Novel developments in the chemo-enzymatic synthesis of enantiopure -hydrogen- and -disubstituted -amino acids and derivatives
Sonke, T.

Citation for published version (APA):
Summary and conclusions

Proteins, in addition to nucleic acids, lipids and carbohydrates, are biomolecules that play a crucial role in virtually all biological processes. Functions mediated by these remarkably versatile molecules are found, for instance, in enzymatic catalysis, transport, immune protection, and transmission of nerve impulses. All proteins are built from the same set of twenty amino acids, the so-called proteinogenic amino acids, which are linked by peptidic bonds. The characteristics of a protein, and thus its final biological function, are laid down in the sequence of these amino acids, usually over 100 per protein, which is unique for every protein.

Because of their key role as building blocks of life, proteinogenic amino acids are nutritionally important compounds. Consequently, they are extensively used in the food and feed industry to improve the nutritional value of certain foodstuffs, but also, for example, as flavor enhancer. Amino acids, especially the nonproteinogenic (synthetic) ones, and derivatives thereof, are also increasingly exploited in the cosmetics, agrochemical, pharmaceutical, and biomedical industries, due to their unique chemical and physical properties, including chirality. For most of these new applications, it is essential that the amino acid is enantiomerically pure.

Although most of the proteinogenic amino acids are currently produced by microbial processes, mainly by fermentation [216], processes based on enzymatic catalysis are still preferred for the production of enantiomerically pure nonproteinogenic amino acids [281]. One of such processes is in operation at DSM since the 1980s. It is based on the kinetic resolution of racemic amino acid amides by the action of enantioselective (α-amino) amidases (Scheme 1). This process offers various important advantages, such as biocatalyst versatility (different biocatalysts with complementary broad substrate range and high enantioselectivity are available), product and enantiomer flexibility (both D- and L-stereoisomers of α-H- and α,α-disubstituted amino acids can be prepared), cost-efficient substrate synthesis, and favorable position of the reaction equilibrium. These advantages
Summary and conclusions

make the amidase process an attractive production platform for different types of enantiopure amino acids. At the start of the work described in this thesis, however, various aspects of this process were not optimal. Options for improvement were the still inefficient production of an attractive amidase by fermentation of the wild-type microorganism, and the operation of this amino acid production process as a classical kinetic resolution instead of a dynamic kinetic resolution (DKR), the latter having a theoretical yield of 100% of the desired amino acid enantiomer. The first part of the work described in this thesis was focused on the optimization of DSM’s amidase process to further increase its competitiveness.

Beyond any doubt, *O. anthropi* NCIMB 40321 is the most versatile whole-cell amidase biocatalyst available at DSM. These cells combine high amidase activity toward α-H- and α,α-disubstituted α-amino acid amides, α-hydroxy acid amides and α-N-hydroxy-amino acid amides with strict L-selectivity. The industrial potential of this biocatalyst system is further reinforced by its broad pH optimum, excellent temperature stability, and high salt tolerance. Its application on commercial scale, however, was severely hampered by the fact that *O. anthropi* is a pathogenic micro-organism of risk class 2. Furthermore, formation of a sufficient level of the enzyme in the wild-type strain required a complicated fermentation protocol with an expensive inducer. **Chapter 2** details the way in which these limitations were solved. To enable identification of the gene(s) responsible for this broad amidase activity, an L-specific amidase was purified from *O. anthropi* NCIMB 40321. Characterization studies showed that this L-amidase enantioselectively hydrolyzes the same broad range of substrates as the wild-type *O. anthropi* cells. Peptides, however, while being efficiently hydrolyzed by *O. anthropi*, were not hydrolyzed by the isolated enzyme. The gene encoding this L-amidase was cloned via reverse genetics, and appeared to code for a protein of 314 amino acids with a calculated molecular weight of 33,870. This protein has clear homology to a few other stereoselective amidases and the acetamidase/formamidase family of proteins. Among these are the *Klebsiella oxytoca* (R)-amidase used by Lonza for the resolution of (R,S)-3,3,3-trifluoro-2-hydroxy-2-methylpropionamide, and an L-selective amidase from an *Enterobacter cloacae* strain, which has been developed in the laboratories of Mitsubishi for the production of (S)-tert-leucine. Finally, 17-fold overexpression of the *O. anthropi* L-amidase in *Escherichia coli*
was realized by cloning of the gene in a specialized expression vector with an inducible promoter system (pTrc99A). Decoupling of growth and induction phases, which is possible with this type of vector, appeared to be essential. Thus, we succeeded in establishing an efficient production of this attractive biocatalyst in one of DSM’s enzyme production hosts.

As already indicated above the first generation amidase process is a resolution process with an intrinsic maximum yield of 50% of the desired enantiomer per pass. Because the competitiveness of chemo-enzymatic production methods to enantiomerically pure amino acids depends primarily on the costs of substrate manufacturing, chemical racemization procedures for the recycling of the unreacted amino acid enantiomer have been developed. Unfortunately, the racemization conditions are incompatible with the enzymatic resolution step, and therefore additional hardware is required which adds to the overall process costs. We envisioned that a fully enzymatic dynamic kinetic resolution (DKR) process combining the simultaneous action of an enantioselective amidase and an α-H-α-amino acid amide racemase in one vessel would be a more cost-efficient process (Scheme 2). However, an α-H-α-amino acid amide racemase was still unknown at the start of this work. Chapter 3 describes the identification of this novel enzyme, as well as the subsequent development of the desired DKR process by combining an optimized variant of this amino acid amide racemase with one of DSM’s aminopeptidases, P. putida PepA.

![Scheme 2](image)

Scheme 2. Two enzyme-catalyzed dynamic kinetic resolution process for production of L-α-H-α-amino acids in 100% theoretical yield.

Many different approaches to identify the required α-H-α-amino acid amide racemase failed, and we finally turned to a strategy that was aimed at identification of an enzyme named α-amino-ε-caprolactam racemase (ACL racemase). However, in the literature it was clearly stated that the ACL racemase from Achromobacter obae did not racemize linear amino acid amides. Therefore, we envisaged the ACL racemase as a starting backbone only, and anticipated the need for a subsequent protein engineering program to change it into an enzyme with racemase activity for the desired type of substrates. Thus, micro-organisms were enriched on minimal medium containing DL-α-amino-ε-caprolactam (ACL) as sole nitrogen source, which resulted in the isolation of
Ochrobactrum anthropi strain NCIMB 41129. To our surprise testing of this novel strain showed that it already had a low $\alpha$-H-$\alpha$-amino acid amide racemase (AmaR) activity. Following isolation of the amaR gene from an O. anthropi genomic library and its heterologous expression in E. coli, the racemase was partially purified and characterized. The amaR gene encodes a 439 amino acids long polypeptide with a calculated molecular weight of 46,810 and an aminotransferase class-III pyridoxal-phosphate attachment site. Furthermore, highest homology (52% sequence identity) was observed with the ACL racemase from Achromobacter obae. The racemase combines a good thermostability with a broad pH optimum, which are two prerequisites for large-scale application. Nevertheless, two hurdles had to be taken to establish the desired DKR process. Firstly, the $\alpha$-H-$\alpha$-amino acid amide racemase was inhibited by exactly those divalent metal ions that are essential for the activity of the P. putida L-aminopeptidase. Careful tuning of the metal ion concentration, which showed that 0.6 mM of Mn$^{2+}$ is the optimal compromise, solved this problem. The second problem, a moderate to low activity toward linear $\alpha$-H-$\alpha$-amino acid amides with a longer and/or C$_{\beta}$-branched aliphatic side chain, was tackled by protein engineering. Two rounds of directed evolution were sufficient to obtain a mutant racemase with a 35-fold improved activity toward the long-chain unsaturated $\alpha$-H-$\alpha$-amino acid amide 2-amino-non-8-enoic acid amide (ANEA-NH$_2$). By combining this improved racemase mutant with the L-aminopeptidase from P. putida, it was possible to synthesize (S)-ANECA from the racemic amide in 94% conversion and 98% enantiomeric excess. Work-up of the reaction mixture provided the enantiopure (S)-ANECA (ee >99.9%) in gram quantity. Thuswise the work described is Chapter 3 clearly resulted in proof-of-principle for the desired DKR amidase process.

The second part of this thesis describes the explorative work on breakthrough concepts for the production of enantiomerically pure amino acids and derivatives thereof. The work detailed in Chapter 4 explored the potential of E. coli peptide deformylase (EcPDF) for the enantioselective synthesis of a wide range of such compounds, with focus on $\alpha$-aminonitriles. Because $\alpha$-aminonitriles are the precursors of the $\alpha$-amino acid amides in the Strecker synthesis, application of PDF results in an “earlier resolution” concept with the associated advantages.

In nature, PDF catalyzes the removal of the N-terminal formyl group from nascent polypeptides that are synthesized by the ribosomal machinery. Because PDF activity is essential for the survival of all eubacteria and a functional role in protein (de)formylation has not been found for eukaryotic PDF orthologs up to now, this enzyme is regarded an attractive target for new antibacterial agents. This has led to a huge number of papers on these hydrolytic enzymes. However, at the start of this work nothing had been published about the application of a PDF in organic synthesis, most likely because it took until 1995 before the enzyme could be obtained in stable form.
To enable exploration of EcPDF’s biocatalytic potential we developed a quick and efficient purification method that was based on polyethylenimine mediated removal of nucleic acids followed by a single affinity chromatographic step applying Met-Lys-Sepharose as affinity matrix. With this method about 200 mg of pure EcPDF per liter of E. coli culture was obtained. The purified enzyme appeared to hydrolyze N-formyl derivatives of non-functionalized amines and β-amino alcohols with low to moderate activity and enantioselectivity, whereas the resolution of N-formylated α- and β-amino acids, α-amino acid amides, and α-aminonitriles proceeded with good activity and almost complete enantioselectivity (E ratio >500). Of these substrates, N-formyl-α-aminonitriles were hydrolyzed with the highest activity. Unfortunately, stability problems of the formed α-aminonitriles rendered the PDF catalyzed deformylation reaction suitable for the synthesis of the remaining and more stable (R)-N-formyl-α-aminonitrile only. Synthesis of the corresponding (S)-enantiomer was possible by the reverse reaction, i.e. resolution of the racemic α-aminonitrile via PDF-mediated formylation, when thermodynamic conditions and a high formate concentration (6 M) were applied. Thus both enantiomers of the more stable N-formyl-α-aminonitriles can now be obtained in enantiopure form by either the stereoselective hydrolysis of the racemic N-formylated precursor molecules or the stereoselective formylation of the racemic α-aminonitriles (Scheme 3). These results are of great interest because α-aminonitriles are valuable building blocks for chemical synthesis, since they are readily accessible using Strecker chemistry starting from aldehydes or ketones, and can be converted into a broad range of interesting synthons including amino acids. The fact that α-aminonitriles are not very (optically) stable severely limits the number of suitable enzymatic resolution routes to enantiopure (N-acyl)-α-aminonitriles. Therefore, the PDF catalyzed reactions are a very valuable addition to the limited toolbox of enantioselective α-aminonitrile converting reactions known today.

\[ \text{Scheme 3. PDF catalyzed formation of enantiopure aminonitriles by: (A) deformylation of the unwanted enantiomer, and (B) formylation of the desired enantiomer.} \]
Besides the use of PDF in resolution reactions for the synthesis of enantiopure amines and amino acid derivatives, we also explored its application in the removal of the N-terminal formyl protecting group from peptides. The formyl group is an amino protective group of particular industrial interest because it can easily and cheaply be introduced using formic acid and acetic anhydride. However, a mild way to remove this group after completion of the peptide bond forming steps was still lacking. It turned out that this enzyme is highly suited for the mild and selective deprotection of N-formyl peptides as was shown for \( N \)-formyl-L-Leu-L-Tle-NHCH₃ (Tle = tert-leucine). Application of the purified EcPDF resulted in smooth hydrolysis of the formyl group without any concomitant peptide bond hydrolysis, a result not obtainable with acid mediated chemical hydrolysis. Furthermore, the diastereomeric excess of the dipeptide, which was unsatisfactory because of racemization of the N-terminal amino acid in the chemical peptide coupling step, significantly increased by this enzymatic deprotection step. Use of a purified EcPDF preparation turned out to be essential for this type of application, because with a cell free extract from the PDF expressing \( E. \ coli \) strain extensive simultaneous peptide bond hydrolysis was observed.

\( \alpha \)-H-\( \alpha \)-Amino acids can be manufactured on commercial scale by four basic methods: isolation from protein hydrolysates (frequently called extraction process), chemical synthesis, chemo-enzymatic synthesis, and fermentation. Although all four methods have been in operation in the past, most proteinogenic amino acids are nowadays industrially manufactured by fermentation. In these processes, overproducing microorganisms make the desired L-amino acid in high yield and absolute enantioselectivity directly from cheap carbohydrates (mainly glucose) and ammonia, which makes fermentation in many cases the most cost-efficient manufacturing method. Mainly because of this economic driver, bulk amino acids, such as monosodium L-glutamate (MSG), L-lysine and L-threonine, are manufactured by fermentation for decades now. Thanks to the recent astonishing development of the methodologies scientists can choose from to force microorganisms to overproduce metabolites, fermentation is increasingly used for the smaller scale production of proteinogenic amino acids such as L-phenylalanine, L-arginine, L-valine, L-isoleucine, and L-tryptophan \(^{122,216}\).

Triggered by the lowest-cost potential offered by fermentation, we investigated whether such process is also feasible for the production of certain nonproteinogenic amino acids. D-Phenylglycine (D-Phg) was chosen as model system for the work described in Chapter 5, since it is an important building block of semi-synthetic antibiotics like ampicillin and cephalexin.

Since a direct biosynthetic pathway to free D-Phg is not known in nature, an artificial one was constructed. This three-step hybrid pathway, which starts with the L-phenylalanine precursor phenylpyruvate, consists of the first two steps of the L-HPG
biosynthetic pathway found in many actinomycetes, combined with a D-4-hydroxyphenylglycine aminotransferase (Scheme 4). The genes encoding these first two enzymes, hydroxymandelate synthase (HmaS\textsubscript{Ao}) and hydroxymandelate oxidase (Hmo\textsubscript{Sc}), were cloned by PCR from Amycolatopsis orientalis and Streptomyces coelicolor, respectively. Both genes could be expressed in E. coli, although the enzymatic activities obtained toward phenylpyruvate (HmaS\textsubscript{Ao}) and L-mandelate (Hmo\textsubscript{Sc}) were low. The third enzyme needed, a D-(4-hydroxy)phenylglycine aminotransferase (HpgAT), had been demonstrated before in different Pseudomonas strains, but the gene encoding this enzymatic activity was not known. We isolated the HpgAT gene by screening a P. putida LW-4 genomic library in E. coli using a spectrophotometric assay. By testing both D- and L-glutamate as amino donor, the unique stereoinverting property of this aminotransferase was confirmed: while the enzyme is inactive with D-glutamate, it efficiently converts phenylglyoxylate to D-Phg (ee >99%) in the presence of L-glutamate.

![Scheme 4. Artificial D-Phg biosynthesis pathway.](image)

HmaS: Hydroxymandelate synthase; Hmo: Hydroxymandelate oxidase; HpgAT: D-(4-Hydroxy)phenylglycine aminotransferase.

To realize combined expression of the three heterologous enzymatic activities in E. coli, the hmaS\textsubscript{Ao}, hmo\textsubscript{Sc} and hpgAT genes were cloned into an IPTG inducible E. coli expression vector (pJF119EH), resulting in a D-Phg synthesis operon which is transcribed into one polycistronic mRNA. Functional expression of this artificial operon in E. coli DH5\alpha was then conclusively demonstrated by in vitro incubation of phenylpyruvate with the cell free extract of this recombinant strain, which resulted in the formation of D-Phg. To produce D-Phg ultimately from glucose, it is essential that an E. coli production strain is used in which the amount of phenylpyruvate is not limiting. Since phenylpyruvate is the direct precursor of L-phenylalanine, the artificial D-Phg operon, enlarged with a copy of the pheA\textsuperscript{fbr} gene encoding a feedback deregulated chorismate mutase/prephenate dehydratase mutant ensuring sufficient flow from chorismate to phenylpyruvate (Scheme 4), was expressed in L-phenylalanine overproducing E. coli strain KB532. Cultivation of this strain in a mineral salts medium containing glucose as carbon source, L-tyrosine as compensation for this strain’s auxotrophy, and IPTG as inducer led to in vivo formation of
clearly detectable amounts of D-Phg. Further metabolic engineering of this D-Phg production strain to eliminate competing pathways increased the amount of D-Phg by 100% to 102 mg/g biomass and reduced the formation of the main by-products L-phenylalanine (by 35%) and 4-hydroxyphenylglyoxylate (by 85%).

This work shows that it is indeed possible to synthesize D-Phg (and other nonproteinogenic amino acids) from renewable feedstocks by a completely fermentative route. However, it also revealed a number of difficulties associated with the development of this type of fully fermentative routes. Clearly, the formation of by-products formed by the action of strain endogenous enzymes is one of these difficulties. Although we have demonstrated that the formation of these side products can be significantly reduced by elimination of the interfering endogenous activities, this approach will frequently result in less fitter strains with more auxotrophies, and, consequently, higher raw materials costs. A more delicate tuning of the heterologous and endogenous enzyme activities will certainly enable a more elegant solution of this problem. But also other aspects, like transport of the product out of the cell, product toxicity, and product recovery need to be addressed. The recent tremendous progress in tools like genomics, proteomics, genetic and protein engineering, and high throughput screening, as well as the fundamental information obtained, will finally lead to a more rationalized, and, consequently, a more straightforward, more reliable, and thus quicker development of fermentation processes [216]. However, we agree with the opinion of Leuchtenberger et al. that, at least for the next five to ten years, chemo-enzymatic processes will remain the preferred production method for nonproteinogenic amino acids [281].

Rapid identification of enzymes suitable for commercial scale applications is a frequently recurring theme throughout the work described in this thesis. Because of the pivotal importance of the development of enzymatic assays and activity-based screenings, Chapter 6 is devoted to this with special focus on Chemical Custom Manufacturing (CCM). This is the industry segment in which the fine chemical industry manufactures compounds on demand for a customer. CCM is one of the outlets for the enantiopure amino acids and derivatives, which are the target compounds of the synthetic methods detailed in this thesis. Because of the short timelines in CCM the time available for biocatalyst identification and (small-scale) process development is very limited, typically a couple of weeks to a few months at most. But also for other application areas, like metabolic pathway engineering or development of 2nd generation production processes, time for identification of new or improved biocatalysts is usually limited.

In chapter 6, two different screening approaches are described that meet the very demanding timelines in CCM. The first approach is based on the use of instrumental assay techniques such as NMR and MS. These types of techniques are very broadly applicable (“generic”) because they generate distinct signals for many chemical functionalities
enabling straightforward differentiation between substrate and product of an enzymatic reaction. Therefore, development of a screening method based on such instrumental technique usually requires a few days only. In combination with a sufficient screening throughput, these methods generate the first screening results in a short period of time, leading to rapid hit identification and thus prompt start of application work and sample preparation.

Rapid and responsive biocatalyst identification is also possible with much less generic assay methods, but this requires a lot of proactive development work with the accompanying expenses. Therefore, this second approach is only feasible for a number of frequently encountered chemical group conversions in, for example, the synthesis of amino acids or in reactions giving ammonia as co-product. This situation is typical for plate-reader assays, which frequently require lengthy set-up of a chemical derivatization or enzymatic cascade reaction because the actual screening reaction is not accompanied by a clear change in optical signal.

Both screening approaches have been successfully applied in the screening projects described in this thesis. We are therefore convinced that the use of these two approaches in combination with high-quality biodiversity as input will result in the continued delivery of useful biocatalysts, thereby leading to further implementation of biocatalysis in the chemical industry as cheap and environmentally benign technology.