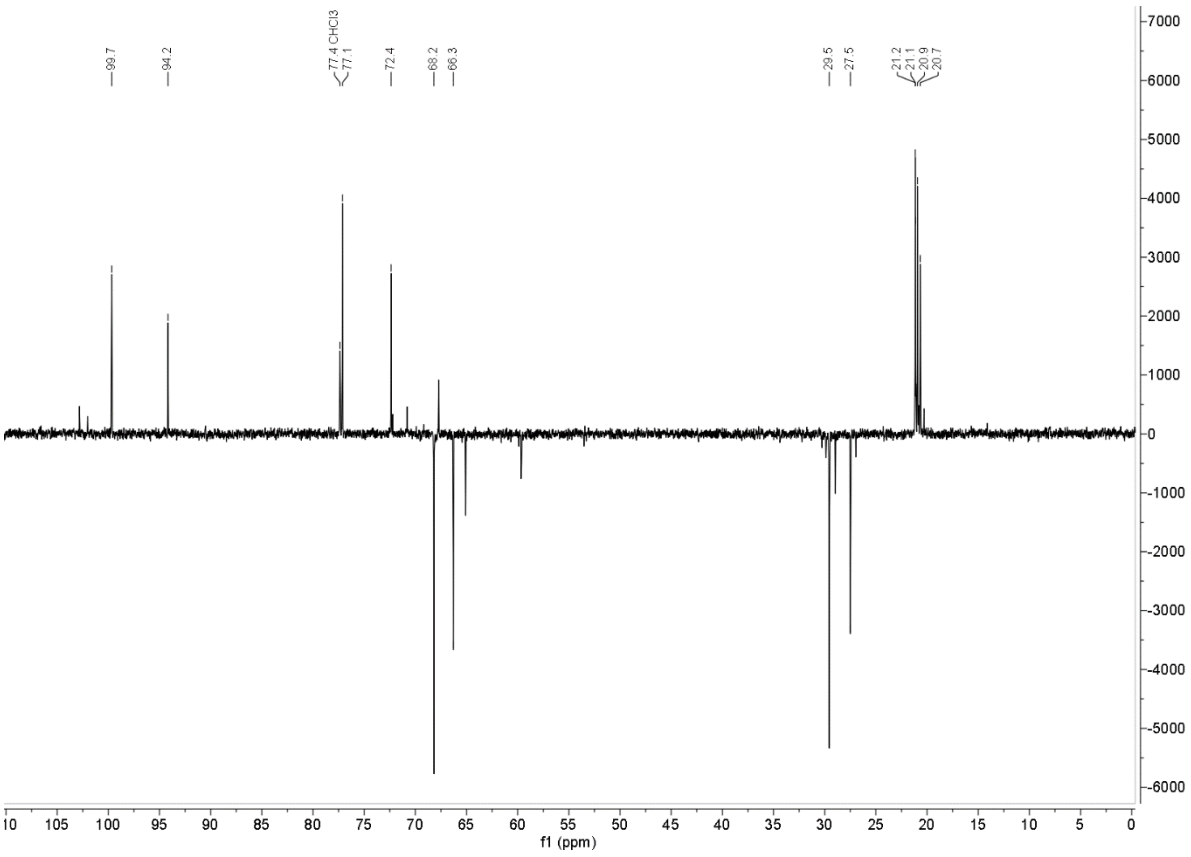
^{1}H NMR (500 MHz, $CDCl_3$):



^{13}C NMR (127 MHz, $CDCl_3$):

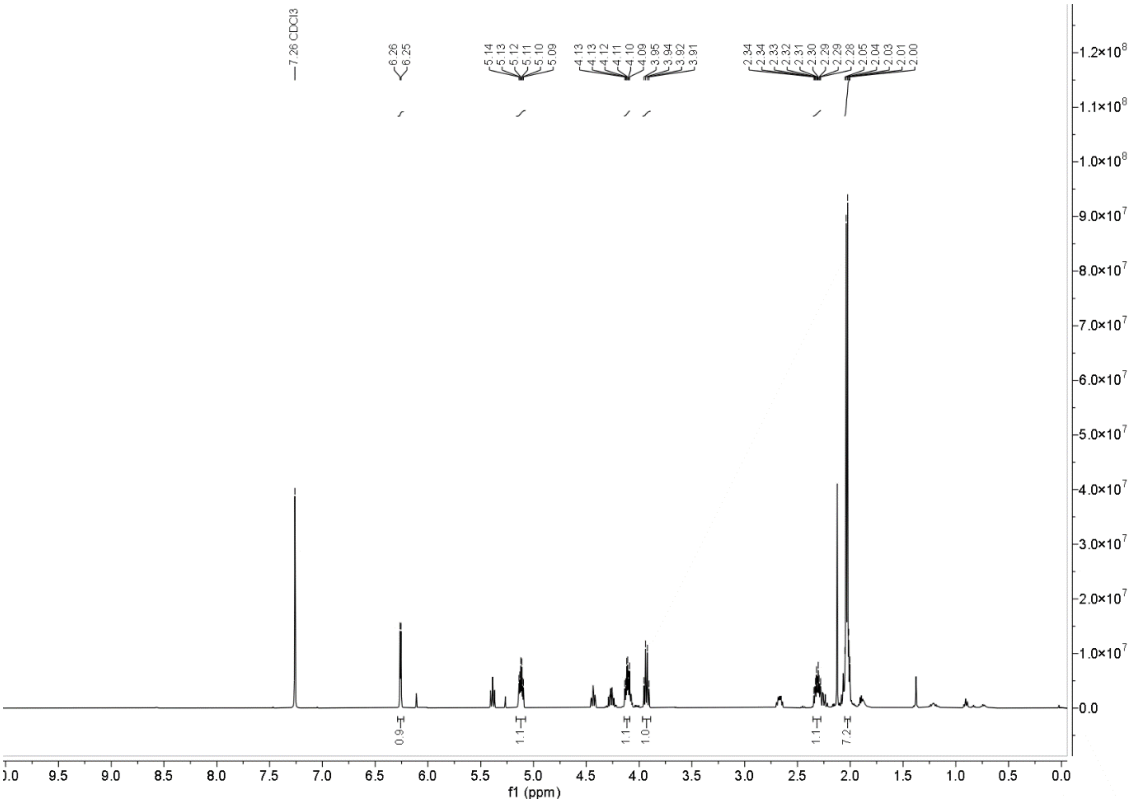


^{13}C NMR (DEPT, 127 MHz, $CDCl_3$, $CHCl_3$ – 77.3574 ppm):

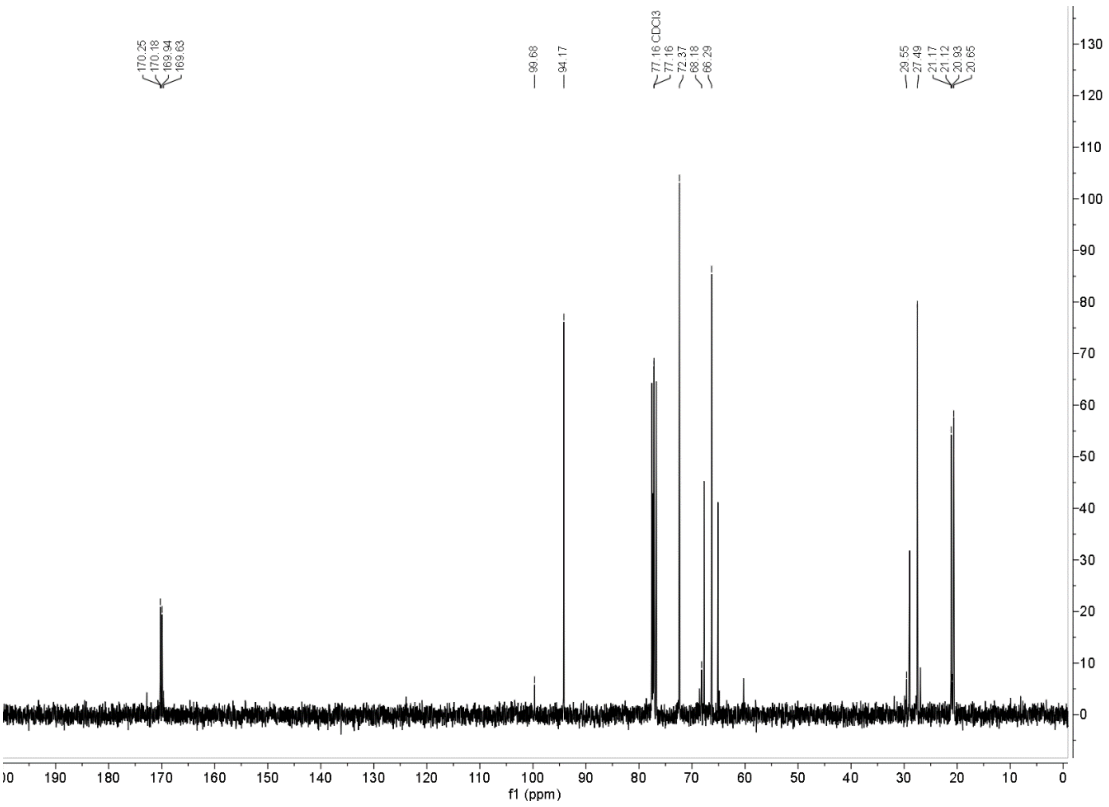


2,3(*R*)-Diacetoxytetrahydrofuran (***R*-13**, product B):

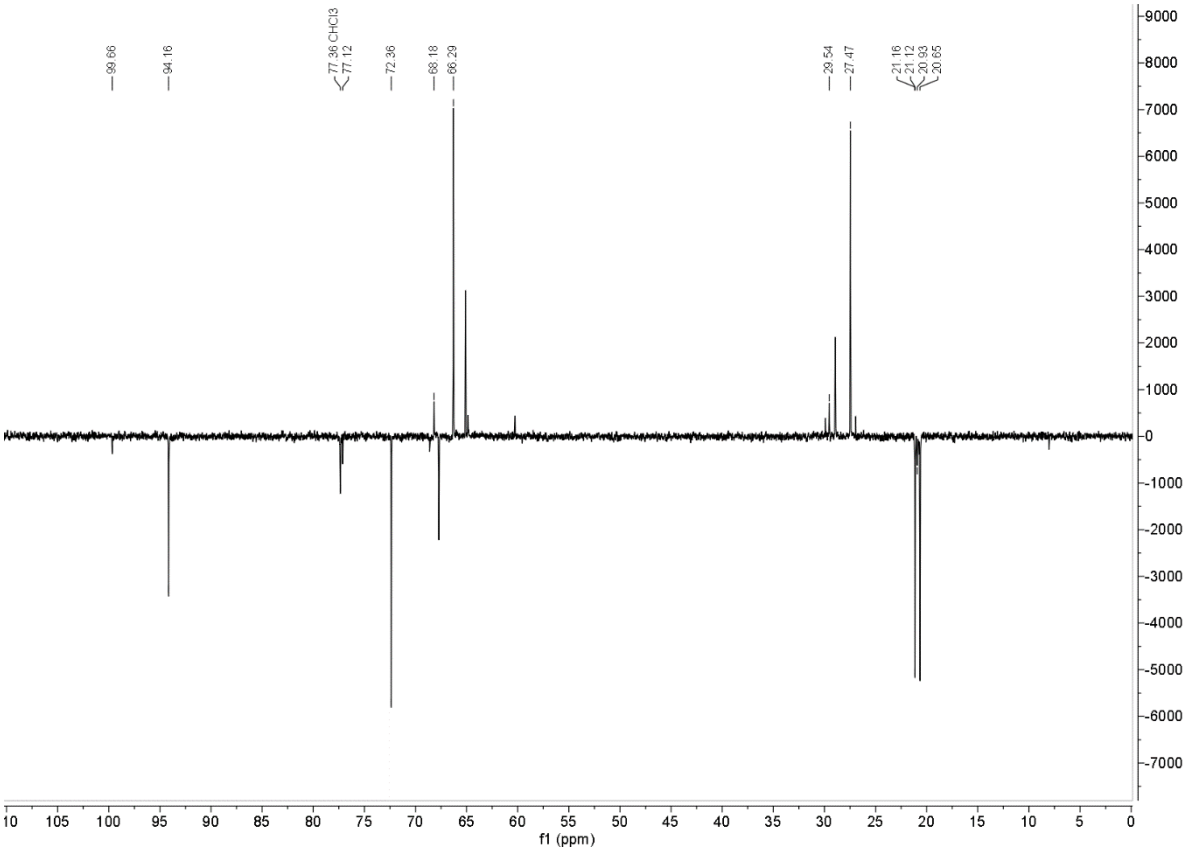
^{1}H NMR (500 MHz, $CDCl_3$):



^{13}C NMR (127 MHz, $CDCl_3$):

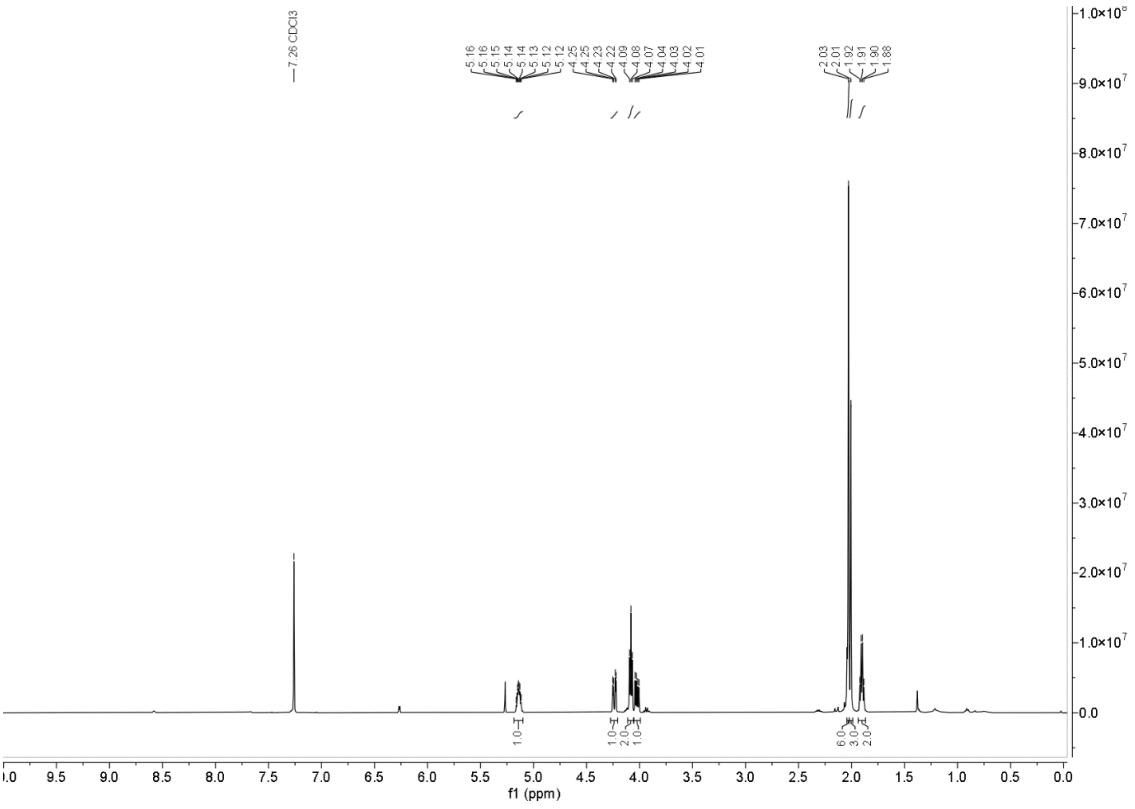


^{13}C NMR (DEPT, 127 MHz, $CDCl_3$, $CHCl_3$ – 77.3574 ppm):

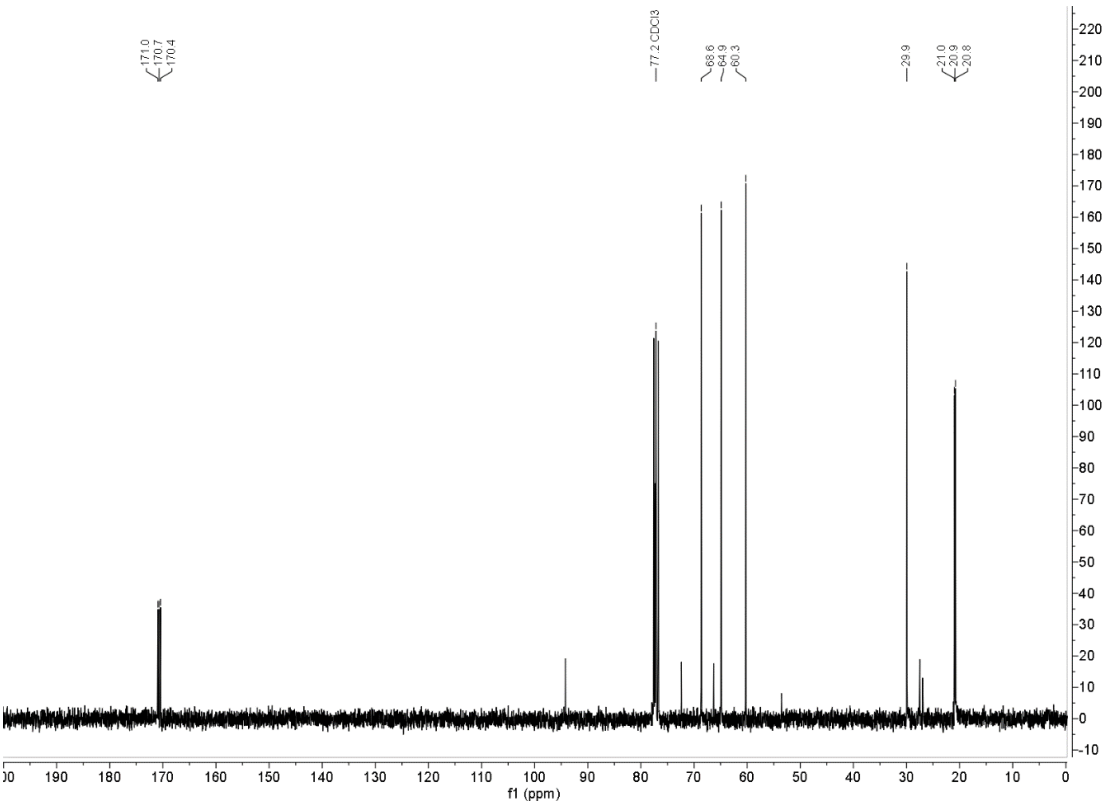


(R)-1,2,4-Butanetriyl acetate (product C):

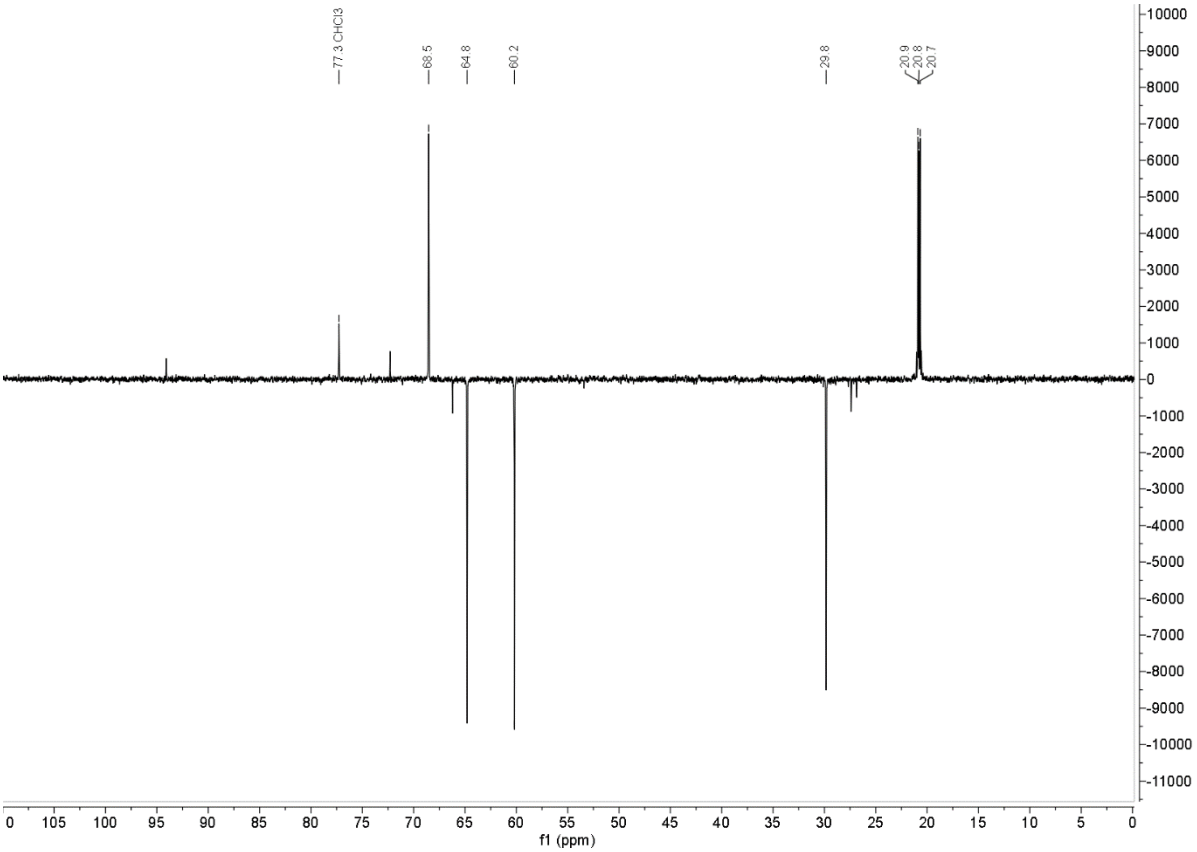
^{1}H NMR (500 MHz, $CDCl_3$):



^{13}C NMR (127 MHz, $CDCl_3$):

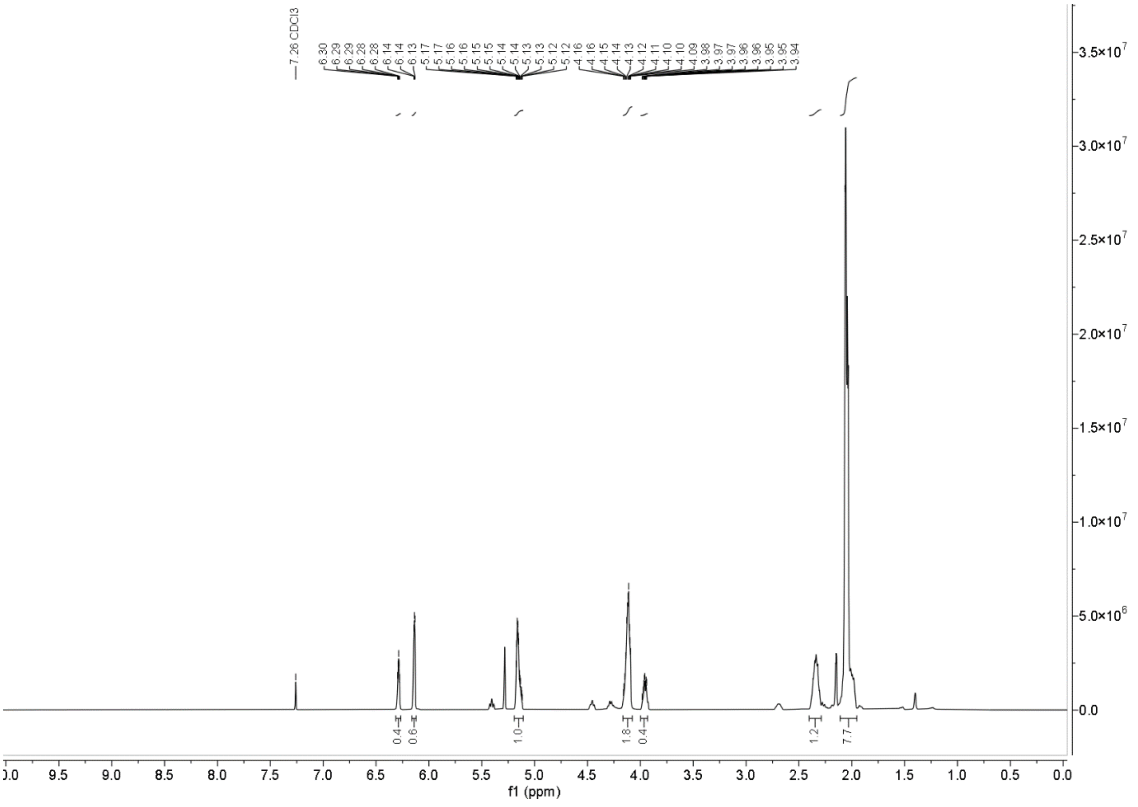


^{13}C NMR (DEPT, 127 MHz, $CDCl_3$, $CHCl_3$ – 77.3574 ppm):

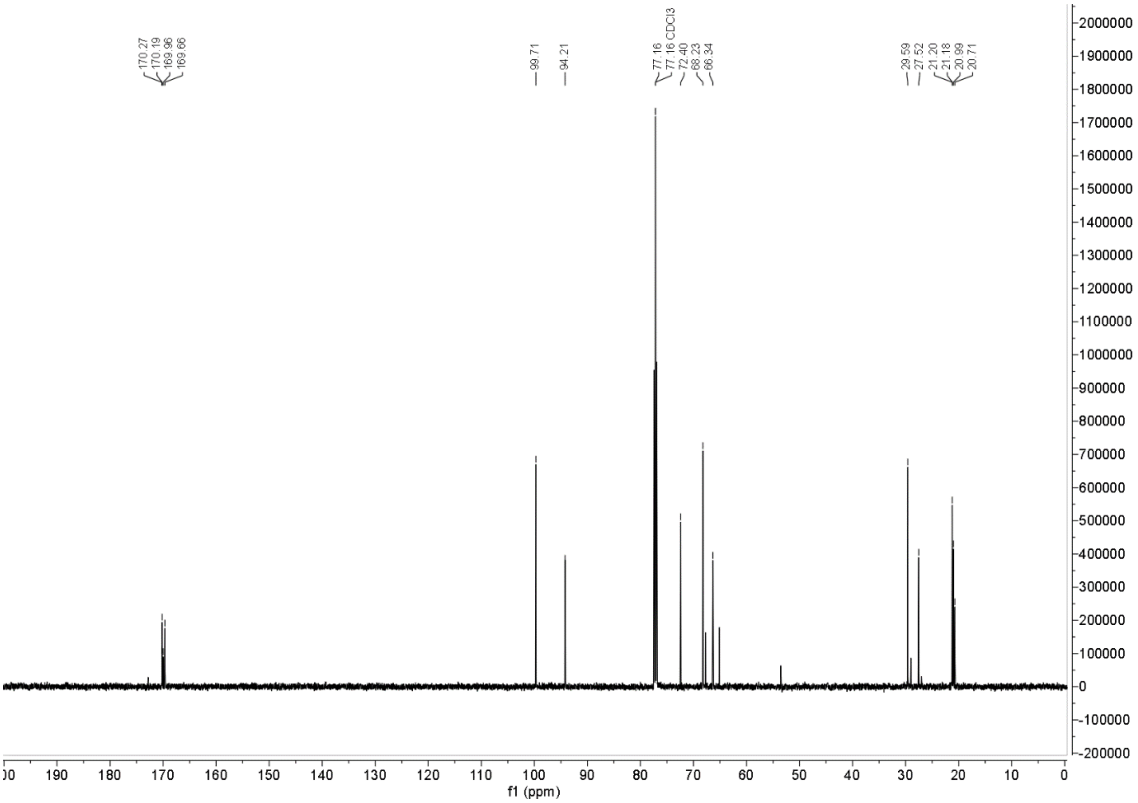


2,3(*R*)-Diacetoxytetrahydrofuran (***R*-13**)

^{1}H NMR (500 MHz, $CDCl_3$):

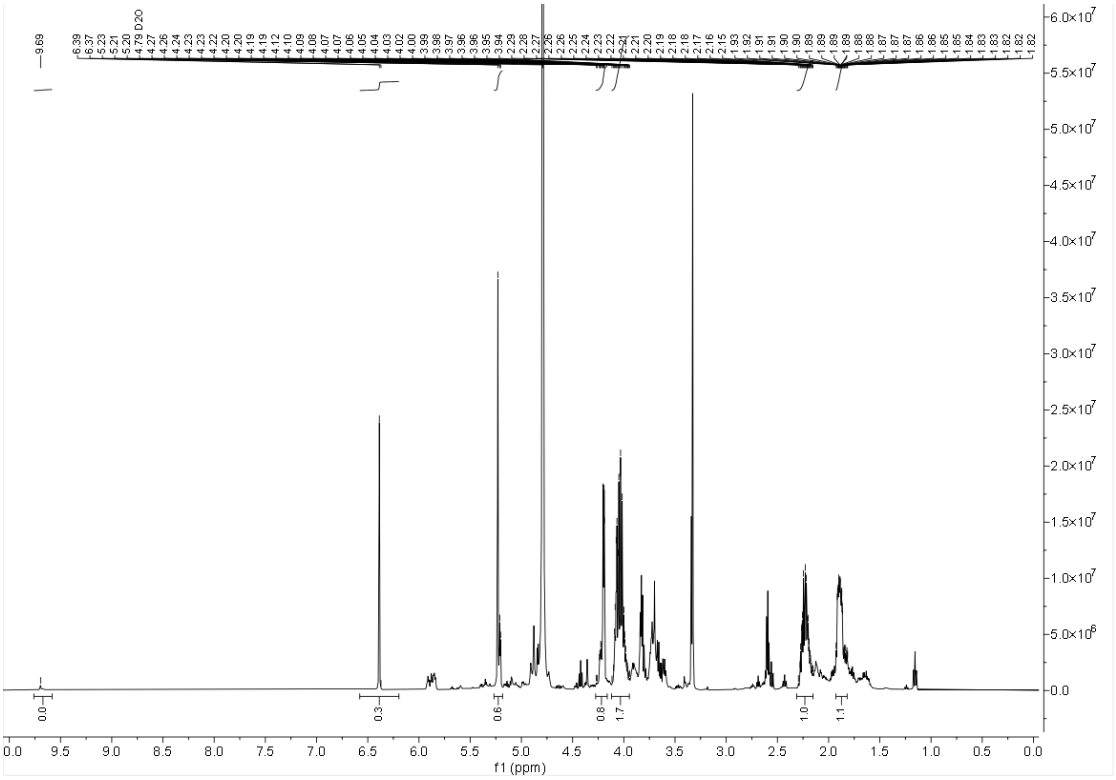


^{13}C NMR (127 MHz, $CDCl_3$):

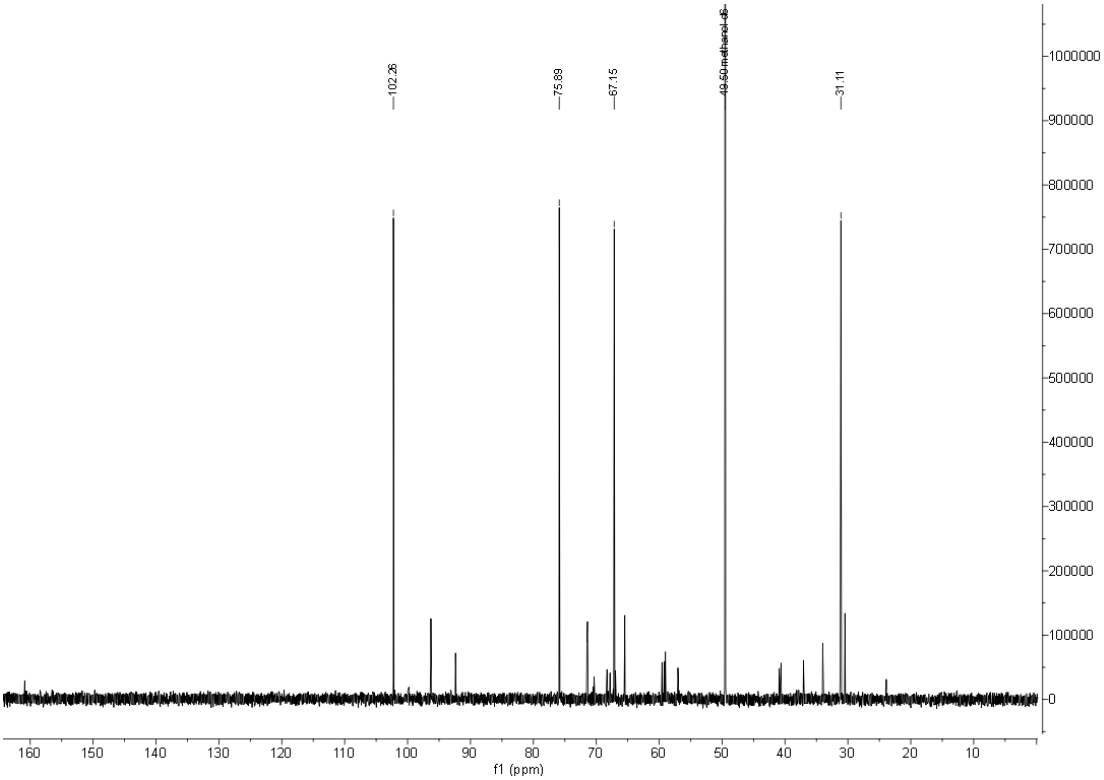


3-Deoxy-D/L-*glycero*-tetrose (***rac*-11**)

^{1}H NMR (500 MHz, D_2O):

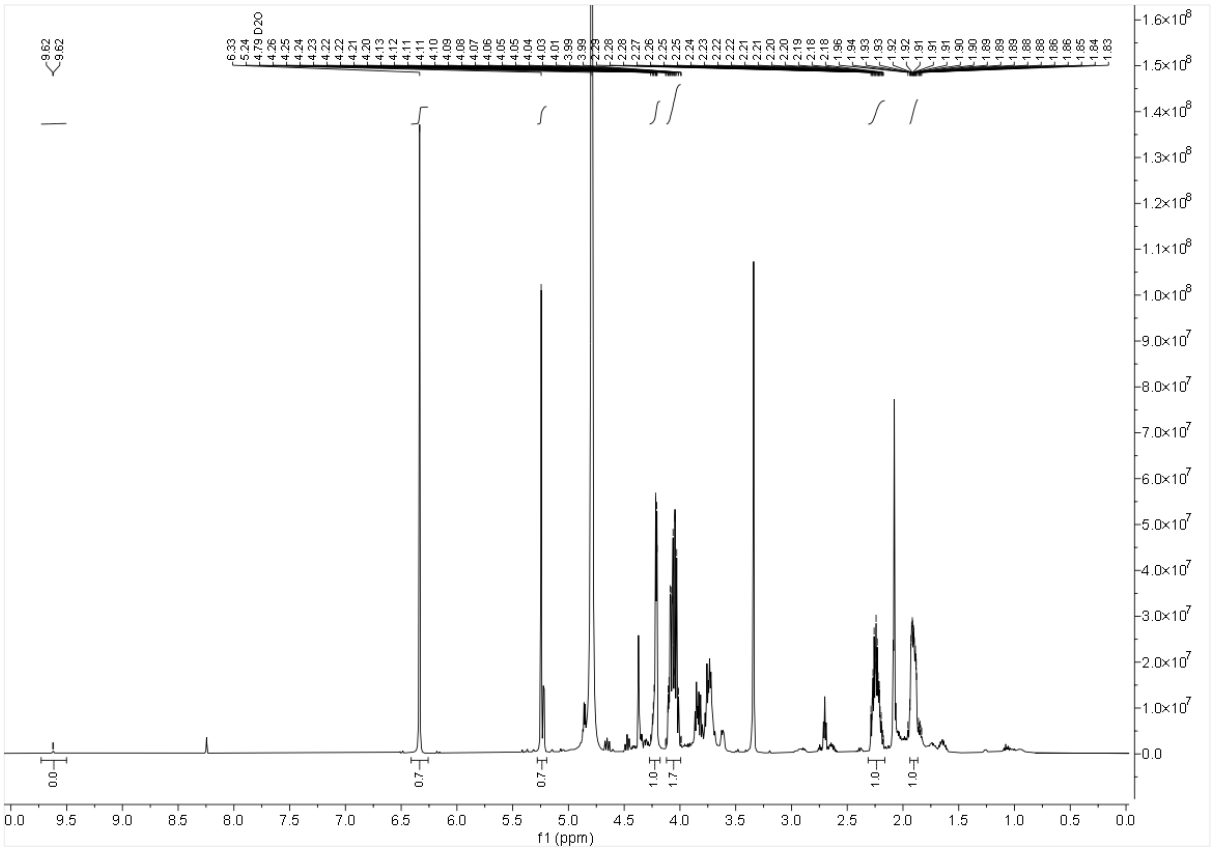


^{13}C NMR (126 MHz, D_2O, methanol at 49.50 ppm, spectrum from previous synthesis, June 2022):



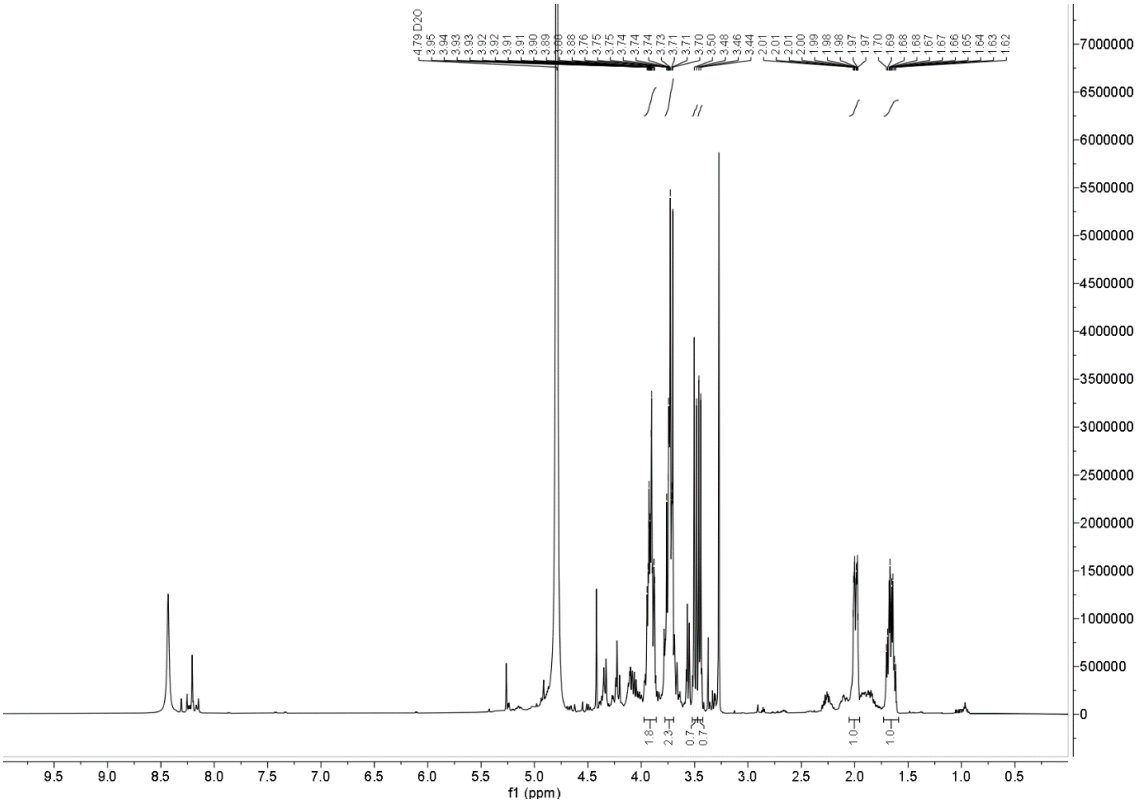
3-Deoxy-D-*glycero*-tetrose (***R*-11**)

^{1}H NMR (500 MHz, D_2O):

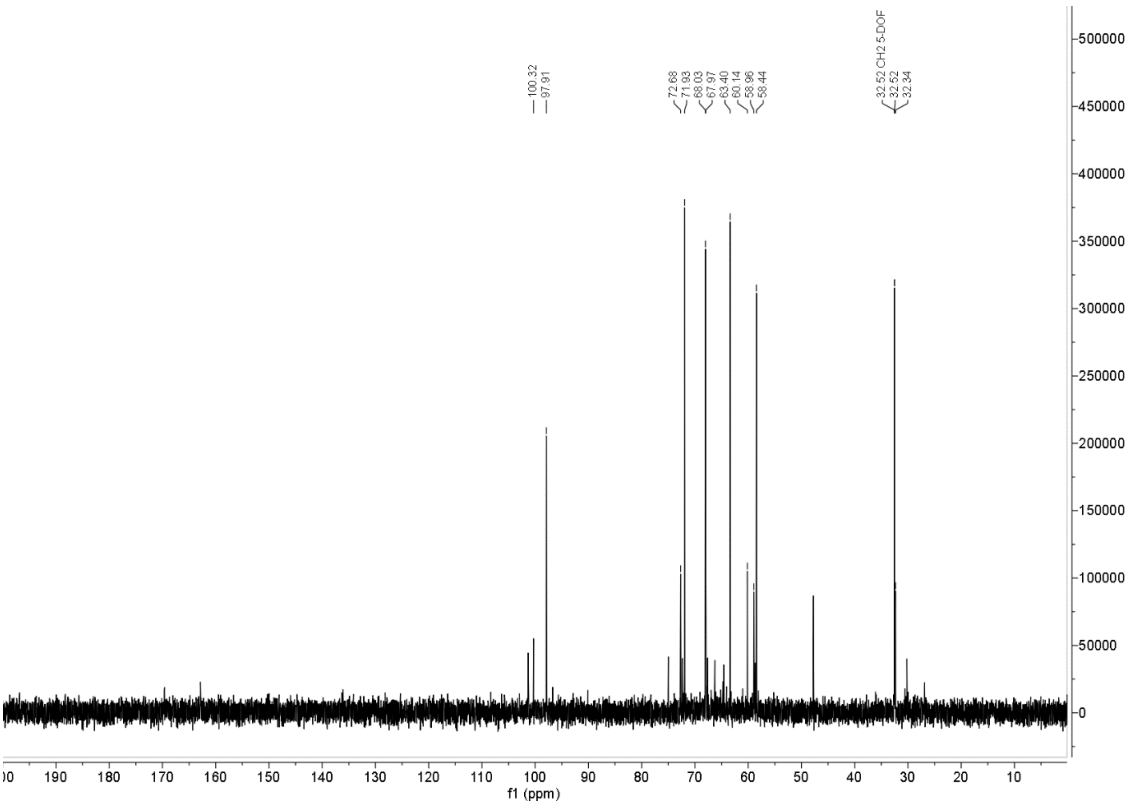


5-Deoxy-D-*threo*-hex-2-ulose (5-deoxyfructose, **30** from ***rac*-11**, with DAAO)

^{1}H NMR (500 MHz, D_2O):

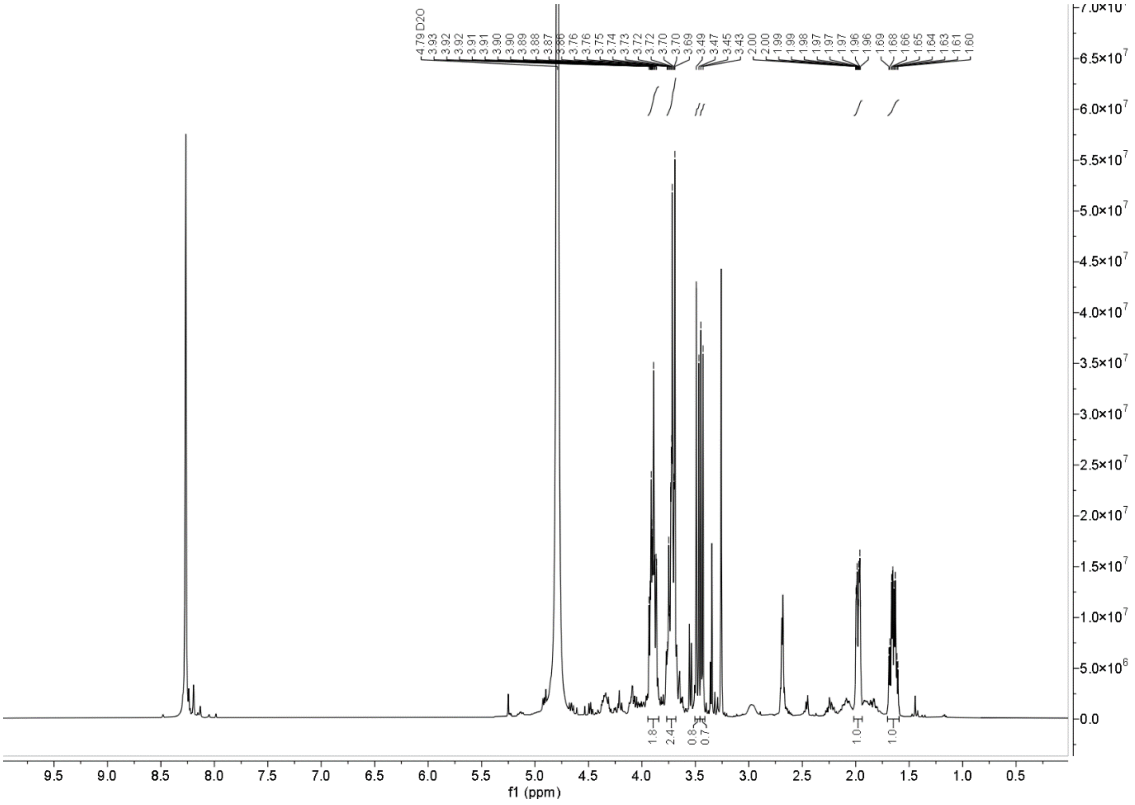


^{13}C NMR (126 MHz, D_2O):

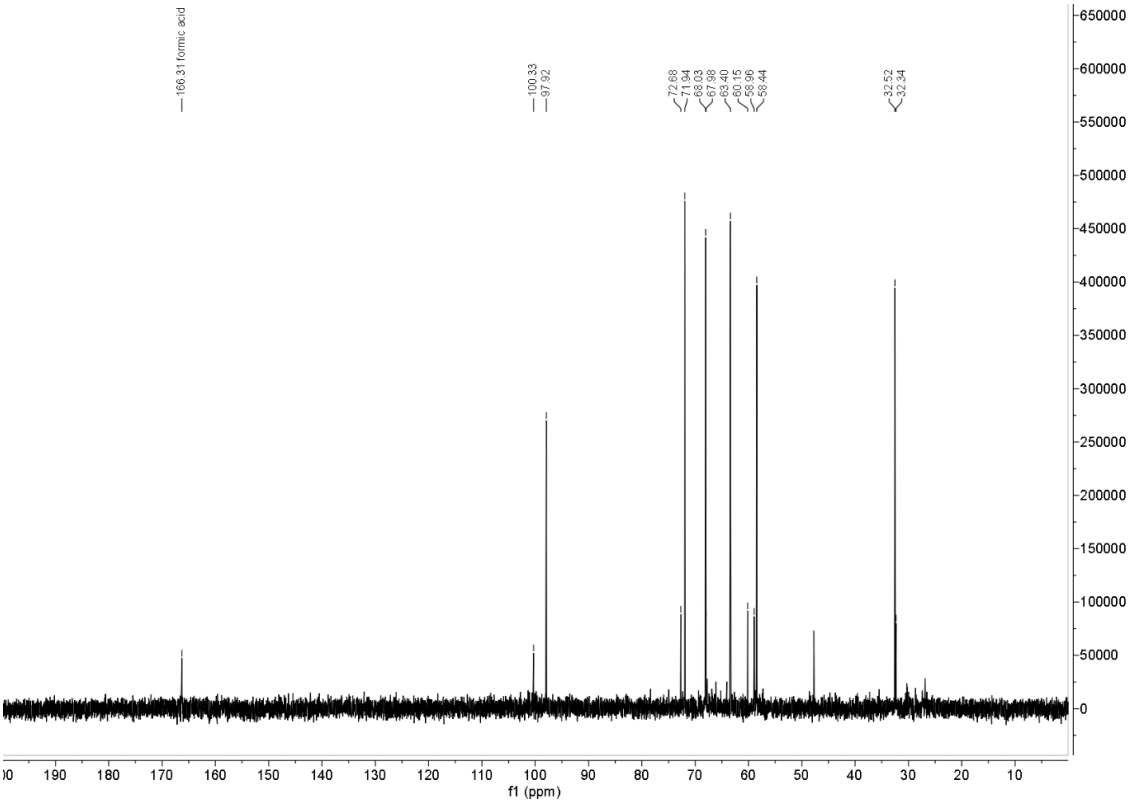


5-Deoxy-D-*threo*-hex-2-ulose (5-deoxyfructose, **30** from ***rac*-11**, with transaminase)

^{1}H NMR (500 MHz, D_2O):

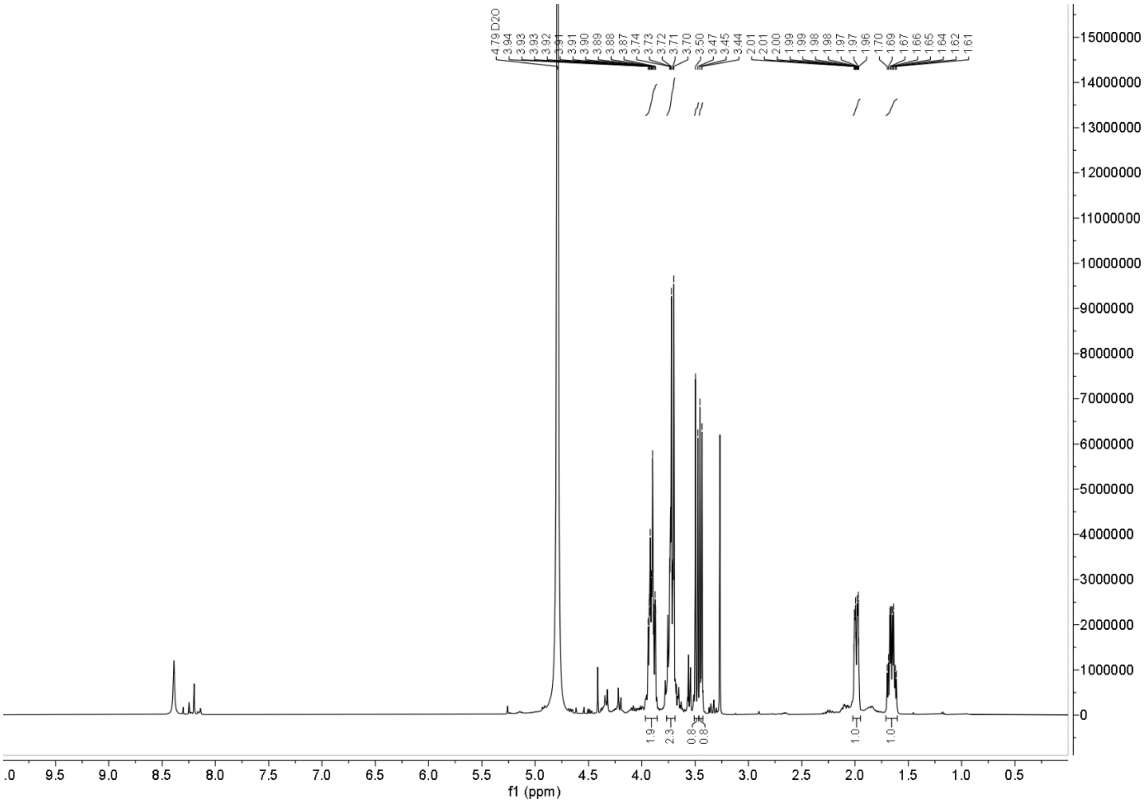


^{13}C NMR (126 MHz, D_2O):

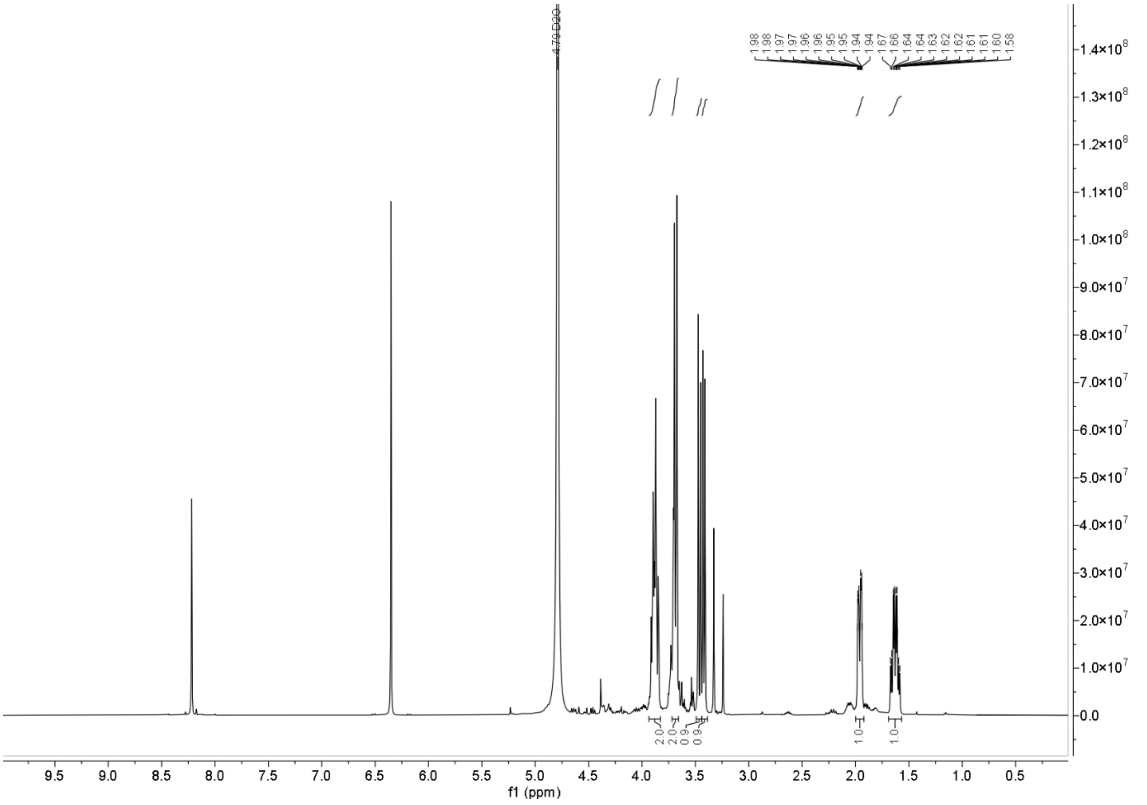


5-Deoxy-D-*threo*-hex-2-ulose (5-deoxyfructose, **30** from ***R*-11**, with DAAO)

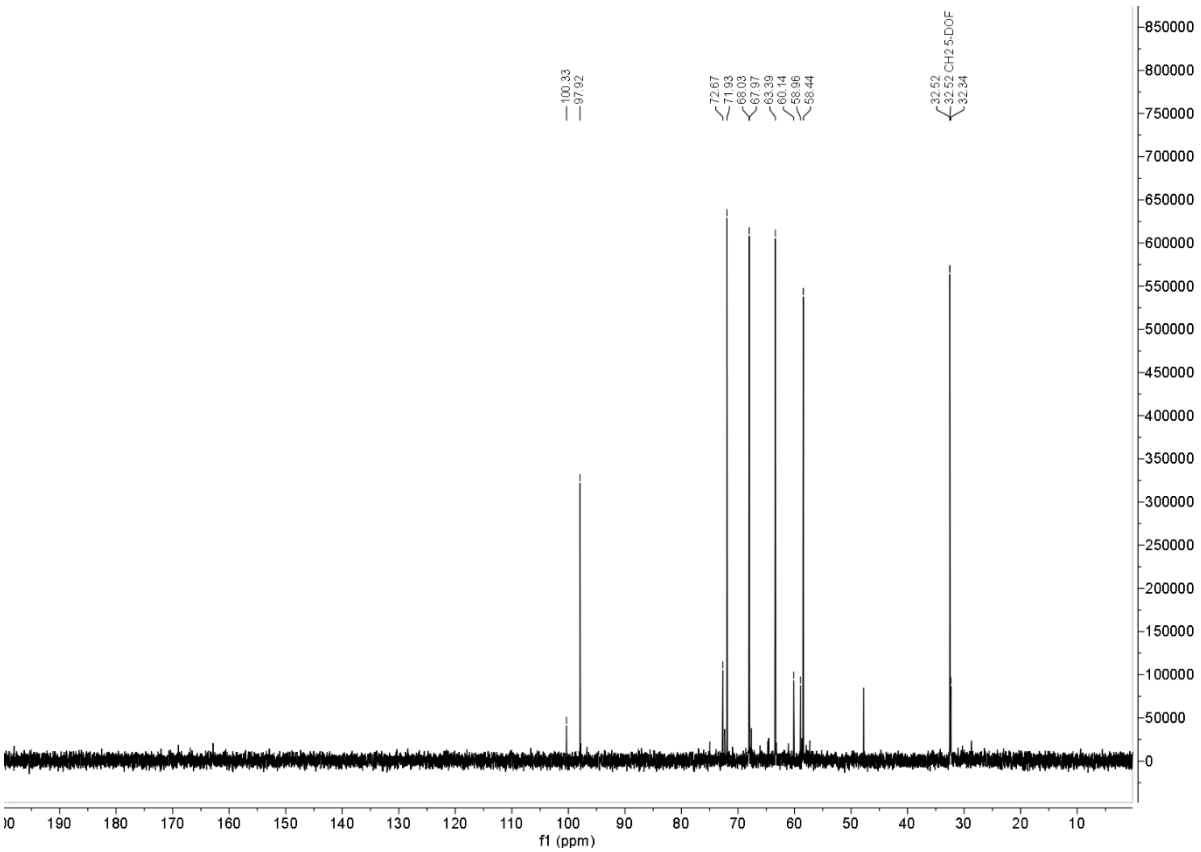
^{1}H NMR (500 MHz, D_2O):



q-^{1}H NMR (500 MHz, D_2O, maleic acid standard):

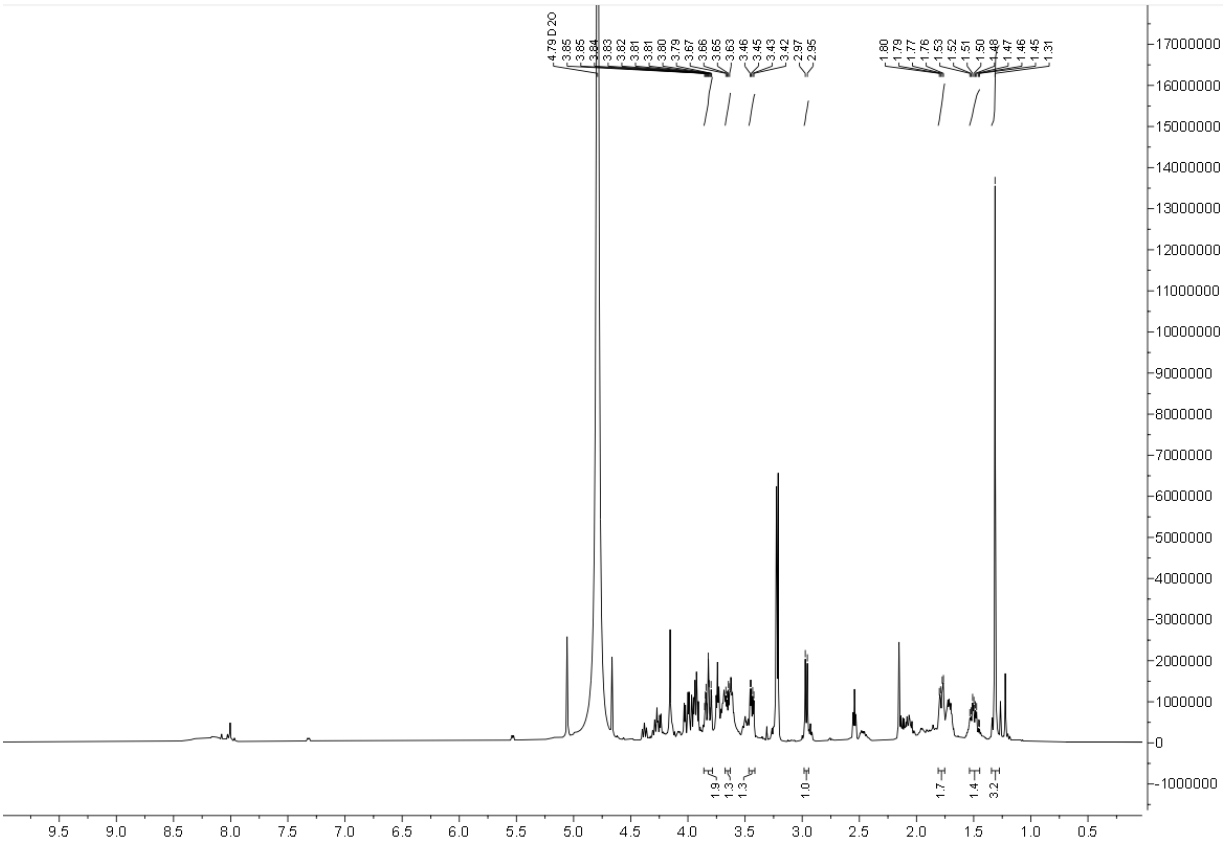


^{13}C NMR (126 MHz, D_2O):

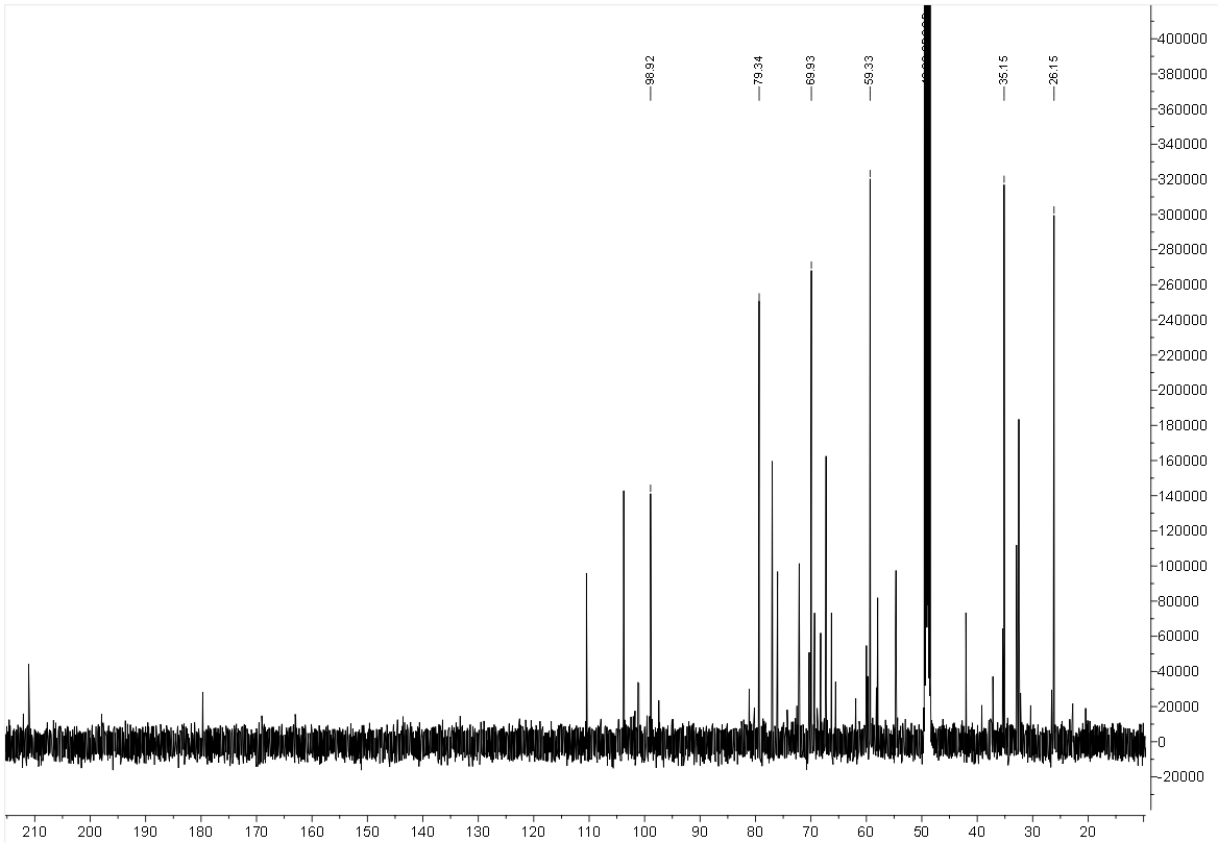


1,5-Dideoxy-D-*threo*-hex-2-ulose (1,5-dideoxyfructose, **31** from ***R*-11**, with DAAO)

^{1}H NMR (500 MHz, D_2O, CD_3OD):

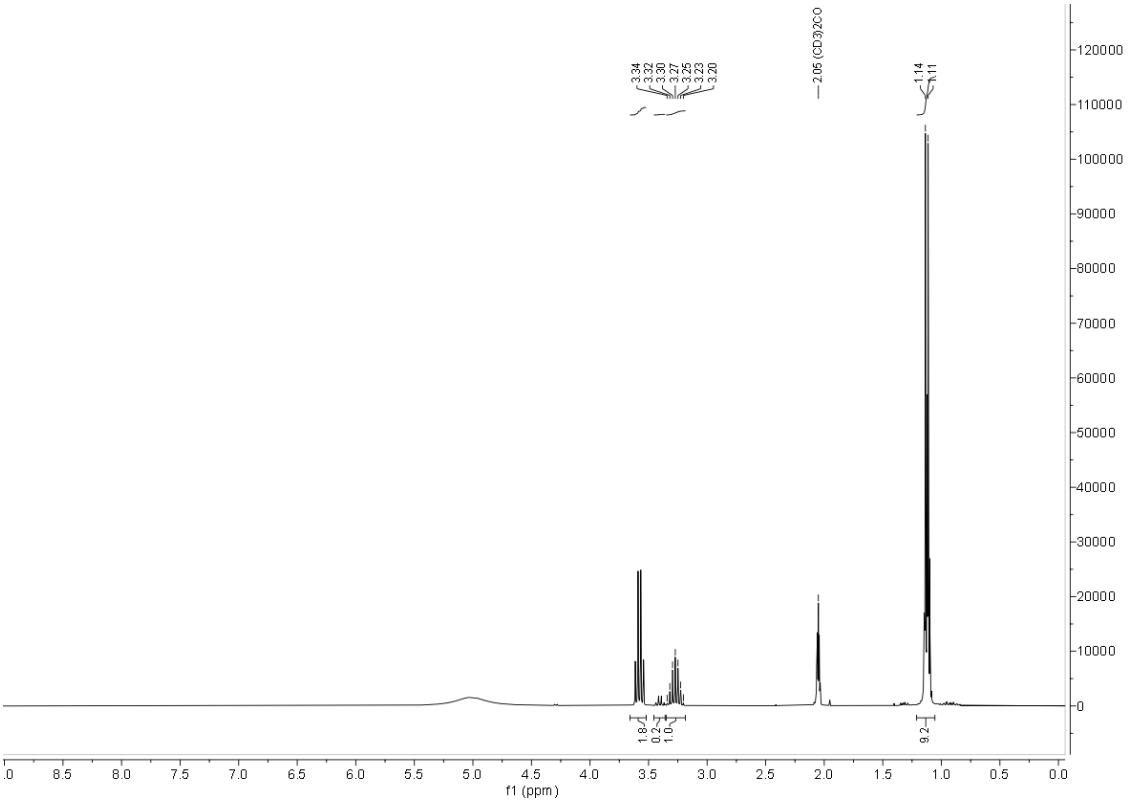


^{13}C NMR (126 MHz, D_2O, CD_3OD):

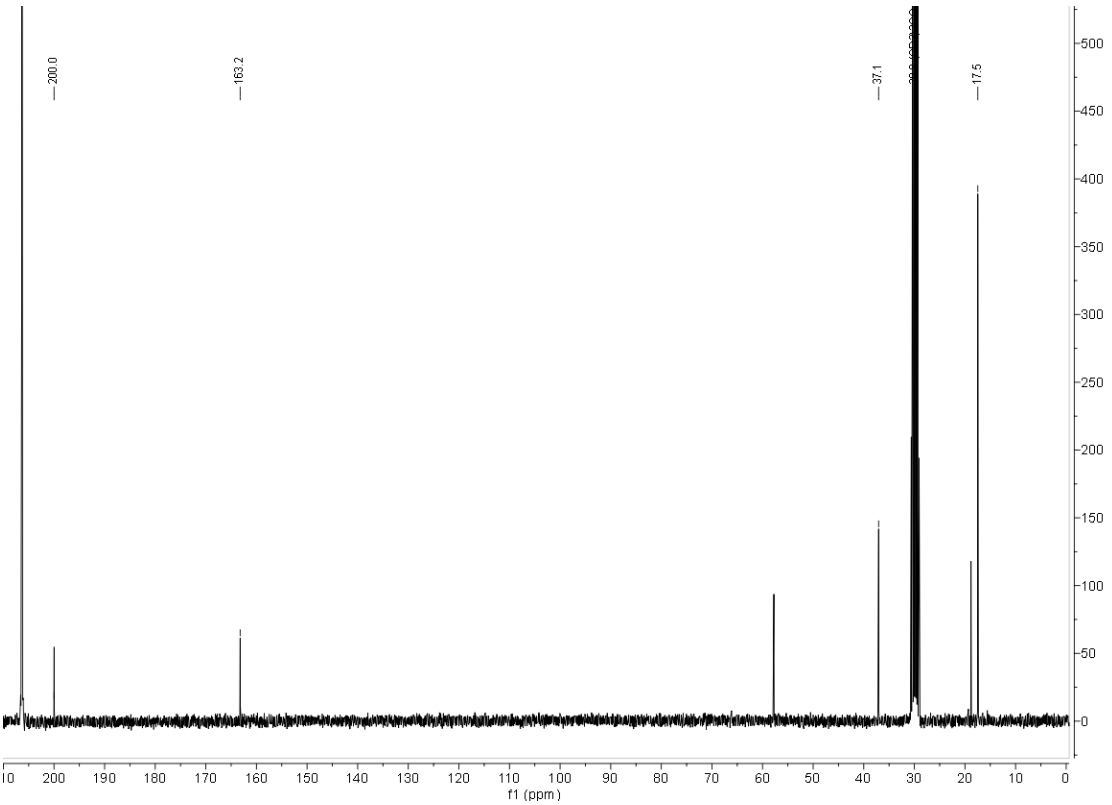


3-Methyl-2-oxobutyric acid (**34**)

^{1}H NMR: (300 MHz, acetone-d^{6})

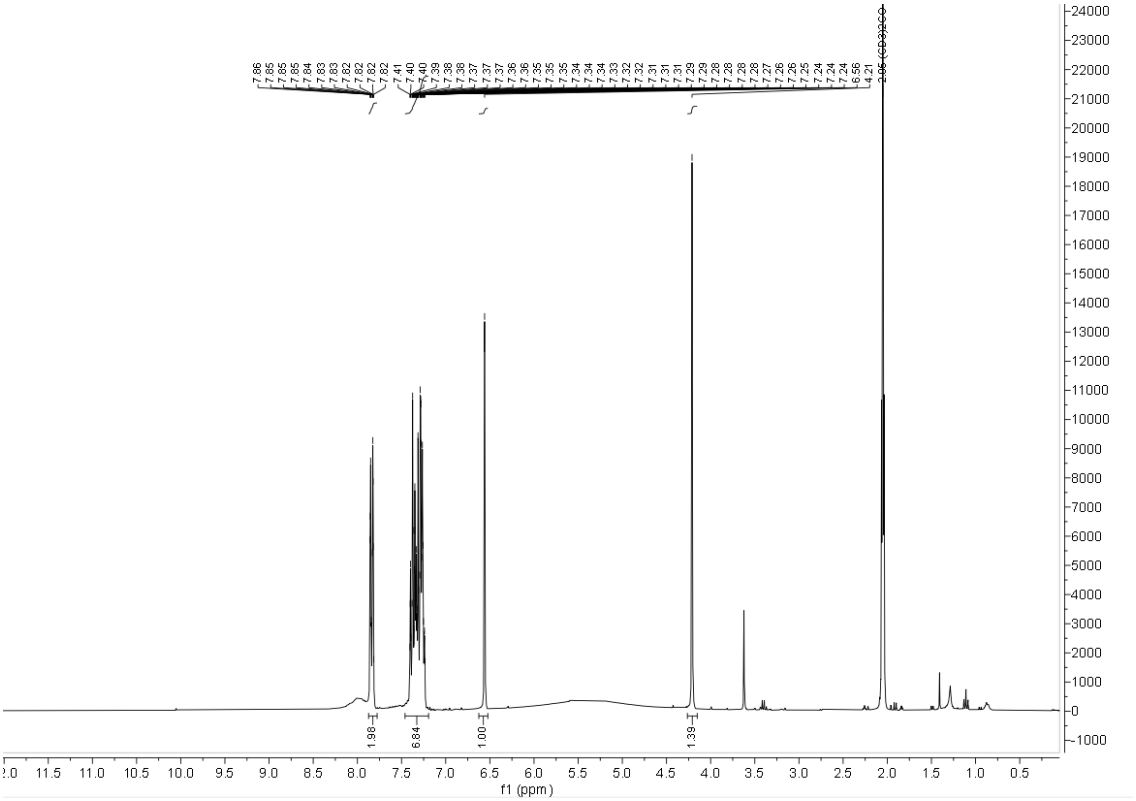


^{13}C NMR: (75 MHz, acetone-d^{6})

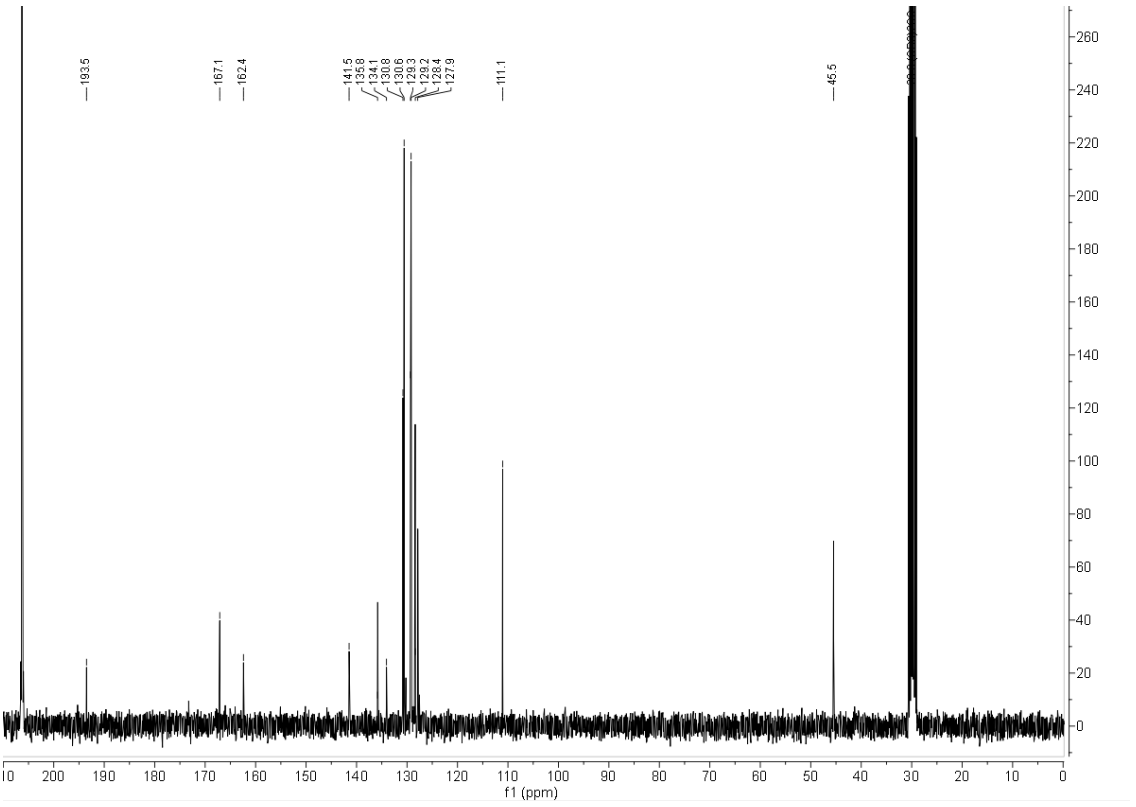


2-Oxo-3-phenylpropionic acid (**33**, keto-enol)

^{1}H NMR: (300 MHz, acetone-d^{6})

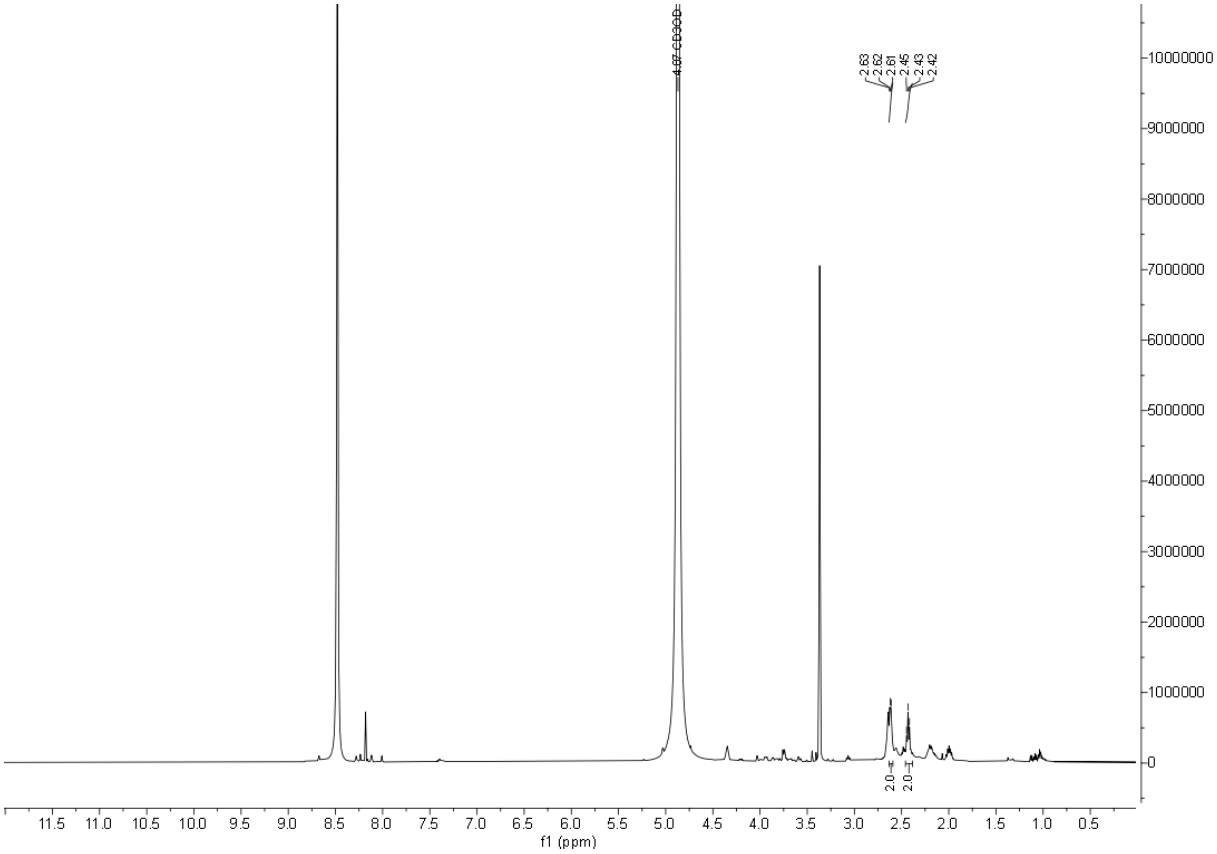


^{13}C NMR: (75 MHz, acetone-d^{6})



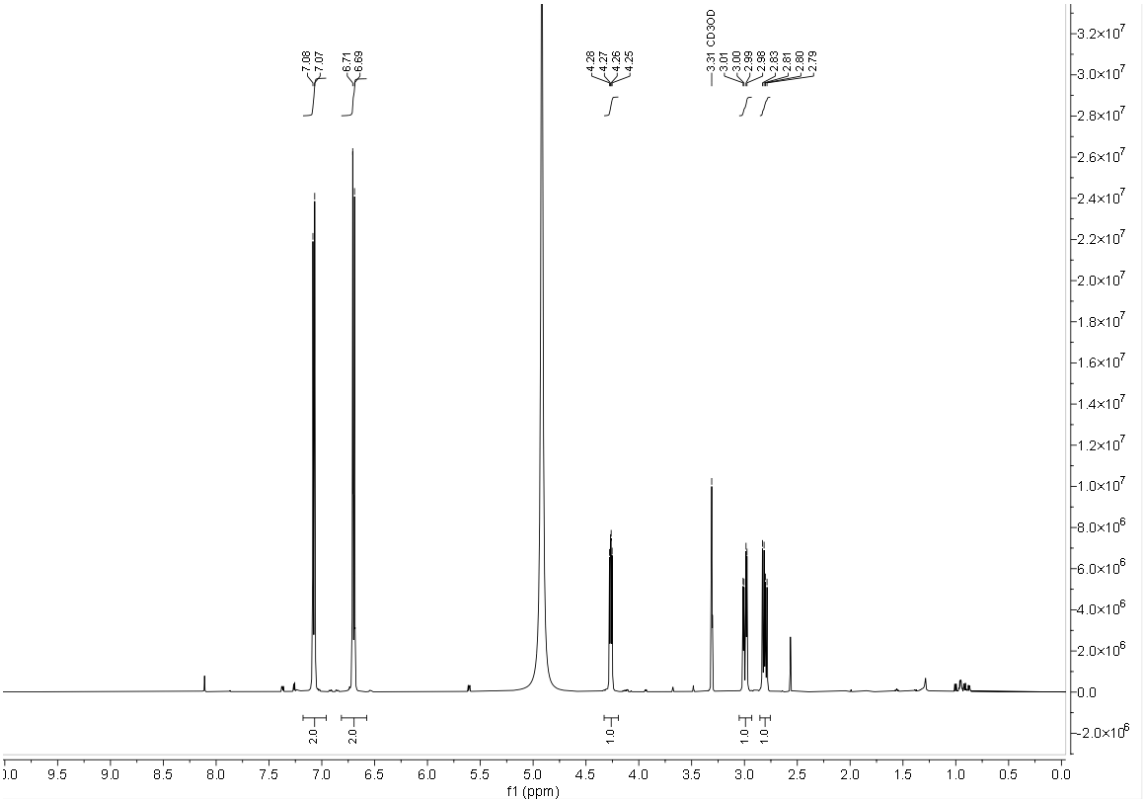
2-Oxoglutaric acid (**35**)

^{1}H NMR: (500 MHz, methanol-d^{4})

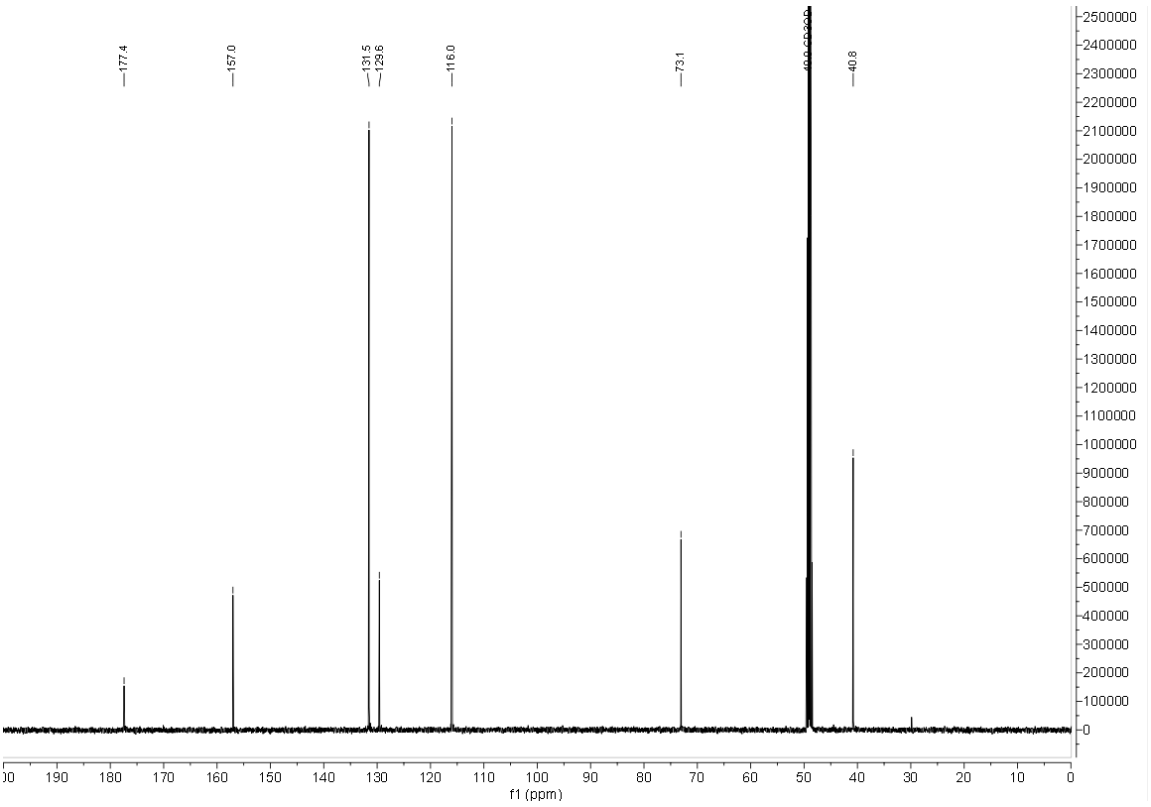


D-2-Hydroxy-3-(4-hydroxyphenyl)propionic acid (**37**)

^{1}H NMR: (500 MHz, methanol-d^{4})

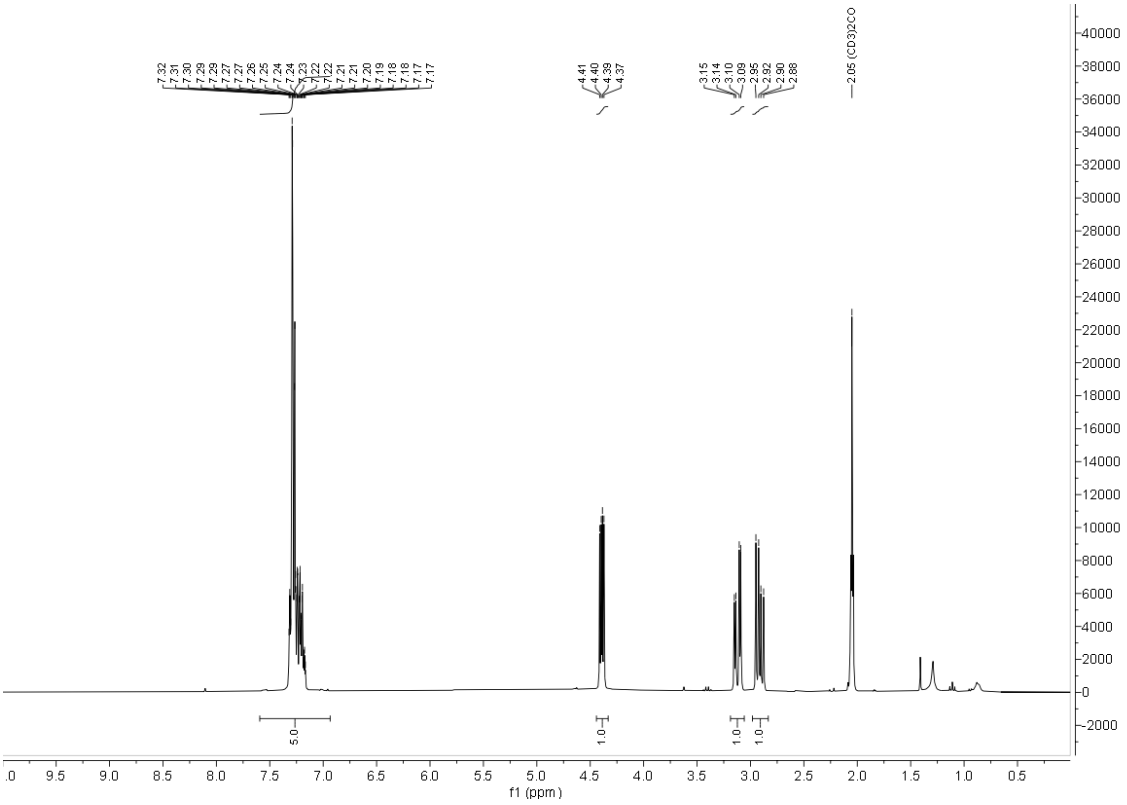


^{13}C NMR: (126 MHz, methanol-d^{4})



D-2-Hydroxy-3-phenylpropionic acid (**36**)

^{1}H NMR: (300 MHz, acetone-d^{6})



^{13}C NMR: (75 MHz, acetone-d^{6})