

Silvia Maitz, BSc

Engineering of hydroxynitrile lyases for nitroaldolase activity

MASTER'S THESIS

to achieve the university degree of

Diplom-Ingenieurin

Master's degree programme: Biotechnology

submitted to

Graz University of Technology

Supervisor

Schwab, Helmut, Univ.-Prof. Dipl.-Ing. Dr.techn.

Institute of Molecular Biotechnology

Steiner, Kerstin, Dipl.-Ing. Dr.nat.techn.

AFFIDAVIT

I declare that I have authored this thesis independently, that I have not used other than the declared sources/resources, and that I have explicitly indicated all material which has been quoted either literally or by content from the sources used. The text document uploaded to TUGRAZonline is identical to the present master's thesis.

Date

Signature

0. Acknowledgements

First and foremost I want to thank Professor Helmut Schwab for giving me the opportunity to carry out my diploma thesis at the austrian centre of industrial biotechnology and thus giving me the necessary space, equipment and above all the best professional support I could ask for.

My special thanks go to Dr. Kerstin Steiner who managed to guide me through the necessary working steps and support me with her expertise without restricting my desire to work autonomously and independently. Without her support and patience finishing this work would have been impossible.

I also want to thank the acib staff and especially Julia, Karin, Myria, Sabine, Sarah and Svitlana for their valuable advice and company during the longer working hours.

Additionally I want to thank Myria and Gernot for helping me with the HPLC measurements necessary for the final characterization of my variants at the institute of organic chemistry.

Of course I also want to thank my mother and my siblings for their unwavering support and encouragement. I also want to thank my friends Miriam, Sara and Sarah for helping me to occasionally divert my thoughts towards the more important parts of life.

However, most of all I want to thank my daughter Aurea, for putting up with a mother who spent half a year either physically absent or absent-minded. Thank you for your continuous cheerfulness and patience. If it was not for you, I would never have come this far!

1. Abstract

Hydroxynitrile lyases (HNLs) are commonly known to catalyze the formation and cleavage of cyanohydrins and are widely applied in industry and research. Not too long ago it was shown that *HbHNL* from *Hevea brasiliensis* is also able to catalyze the so called Henry reaction (nitroaldol reaction), where an aldehyde reacts with a nitroalkane yielding a β -nitro alcohol. Enantiopure β -nitro alcohols are versatile precursors in the pharmaceutical industry and hence they constitute interesting targets for production by biocatalysis. Since *HbHNL*'s nitroaldolase activity was verified a decade ago, several approaches have been made to also apply other enzymes in a biocatalytic, asymmetric nitroaldol reaction, but an enzyme promoting high yields combined with high enantioselectivity has yet to be found. Recently two bacterial cupin-fold HNLs, *GtHNL* from *Granulicella tundricola* and *AcHNL* from *Acidobacterium capsulatum* have been found of which certain variants are able to catalyze the nitroaldol reaction with nitroethane and benzaldehyde as substrates, yielding 2-nitro-1-phenyl-1-propanol. The two most promising variants, *GtHNL*-A40R and *AcHNL*-A40H, were subjected to enzyme engineering by directed evolution aimed to find variants that showed higher conversion and selectivity towards the (1*R*,2*S*) 2-nitro-1-phenyl-1-propanol product compared to the parent variants. However, although the screening of random mutagenesis libraries of both variants did not yield a single variant with higher enantioselectivity, all the identified hits achieved higher benzaldehyde conversions than the parent variants. The best variant, *AcHNL*-A40H/T50S, converted nearly six times the amount of benzaldehyde during the reference time compared to the parent variant.

2. Kurzzusammenfassung

Hydroxynitril-Lyasen (HNLs) sind im Allgemeinen dafür bekannt, die Bildungs- und Spaltungsreaktion von Cyanohydrinen zu katalysieren und finden sowohl in der Industrie als auch in der Forschung breite Anwendung. Kürzlich wurde auch gezeigt, dass die *HbHNL* von *Hevea Brasiliensis* zusätzlich die sogenannte Henry Reaktion (Nitroaldol Reaktion) katalysieren kann, bei der ein Aldehyd und ein Nitroalkan zu einem β -Nitroalkohol reagieren. Enantiomerenreine β -Nitroalkohole sind vielfältige Zwischenprodukte in der pharmazeutischen Industrie, was deren Produktion mittels Biokatalyse besonders interessant macht. Seit die Nitroaldolase Aktivität der *HbHNL* vor rund zehn Jahren nachgewiesen wurde, gab es mehrere Ansätze auch andere Enzyme in einer biokatalytischen, asymmetrischen Henry-Reaktion einzusetzen, aber ein Enzym das hohe Umsätze in Kombination mit einer hohen Enantioselektivität bietet, wurde bis dato noch nicht gefunden. Allerdings wurden erst kürzlich zwei bakterielle cupin HNLs, *GtHNL* von *Granulicella tundricola* und *AcHNL* von *Acidobacterium Capsulatum*, identifiziert von denen spezielle Varianten die Nitroaldol Reaktion mit Nitroethan und Benzaldehyd als Substrate katalysieren können und somit das Produkt 2-Nitro-1-Phenyl-1-Propanol bilden. Die beiden vielversprechendsten Varianten, *GtHNL-A40R* und *AcHNL-A40H*, wurden einem Enzym-Engineering durch gerichtete Evolution unterzogen um Varianten zu finden die einen höheren Umsatz und eine höhere Selektivität für das (1*R*,2*S*) 2-Nitro-1-Phenyl-1-Propanol Produkt im Vergleich zu den Ausgangsvarianten zeigen. Obwohl das Screening von, durch ungerichtete Mutagenese erzeugten, Genbibliotheken beider Varianten keine einzige Variante mit höherer Enantioselektivität lieferte, konnten alle identifizierten Hits höhere Umsätze erzielen als die jeweiligen Ausgangsvarianten. Die beste Variante, *AcHNL-A40H/T50S*, konnte während des gleichen Zeitraums beinahe sechs Mal so viel Benzaldehyd umsetzen wie die Ausgangsvariante.

Table of Contents

0. Acknowledgements	ii
1. Abstract	iii
2. Kurzzusammenfassung	iv
Table of Contents	v
3. Introduction.....	1
3.1. Characteristics and application of hydroxynitrile lyases	1
3.1.1. Cupin-fold hydroxynitrile lyases from <i>Granulicella tundricola</i> and <i>Acidobacterium capsulatum</i>	2
3.2. The nitroaldol reaction.....	3
3.2.1. Catalysts for the asymmetric nitroaldol reaction.....	4
3.3. Enzyme engineering	6
3.3.1. Directed evolution.....	7
3.3.2. Selection and Screening systems	7
4. Materials and Methods	9
4.1. Materials.....	9
4.2. Preparation of random mutagenesis libraries	9
4.2.1. Isolation of plasmid templates	10
4.2.2. Error prone PCR.....	11
4.2.3. Gel purification using the Wizard SV Gel and PCR Clean-Up System.....	12
4.2.4. QuikChange PCR	12
4.2.5. DpnI digest.....	13
4.2.6. Transformation of <i>E. coli</i> Top10F'	14
4.2.7. Generation of mixed plasmid preparation.....	14
4.2.8. Transformation of <i>E. coli</i> BL21 (DE3) Gold	15
4.2.9. Colony picking.....	15
4.3. Sequencing	16
4.4. Screening	17
4.4.1. Deep well plate fermentation	17
4.4.2. The standard screening assay.....	18
4.4.3. Standard screening assay with cell free lysate.....	23
4.4.4. Time dependent screening assay	23
4.5. Site specific mutagenesis.....	24
4.6. Site saturation mutagenesis	26
4.7. Preparation of cell free lysate	27
4.7.1. Shake flask fermentation.....	27

4.7.2.	Cell harvest and disruption.....	28
4.7.3.	Bradford protein assay	28
4.7.4.	SDS-PAGE analysis	28
4.8.	Final characterization of hits	29
4.8.1.	Biphasic reaction system	29
4.8.2.	HPLC analysis	30
5.	Results and Discussion	32
5.1.	Characterization of parent variants in aqueous system.....	32
5.2.	Library preparation.....	33
5.3.	AcHNL-A40H library screening	36
5.4.	AcHNL-A40H library rescreening	40
5.5.	Preparation and cultivation of AcHNL-T50A and control variants	46
5.6.	Site saturation mutagenesis: AcHNL-A40H/T50X.....	48
5.7.	Cultivation of AcHNL-A40H mutants	51
5.8.	GtHNL-A40R library screening.....	53
5.9.	GtHNL-A40R library rescreening	55
5.10.	Cultivation of GtHNL-A40R variants	60
5.11.	AcHNL-A40H variants: conversion analysis in aqueous system	61
5.12.	GtHNL-A40R variants: conversion analysis in aqueous system.....	62
5.13.	Final characterization in biphasic system and HPLC analysis	63
6.	Conclusion	72
7.	List of Literature	73
8.	Appendix.....	76
9.	Table of Figures	84
10.	List of Tables	88

3. Introduction

3.1. Characteristics and application of hydroxynitrile lyases

Hydroxynitrile lyases, also called oxynitrilases (E.C. 4.1.2.X), are known for their ability to catalyze the reversible cleavage of cyanohydrins, yielding the corresponding carbonyl compound and hydrogen cyanide (HCN). The reverse reaction is the asymmetric addition of HCN to an aldehyde or ketone group, yielding a potentially chiral cyanohydrin. This reaction thus introduces a new C-C bond in addition to the nitrile functional group, which renders HNLs very interesting catalysts for numerous applications in pharmacy, agriculture and cosmetic industry [1-4]. More specifically, enantiopure cyanohydrins act as key intermediates for compounds like α -hydroxy acids, primary and secondary β -hydroxy amines, α -hydroxy esters, α -aminonitriles and several other chemicals [3,5].

The first research on HNLs had already been carried out in the middle of the nineteenth century, when German scientists prepared an HNL-containing extract of bitter almonds and called it "emulsin". This extract propagated the cleavage of amygdalin, yielding a sugar component, benzaldehyde and HCN [6]. In 1908 the first synthesis reaction had been described by Rosenthaler, who used a bitter almond extract to synthesize (*R*)-mandelonitrile [7]. Other than almonds, over 3000 different plants are known to have the ability to produce HCN, though many of them do not use HNLs for cyanogenesis. The HCN produced by these plants does not only function as a nitrogen source, but also acts as part of a defense mechanism against herbivores and microorganisms as it is released upon damage of plant tissue [3,8-10].

Some HNLs, like (*S*)-*Hb*HNL from rubber tree and (*R*)-*Pa*HNL from almonds are currently used on preparative scale [3]. Even though cyanohydrins can also be synthesized with certain lipases or non-enzymatic catalysts like chiral transition metal complexes or cyclic dipeptides [11,12], there are several benefits that arise from the use of HNLs for the synthesis of cyanohydrins. Not only are recombinant HNLs available in reproducible quality and large amounts, they also comprise a high stereoselectivity and a large substrate spectrum while being able to use inexpensive sources of HCN [11-13].

Most of the HNLs that have been characterized so far are originating from plants, but more recently also HNLs with bacterial origin have been identified [2,14-16]. These enzymes not only have many different sources, but also very diverse sequences and structures, which indicates that they emerged from convergent evolution. The structural similarities of HNLs range from folds similar to, for example, α/β -hydrolases, alcohol dehydrogenases or GMC-oxidoreductases, while some of them are also dependent on cofactors like FAD, NAD or metal ions [3,15,16].

3.1.1. Cupin-fold hydroxynitrile lyases from *Granulicella tundricola* and *Acidobacterium capsulatum*

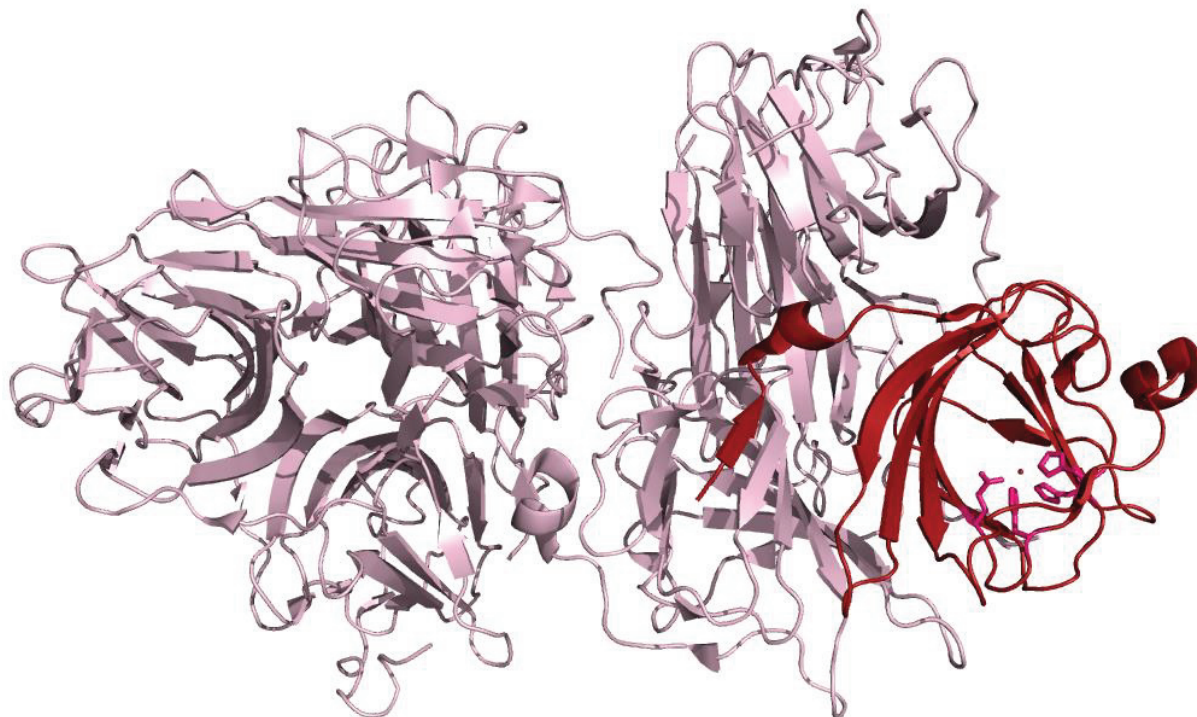


Figure 1. Graphical representation of the 3-dimensional structure of two homotetramers of *GtHNL*. One monomer is depicted in red, the four corresponding metal binding amino acids and the manganese ion are highlighted in purple. The Figure was prepared using the PyMOL Molecular Graphics System and is based on the PDB entry 4bif.

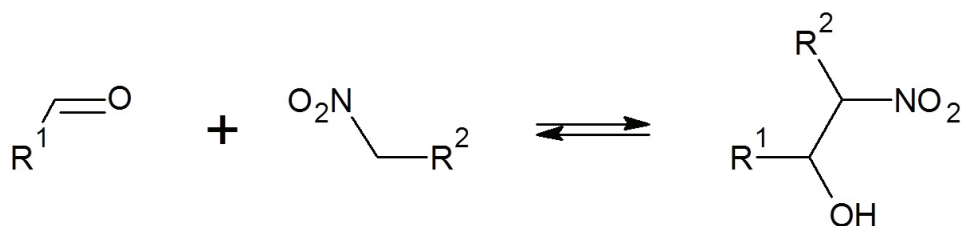
The bacterial HNL from *Granulicella tundricola*, *GtHNL*, is dependent on a Mn^{2+} ion and was the first HNL where a cupin-fold (from the Latin word ‘cupa’, small barrel) was verified by X-ray crystallography [16]. Enzymes with a cupin fold comprise a beta barrel structure and are generally small and rather stable, they also occur very frequently in nature where they fulfill many different functions [17]. Similar to *GtHNL*, most cupins comprise metal ions that are bound in the active site [18].

A graphical representation of the *GtHNL*’s structure is depicted in Figure 1. The active form of *GtHNL* is present as a homotetramer and each monomer includes eleven β -strands and comprises four metal-binding amino acids, the H53, H55, Q59 and H94 [16]. Another HNL originating from a bacterium, *AchNL* from *Acidobacterium capsulatum*, has been identified by database mining for protein sequence similarity to *GtHNL* by Wiedner and coworkers. Similar to *GtHNL*, the *AchNL*’s activity depends on Mn^{2+} and both enzymes share 84 % sequence similarity on amino acid level. Thus, it can be assumed that also *AchNL* comprises a cupin fold as *GtHNL* does [15]. These two cupins have a length of 131 amino acids, show a high overall stability and can be expressed soluble and active in exceptionally high amounts in *E. coli*, rendering both of them very interesting for further research and eventual industrial application.

Furthermore, it has been shown that *AchNL* wild type as well as certain variants of both enzymes are capable of catalyzing the so-called nitroaldol- or Henry reaction [19].

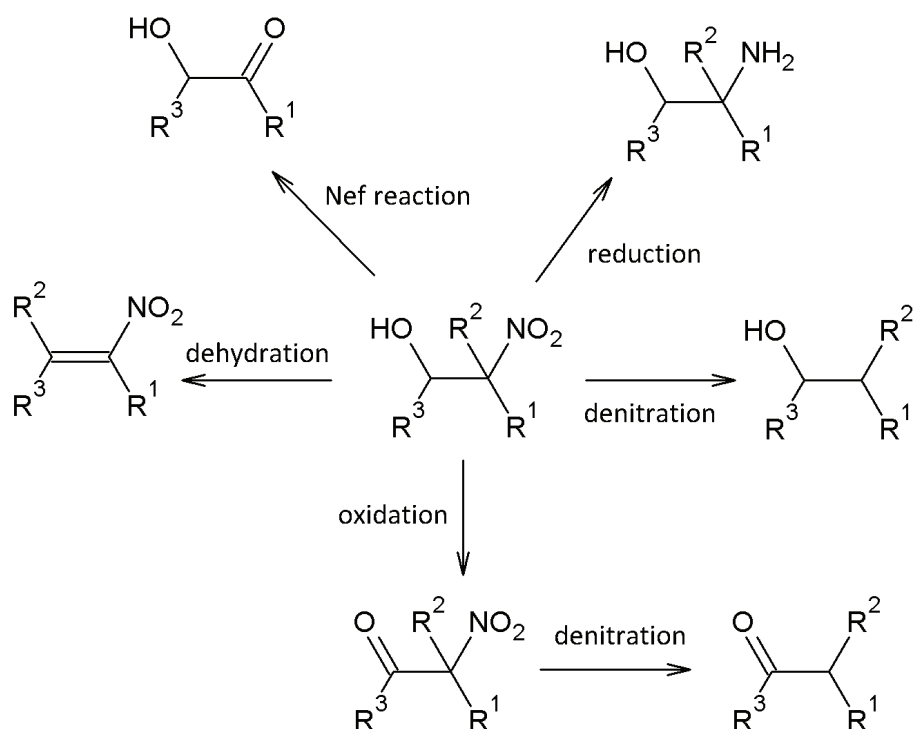
3.2. The nitroaldol reaction

In the nitroaldol reaction, a ketone or aldehyde compound reacts with a nitroalkane that acts as a nucleophile, and thus forms a β -nitro alcohol compound [20] as is shown in Scheme 1. Depending on the substrates, this reaction yields up to two chiral centers in the product, thus rendering it an interesting target for enzyme catalysis which is known to propagate high enantioselectivity in many applications.



Scheme 1. Reaction scheme of the nitroaldol reaction.

Using the resulting β -nitro alcohol as a precursor, it is possible to produce many valuable intermediates and building blocks, as is shown in Scheme 2. Subsequently, these intermediates can be employed in the synthesis of biologically active compounds like insecticides, antibiotics and fungicides, and pharmaceuticals like sphingosine and ephedrine. [20]



Scheme 2. Graphical representation of the follow-up reactions and corresponding products that can be formed based on β -nitro alcohols [21].

3.2.1. Catalysts for the asymmetric nitroaldol reaction

The first enantioselective nitroaldol reaction was described in 1995 by Shibasaki *et al.* who applied heterobimetallic catalysts to propagate this reaction [22]. Since then, several other potential catalysts have been described, including transition-metal catalysts, organocatalysts [23] and biocatalysts. For example R. Arunachalam *et al.* have recently described the use of binuclear Cu(II) helicates as diastereoselective catalysts for the nitroaldol reaction with 2-pyridinecarboxaldehyde and 1-nitropropane as substrates [24]. Another recent example is the catalyst developed by Mei and coworkers, who showed that another copper complex, N'-N-dioxide/Cu(I), also promotes the asymmetric nitroaldol reaction with high yield and enantiomeric excess upon use of nitroethane and benzaldehyde as substrates [25].

However, also a variety of biocatalysts has been described to catalyze the asymmetric nitroaldol reaction. In general, there are two different approaches to the acquisition of enantiopure nitro alcohols using biocatalysts, on the one hand the enzymatic kinetic resolution of chemically synthesized β -nitro alcohols, on the other hand the direct enzyme catalyzed nitroaldol reaction.

Numerous different hydrolases originating from a variety of organisms have been tested for kinetic resolution of aromatic as well as aliphatic nitro alcohols, with partially good enantioselectivities and yields. However, a maximum overall yield of 50 % and a comparably small substrate range are still limiting factors of the kinetic resolution method [21].

The enzymes tested for the direct enzymatic formation of nitro alcohols included lipases, acylases, a transglutaminase, and a glucoamylase. However, most of the tested enzymes showed either only negligible enantioselectivity, or no optical purity data was reported at all. So far, the most promising enzymes to be applied in the direct enzyme-mediated nitro aldol reaction seem to be the HNLs.

Already in 2006, Purkarthofer and coworkers [26] found the first example of a biocatalytic nitroaldol reaction. They examined that *HbHNL* from *Hevea brasiliensis* can catalyze the addition of nitromethane to several aliphatic and aromatic aldehydes with decent yield and enantiomeric excess. Subsequently, the same group increased the scope of tested substrates to a variety of different aldehyde electrophiles and four different nitro alkanes as nucleophiles [27]. In general, the enzymatic conversion was carried out over 48 h in a biphasic reaction system with an aqueous and an organic phase at pH 7 or pH 5.5. As an organic phase, *tert*-butyl-methyl ether (MTBE) was added to the aqueous enzyme solution in ratios of around 50-70 %. The use of an organic solvent was found to be beneficial with respect to reaction rate, yield and *ee*, possibly caused by the related increase of substrate solubility. Additionally, the organic phase suppressed the oxidation of benzaldehyde to benzoic acid and the formation of benzyl alcohol, which was a problem in systems with a higher ratio of aqueous phase. The unspecific background reaction at higher pH values, which is a problem in the cyanohydrin formation reaction, was negligible small in the biphasic system even at a pH of 7 [26]. The obtained yield of nitro

alcohol was 63 % upon use of benzaldehyde and nitromethane as substrates, and the *ee* was 92 % at pH 7, while a pH of 5.5 decreased the yield to 32 % and increased the *ee* to 97 %. *HbHNL* was found to be (*S*)-selective for the nitroaldol reaction, which is in accordance with the selectivity found for the formation of cyanohydrins. If nitro ethane was used as a nucleophile, the corresponding nitro alcohol compound comprised two stereocenters. The overall yield of the reaction of benzaldehyde and nitroethane was 67 % and the ratio of *anti/syn* was found to be 9:1 with an *ee* of the *anti* isomer of 95 %, corresponding to an overall content of 90 % (1*S*,2*R*) 2-nitro-1-phenyl-1-propanol.

The *HbHNL* can be overexpressed in *Pichia pastoris* in high amounts, but the required portion of enzyme for a reasonable conversion in the nitroaldol reaction is much higher than the amount necessary for the cyanohydrin formation reaction, which is a considerable disadvantage [27].

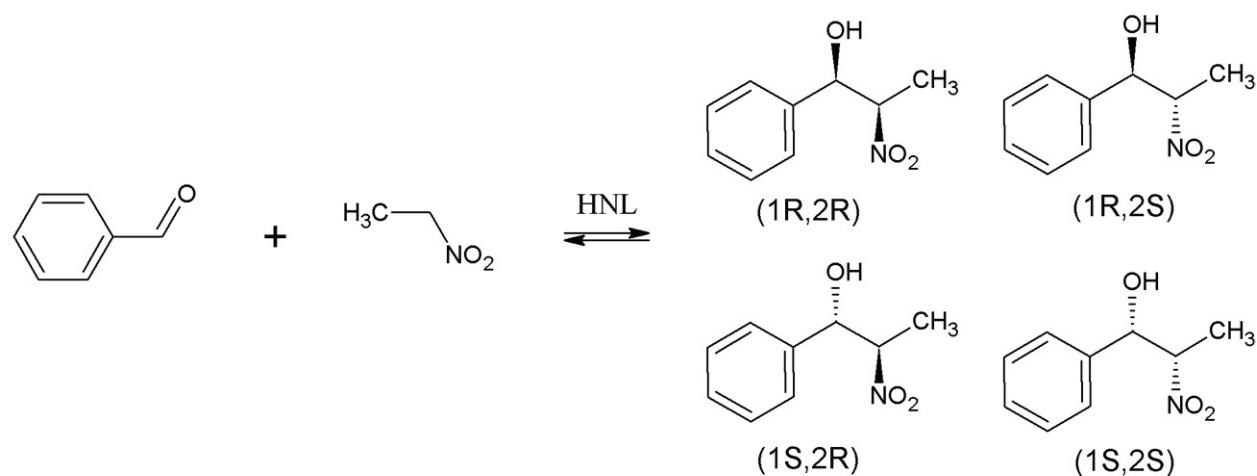
So far, also several other HNLs have been tested for nitroaldol activity, for example *MeHNL* from *Manihot esculenta* which shares a 75 % amino acid sequence similarity with *HbHNL*. However, even though an activity could be confirmed, it was about 4 times lower than the nitroaldolase activity of *HbHNL*. The same group also tested *Prunus amygdalus* HNL, but could not detect any nitroaldolase activity at all [27].

Another asymmetric nitroaldol reaction catalyzed by *AtHNL* from *Arabidopsis thaliana* was described by Fuhshuku and Asano. They showed that the *R*-selective *AtHNL* is also able to catalyze the addition of nitromethane to several different aldehyde compounds, forming the corresponding *R*- β -nitro alcohols. For this reaction, *n*-butyl acetate in a 50:50 ratio with buffer (pH 7) was applied and the reaction was incubated for 2 hours at 30 °C yielding 30 % conversion and an *ee* of 91 % with benzaldehyde as electrophile. The amount of enzyme that had to be applied for a good conversion and *ee* was once again very high (4000 U/mmol) [28].

Bekerle-Bogner *et al.* investigated the conversion of benzaldehyde and nitromethane or nitroethane with certain variants of *GtHNL* and *AcHNL* in the biphasic system. After 2 h of reaction time and with nitromethane as nucleophile, the best variant, *AcHNL*-A40H, yielded with 63 % conversion and >99 % *ee* (*R*) very promising results. With nitroethane as nucleophile the conversion was reduced to 16 % but the reported *ee* (*R, anti*) and *de* (*R,S*) (*R,R*) were 86 % and 54 %, respectively. Compared to *AcHNL*-A40H, another promising variant, *GtHNL*-A40R, achieved with 27 % a higher conversion of benzaldehyde but the *ee* (*R, anti*) and *de* (*R,S*) (*R,R*) were only 74 % and 10 %, respectively. Of all the variants tested by that group, with nitroethane as nucleophile, the *AcHNL*-A40H was the most promising with respect to selectivity and the *GtHNL*-A40R with respect to activity.

The biocatalytic nitroaldol reaction with benzaldehyde and nitroethane as substrates was investigated for this thesis. As biocatalysts, the above mentioned *R*-selective cupin-fold enzyme variants *AcHNL*-A40H and *GtHNL*-A40R were applied as basis. Since the pure (1*R*,2*S*) 2-nitro-1-phenyl-1-propanol isomer (see Scheme 3) is a valuable building block in, for example, the pharmaceutical

industry, the main goal of this diploma thesis was to find an enzyme variant that promotes a high conversion with benzaldehyde and nitroethane as substrates and additionally yields the (1*R*,2*S*) product in high enantiomeric and diastereomeric excess.



Scheme 3. The HNL catalyzed nitroaldol reaction with benzaldehyde and nitroethane as substrates, yielding four possible products.

3.3. Enzyme engineering

The enzymes AcHNL-A40H and GtHNL-A40R were used as parent variants for engineering by directed evolution in search of amino acid exchanges leading to a higher substrate conversion. Even though enzyme engineering could generally also be accomplished by rational design, the method of choice was directed evolution by random mutagenesis since no reliable reaction mechanism or substrate binding model was available for any of the two enzymes. Compared to rational design that relies on knowledge about the residues involved in catalysis, directed evolution can be carried out without such information that is partially difficult to achieve, but the need for suitable mutagenesis- and high throughput screening methods is still a major disadvantage [29].

In between rational design and directed evolution lies the so called semi-rational design, which usually leads to much smaller libraries compared to those prepared with directed evolution. One method that belongs to this category is the site-directed saturation mutagenesis (SSM). For this method, a targeted amino acid residue (e. g. a residue known to influence the enzyme activity) is substituted by all possible 20 canonical amino acids using degenerated primers containing the NNK codon at the respective position. Since N can be any base and K can be either G or T, SMM yields 32 possible codons encoding all 20 amino acids (12 encode duplicate amino acids) and only 94 clones have to be screened in order to cover 95 % of all amino acid substitutions at one position. SSM can also be performed with more than one position simultaneously, although this strongly increases the necessary amount of clones that have to be screened [29,30].

3.3.1. Directed evolution

In general, directed evolution can be divided into the non-recombining methods, where point mutations are randomly introduced into the target gene, and the recombining methods, where several parent sequences are arbitrarily recombined to form mosaic genes. For engineering of the two cupin-fold HNL variants, randomized mutagenesis of arbitrary nucleotides throughout the whole sequence length of the respective genes was achieved by using error prone PCR (epPCR). This method was first described in 1992 by Cadwell and Joyce and is based on the fact that a DNA polymerase may, with a certain probability, randomly insert incorrect nucleotides during DNA amplification [31]. For epPCR the *Taq* polymerase from *Thermus aquaticus* is applied frequently since it lacks proofreading activity and the error rate can be adjusted by the addition of Mn^{2+} , the increase of Mg^{2+} concentration and/or an imbalance of dNTPs [29,32]. A disadvantage of the *Taq* polymerase is that the bases A and T are mutated more frequently than the bases G and C, leading to a bias of the resulting mutations. This problem can be overcome by the use of other polymerases, for example the so called Mutazyme available from Stratagene [29]. An alternative method to introduce point mutations would be the use of a so called mutator strain, for example the *E. coli* derivative *Epicurian coli* XL1-Red which can introduce mutations during replication of a plasmid harboring the target gene sequence [33].

Another method that circumvents biased libraries and also promotes the simultaneous substitution of nucleotides that are positioned directly next to each other is called sequence saturation mutagenesis (SeSaM). For this method, randomly-sized fragments of the target gene DNA are generated and elongated with few units of a universal base, e.g. deoxyinosine, before the strands are elongated to the full length gene in a PCR mixture containing the single-stranded template sequence. During the following PCR-amplification, standard nucleotides are arbitrarily incorporated at the position complementary to the universal base and thus a randomly mutagenized library without bias can be created [34].

During the last decade, several different enzymes have been optimized with random mutagenesis using epPCR. For example the temperature stability of an Endoglucanase was increased by 92 %, SSM and epPCR were applied for engineering of Xylose reductase in order to reduce the enzyme's acceptance for unwanted side reactions and the same methods were used to highly increase the catalytic efficiency of an Aldolase [29].

3.3.2. Selection and Screening systems

After a library is finished, the real challenge of directed evolution is to establish a suitable selection- or screening assay that exhibits a high discriminatory capability and high throughput. Generally, selection assays are to be preferred since they exploit a growth advantage possibly introduced by mutagenesis

and positive variants will automatically be enriched, but selection is only possible with enzymes directly involved in the metabolism respectively growth behavior of an organism and thus not applicable for many objectives.

In contrast to selection, screening systems are used much more frequently. Under certain circumstances, specialized assays can be based on high-throughput GC-MS, HPLC or NMR systems (which require rather expensive equipment) or on methods like ribosome display, phage display and fluorescence activated cell sorting (which are laborious and not generally applicable) [29]. Colony based assays can be carried out in high throughput, when the enzyme-catalyzed reaction itself or a follow-up reaction forms colored or fluorescent products [35,36]. However, this method depends on the cellular uptake and release of reagents and is restricted to fast reactions because otherwise the diffusive effects would strongly impair the results.

Regardless of the considerable number of alternatives, the color- or fluorimetric microtiter plate (MTP) based assays are still the most versatile and widely-used methods for screening, whereas really high throughput can mostly be only achieved in combination with robot assistance. These assays are based on the photometric or fluorimetric measurement of product, side product or (sometimes) substrate concentration, where either an initial rate- or an end point measurement can be performed [29].

The aim of this work was to find variants of the enzymes *AcHNL-A40H* and *GtHNL-A40R* that show increased nitroaldolase activity and promote the formation of high ratios of the (1*R*,2*S*) 2-nitro-1-phenyl-1-propanol product in high enantiomeric and diastereomeric excess. To achieve this, the respective genes were first of all used as templates in an epPCR with *Taq* DNA polymerase and elevated concentrations of Mn^{2+} and Mg^{2+} . The thus resulting library was transformed into *E. coli* and screened for activity in the nitroaldol reaction with nitroethane and benzaldehyde as substrates. The screening assay was a 96-well MTP based optical assay where the enzyme mediated decrease of benzaldehyde was detected in an aqueous solution. For that, benzaldehyde was derivatized with the fluorogenic reagent 4-hydrazino-7-nitrobenzofurazane (NBDH) [32,37,38] and thereafter the absorbance of the corresponding hydrazone was measured photometrically. Drawbacks of this screening assay are that only the decrease of reactant can be measured, that side reactions cannot be completely suppressed and that the results do not provide any information on the ratio of isomers formed (there are four possible products after all). Therefore, the most promising variants were subsequently used for catalysis of the same reaction in an aqueous-organic biphasic system and the reaction products were analyzed via HPLC measurements. The final results of all the hits were then compared and evaluated with regard to overall benzaldehyde conversion and the conformation of the products.

4. Materials and Methods

4.1. Materials

If not stated otherwise, all chemicals were purchased from Sigma-Aldrich (St. Louis, MO, USA) or Roth (Karlsruhe, Germany). In general, all genes had been cloned into the pET-26b(+) vector (Novagen/Merck, Darmstadt, Germany), a graphical representation of the expression plasmid can be found in Figure 39 in the appendix [2]. Protein expression was always done in *E. coli* BL21 (DE3) Gold (Stratagene, LA Jolla, Ca, USA). The nucleotide as well as the protein sequences of both enzymes can be looked up in the appendix, Table 24.

4.2. Preparation of random mutagenesis libraries

Growth media and (stock-) solutions

LB-broth Lennox (pH 7.0 ± 0.2)

5 g/L sodium chloride
10 g/L tryptone
5 g/L yeast extract
In total 20 g dissolved in 1 L H₂O

LB-agar Lennox (pH 7.0 ± 0.2)

15 g/L agar-agar
5 g/L sodium chloride
10 g/L tryptone
5 g/L yeast extract
In total 35 g suspended in 1 L H₂O

SOC-medium

20 g/L bacto tryptone
5 g/L bacto yeast extract
3.46 g/L glucose
2 g/L magnesium chloride (MgCl₂)
2.46 g/L magnesium sulfate (MgSO₄)
0.18 g/L potassium chloride
0.58 g/L sodium chloride

Glycerol

30 %: 300 g of glycerol were mixed with 700 mL of H₂O

50 %: 500 g of glycerol were mixed with 500 mL of H₂O

Kanamycin-stock [40 mg/mL], (Roth)

Sterile filtered solution of 800 mg Kanamycin sulphate in 20 mL deionised H₂O

MnCl₂ [100 mmol/L], (Sigma-Aldrich)

Sterile filtered solution of 198 mg MnCl₂·4H₂O in 10 mL deionised H₂O

dNTPs [10 mM], (Thermofisher Scientific Inc., Waltham, MA, USA)

Starting from 100 mM stock solutions, 20 µL of each dNTP were mixed with 120 µL of deionized H₂O, yielding a total volume of 200 µL.

4.2.1. Isolation of plasmid templates

Two libraries were prepared, one with the parent variant ACHNL-A40H and a second one with the parent variant GtHNL-A40R. Both genes had previously been cloned into the pET-26b(+) expression plasmid and transformed into *E. coli* BL21 (DE3) Gold cells [2].

First of all, the two *E. coli* cryopreserved cell cultures that each comprised one of the genes coding for the parent variants were streaked out on agar plates containing LB-agar supplemented with kanamycin to a final concentration of 40 µg/mL (LB-Kan agar). These cultures were then used for plasmid isolation according to the GeneJET Plasmid Miniprep Kit (Thermofisher Scientific):

The cell material from an agar plate culture or the centrifuged pellet from a liquid culture were resuspended in 250 µL of resuspension solution before 250 µL of lysis solution were added. After very careful mixing, 350 µL of neutralization solution were added, the suspension was again mixed carefully and then centrifuged for 5 minutes. The supernatant was transferred to the provided Spin Columns and centrifuged for another minute. After two washing steps with 500 µL of wash solution, the flow-through was discarded and the empty column was again centrifuged for one minute. Thereafter, the column was transferred to a new tube and the DNA was eluted using 45 µL of deionised water pre-warmed to 60 °C and incubation for 2 minutes, before the assembly was centrifuged for 2 minutes. The resulting flow-through constituted the purified plasmid solution.

The purified plasmids served as templates for the subsequent PCRs. The DNA concentrations were measured with a NanoDrop 2000 UV-VIS Spectrophotometer (Thermofisher Scientific).

4.2.2. Error prone PCR

The method chosen for the preparation of randomly mutagenized megaprimers was epPCR using *Dream Taq* DNA Polymerase (Thermofisher Scientific) and elevated concentrations of MgCl₂ (7 mM) and MnCl₂ (varied between 0.1 mM and 0.4 mM). The exact composition of the epPCR mixtures can be seen in Table 1, the PCR temperature program is summarized in Table 2.

Table 1. Summary of the components of the epPCR mixture.

component	volume
Plasmid template (1 ng/μL)	2 μL
forward primer (10 pmol/μL)	2 μL
reverse primer (10 pmol/μL)	2 μL
dNTPs (10 mM stock)	1 μL
<i>Dream Taq</i> Buffer 10 x (Fermentas)	5 μL
MgCl ₂ (25 mM stock)	10 μL
MnCl ₂ stock (concentration either 4 mM, 2 mM, 1.5 mM, 1.3 mM or 1 mM)	5 μL
<i>Dream Taq</i> Polymerase (5 U/μL, Fermentas)	0.5 μL
H ₂ O	22.5 μL
Total volume	50 μL

Table 2. Temperature program of the epPCR.

step	temperature	time	repetition
Initial denaturation	95 °C	2 min	1
Denaturation	95 °C	30 s	35 cycles
Annealing	55 °C	30 s	
Extension	72 °C	20 s	
Final extension	72 °C	10 min	1
Hold	4 °C	∞	

The primers used for the epPCR were the Cu9_mut_fw1 forward primer and the T7term as reverse primer, producing 533 bp fragments that covered the whole length of the genes:

Cu9_mut_fw1 GTTAACTTTAAGAAGGAGATATACATATG

T7term GCTAGTTATTGCTCAGCGG

The PCR was carried out using a Thermocycler 2720 system (Thermofisher Scientific) and each mixture of mutagenized megaprimers was then loaded on a large preparative agarose gel as a first step of purification. The gel was prepared with 1 % agarose dissolved in TAE buffer (TRIS-Acetate-EDTA buffer) supplemented with ~30 μ L of 1 mg/mL ethidium bromide per 200 mL for detection. Before they were loaded on the gel, the PCR mixtures were mixed with 10 μ L of 6x loading dye (Thermofisher Scientific) each and of each mixture the whole volume (60 μ L) was loaded on the gel. As standard, 5 μ L of O'Gene Ruler Mix DNA ladder ready-to-use (Thermofisher Scientific) were used. Thereafter the gel was run for 80 min at 90 V. The gel was then viewed under UV light and the bands corresponding to fragments with ~530 bp length were cut out and transferred to 1.5 mL reaction tubes using clean scalpels. The gel slices were then weighed and the DNA was extracted and purified using the Wizard SV Gel and PCR Clean-up System (Promega, Fitchburg, WI, USA) before the DNA concentration of the resulting solution was determined.

4.2.3. Gel purification using the Wizard SV Gel and PCR Clean-Up System

For purification of the PCR products, 10 μ L of membrane binding solution were added per 10 mg of gel slice and thereafter the slice was dissolved by heating the mixture to 60 °C. In case of PCR clean-up, the PCR sample was simply diluted with an equal volume of membrane binding solution. Thereafter, the solution was transferred to a SV minicolumn assembly, incubated for one minute and centrifuged for another minute; the resulting flow-through was discarded. Then the minicolumn was washed two times with membrane wash solution (once with 700 μ L and once with 500 μ L) and the assembly was centrifuged one minute after the first and five minutes after the second washing step. Subsequently, the collection tube was emptied and the assembly was centrifuged for another minute prior to elution into a fresh reaction tube with 50 μ L of pre-warmed deionized H₂O.

4.2.4. QuikChange PCR

The purified mutagenized megaprimers were then used as primers in a QuikChange PCR in order to introduce the mutagenized genes into the pET-26b(+) vector. The composition of the QuikChange PCR reaction mixture is listed in **Table 3**, the thermal cycle conditions of the PCR in Table 4. The templates were the parent variant plasmids, the same as for the epPCR mixture. The PCR was carried out using a Thermocycler 2720 system.

Table 3. Summary of the components of the QuikChange PCR mixture.

component	volume
Plasmid template (100 ng/ μ L)	1 μ L
Megaprimer total 300 ng	x μ L
dNTPs (10 mM stock)	1 μ L
<i>Pfu-Ultra</i> Buffer 10 x MgSO ₄	5 μ L
<i>Pfu-Ultra High Fidelity</i> Polymerase (2.5 U/ μ L, Agilent)	1 μ L
H ₂ O	42-x μ L
Total volume	50 μ L

Table 4. Temperature program of the QuikChange PCR.

step	temperature	time	repetition
Initial denaturation	95 °C	2 min	1
Denaturation	95 °C	1 min	
Annealing	60 °C	30 s	30 cycles
Extension	72 °C	11 min	
Final extension	72 °C	10 min	1
hold	4 °C	∞	

4.2.5. *DpnI* digest

The finished PCR products were then purified using Promega's Wizard SV Gel and PCR Clean-up System, but the final elution step was carried out with 30 μ L of pre-warmed water. Thereafter, the remaining methylated plasmid template was digested using *DpnI* (10 U/ μ L, Thermofisher Scientific). The reaction mixture for the digestion is summarized in Table 5, it was carried out for 2 hours at 37 °C using 1.5 mL reaction tubes and a thermomixer device before the enzyme was deactivated for 30 min at 70 °C.

Table 5. Summary of the components of the *DpnI* digest reaction mixture.

component	volume [μL]
<i>DpnI</i> (10 U/ μ L)	2
Buffer Tango 10 x	5
Purified plasmid DNA	~25-30
H ₂ O	13

4.2.6. Transformation of *E. coli Top10F'*

After digestion, the mixture was desalted for 30 minutes using drop dialysis with mixed cellulose esters membranes (Merck). The resulting pure, desalted plasmid was then used for transformation of *E. coli Top10F'* electrocompetent cells. Electroporation was performed using an Electroporator Micro Pulser (Bio-rad, Hercules, USA) with the pre-set program Ec2 for bacteria in electroporation cuvettes. For the transformation, which was carried out three times per sample, 5 μL of plasmid were added to 80 μL of cooled electrocompetent cells and the pulse was applied after one minute of incubation on ice. Immediately after the pulse 600 μL of SOC-medium were added to the cells, the three transformation mixtures were combined and they were incubated at 37 °C, 550 rpm for 1 hour in a thermomixer device. After this regeneration period, kanamycin was added to the transformed cells to a final concentration of 40 $\mu\text{g}/\text{mL}$. Fifty and 100 μL of regenerated cell suspension were plated on LB-Kan agar plates and incubated overnight at 37 °C. The remaining cell suspension was, after addition of antibiotic and another hour of regeneration and selection under the same conditions, used for inoculation of 10 mL LB-Kan broth, which consisted of 20 g/L LB broth and kanamycin with a concentration of 40 $\mu\text{g}/\text{mL}$. This broth, which had been prepared in 50 mL tubes, was inoculated with 1 mL of the cell suspension and this pre-culture was then incubated at 120 rpm, 37 °C overnight in a multitron orbital shaker (Infors HT, Bottmingen, Switzerland).

After incubation, the cultures on the agar plates were used to evaluate the diversity of the respective library. The intended screening of around 7000 clones was only practical when at least the same number of successfully transformed cells had been present in the 1 mL of regenerated culture that had been used for inoculation of the overnight culture. Therefore, the colony forming units (CFUs) on the plates were counted and, depending on whether the found number of CFUs really exceeded 7,000 CFUs/mL, the liquid overnight culture was either used for library generation or discarded.

In case of a sufficient number of different clones per library, ten arbitrary colonies were streaked out on new LB-Kan agar plates and used for isolated plasmid preparations, as already described in part 4.2.1. The resulting plasmid solutions were then sent for sequencing (see part 4.3) and based on the results each library's average mutation rate was evaluated. The corresponding liquid culture was then used for a mixed plasmid preparation.

4.2.7. Generation of mixed plasmid preparation

After incubation, the overnight pre-culture was centrifuged for 5 min in an Eppendorf centrifuge 5810 R (Eppendorf, Hamburg, Germany) at 2,970 rcf before the supernatant was discarded. The cell pellet was taken up with 500 μL of the resuspension solution provided with the GeneJET Plasmid Miniprep Kit and after resuspension the cell suspension was split into two parts which were transferred to two

1.5 mL reaction tubes. Thereafter, the plasmids were isolated according to the standard Miniprep procedure as already described in part 4.2.1. The two resulting eluates, which constituted the mixed plasmid preparation, were pooled and stored at -20 °C until further use.

4.2.8. Transformation of *E. coli* BL21 (DE3) Gold

The general transformation procedure was similar to the transformation of *E. coli* Top10F' cells, but instead of those, electrocompetent *E. coli* BL21 (DE3) Gold cells were used and the total amount of DNA transformed per 80 µL of cells was 220 ng. Additionally, the transformation was carried out six times per library and the regeneration was stopped after 1 hour. Then the cell suspension was plated on six large, square LB-Kan agar bioassay trays in portions of 250, 400, 550, 650, 750 and 1000 µL, before being incubated at 37 °C overnight. The trays had been sterilized beforehand by washing them with ethanol and placing them under UV light for 30 min, before they had been filled with 300 mL of LB-Kan agar per plate.

4.2.9. Colony picking

The colonies that grew on the large agar plates had to be transferred to 384 well microtiter plates (384 well MTPs) for storage until further use. The sterile 384 well MTPs had been filled with 55 µL of LB-Kan broth per well and were then used as destination of the picked colonies. The picking itself was carried out using a Qpix2 picking robot (Genetix) equipped with a 96-position picking head. With the robot, around 7,000 colonies of each library were transferred to around 20 384 well MTPs per library. Of each MTP, the columns 21 and 22 were left sterile and were subsequently inoculated with the respective controls from agar plates using toothpicks. The controls and respective positions on the MTPs can be looked up in Table 6.

The MTPs were thereafter closed with lids and incubated overnight at 37 °C in buckets with a little bit of water to avoid exorbitant evaporation. After incubation, 35 µL of sterile 30 % glycerol were added to each well of the MTPs using a µFill Microplate Dispenser (BioTek Instruments Inc., Winooski, VT, USA), before the 384 well MTPs were sealed with Platesealer Aluminium sealing foil (Greiner Bio-One, Kremsmünster, Austria) and stored at -20 °C.

The 384 well plates that were not completely filled with clones were designated with X and a number, all the other plates that comprised clones at all intended positions were designated in arbitrary order with letters after the alphabet.

For the AchNL-A40H library, 23 of the 384 well plates were picked, of those 17 (designated A-Q) were full and 6 were only partially filled (designated X1-X6). All of those plates were then chosen for the screening.

For the *GtHNL-A40R* library, 24 of the 384 well plates were picked, of those 18 were full (designated A-S, no plate B) and 6 plates were only partially filled (designated X1-X6). Of those plates, the 18 complete ones and the plates X1, X2 and X6 were chosen for screening.

Table 6. Positions of the controls in the 384 well MTPs of the two random mutagenesis libraries.

AcHNL-A40H library			GtHNL-A40R library		
row	column		row	column	
	21	22		21	22
A	sterile		A	sterile	
B					
C	pET-26b		C	GtHNL-A40R	
D					
E					
F					
G	AcHNL-A40H		G	pET-26b	
H					
I					
J					
K	AcHNL		K	AcHNL-A40H	
L					
M					
N					
O	sterile		O	sterile	
P					

4.3. Sequencing

For all the genes that were sequenced, purified plasmid preparations (according to the GeneJET Miniprep kit) of the respective clones were made as a first step. Thereafter, 15 μ L of the eluate containing the purified plasmid were filled into new 1.5 mL reaction tubes and equipped with a barcode label supplied by Microsynth, the company that conducted the sequencing procedure. The primer used for sequencing of the HNL genes was the T7 primer from the Microsynth stock list.

4.4. Screening

4.4.1. Deep well plate fermentation

Since the screening procedure depended on an optical absorbance assay, it could not be conducted in 384 well plates, instead all the steps were carried out in 96 well deep well plates (DWPs) and 96 well microtiter plates.

As a first step, the cultures in the 384 well plates had to be transferred to DWPs for cultivation. For that, a pin stamp with 96 thin steel pins was used. The colonies were simply stamped from the frozen cultures to DWPs that had already been filled with 750 μ L of LB-Kan broth. Out of one 384 well MTP, four 96 well DWPs were stamped.

When the clones were transferred from the 384 well MTP to a 96 well plate, every second position of every second row of the 384 well plate was transferred. Thus, of groups of four wells positioned in a square on the 384 well plate, only one of every group (e. g. always the one in the upper left corner) was transferred. Based on this pattern, the nomenclature of the inoculated DWPs was set according to the stamped position of these groups of four wells. The group in the upper left corner of the 384 well MTP consisted of the positions A1, A2, B1 and B2. If the transferred well of this group was A1, the stamped DWP was designated with the letter code of the 384 well MTP and the additional code A1. In accordance with that the colonies of the 384 well MTP A were transferred to the DWPs A-A1, A-A2, A-B1 and A-B2. After each inoculation of a new DWP, the stamp pins were dipped in ethanol and thereafter flamed in order to re-sterilize the pins.

Immediately after being stamped, these DWPs that constituted the pre-cultures were incubated at 37 °C, 300 rpm overnight in a multitron orbital shaker. For the main cultures, DWPs were filled with 750 μ L LB-Kan broth supplemented with 100 μ M MnCl₂ that had been transferred from a sterile filtered 1000x stock of 100 mM MnCl₂ dissolved in deionized H₂O. The main cultures were inoculated by transferring 25 μ L of the respective pre-culture per well before they were again incubated at 37 °C, 300 rpm for 6 hours. For induction, LB-Kan broth containing 100 μ M MnCl₂ and 1.65 mM IPTG (provided from a 1 M sterile filtered stock solution of 0.238 g IPTG/mL deionized H₂O) was prepared and 50 μ L thereof were used for induction of each well, which led to a final IPTG concentration of 0.1 mM in the main cultures. The DWPs were then again incubated for 17 hours at 25 °C and 300 rpm before the cells were harvested by centrifugation in an Eppendorf centrifuge 5810 R for 20 minutes, 2,970 rcf. Thereafter, the supernatant was discarded and the DWPs with the pellets were stored at -20 °C prior to screening. Every DWP had to be completely frozen before the screening, so that uniform conditions could be assured.

All the rescreening DWPs were inoculated with toothpicks from the 384 well MTP cryo-cultures and also in the case of the AchNL-A40H/T50X site saturation mutagenesis library, the colonies had been

transferred to the DWPs with toothpicks directly from agar plates. Of these plates there were no existing cryo-cultures, therefore they were used to prepare new cryo-cultures out of the DWP pre-cultures: 50 μ L of 50 % glycerol were added to 100 μ L pre-culture in 96-well MTPs before the plates were sealed with Platesealer Aluminium sealing foil and stored at -20 °C.

4.4.2. The standard screening assay

Generally, the assay was based on three steps: cell lysis in DWPs, enzymatic reaction in MTPs and detection via absorbance measurement in MTPs. The substances necessary for screening are listed in Table 7 and the required buffers and screening solutions are specified below.

Table 7. List of chemicals and related data required for the screening assay.

name	formula	molecular weight	additional information
Potassium-di-hydrogen-phosphate	KH_2PO_4	136.086 g/mol	
Di-potassium-hydrogen-phosphate	K_2HPO_4	174.2 g/mol	
Citric acid	$\text{C}_6\text{H}_8\text{O}_7$	192.124 g/mol	
Lysozyme (Sigma-Aldrich)			70,000 U/mg
Benzonase (Merck)			250 U/ μ L
Triton-X 100 (Roth)			
Nitroethane (NE), (Sigma-Aldrich)	$\text{C}_2\text{H}_5\text{NO}_2$	75.07 g/mol	Toxic, use with care and dispose of accordingly
Nitromethane (NM), (Sigma-Aldrich)	CH_3NO_2	61.04 g/mol	
Benzaldehyde (BA), (Sigma-Aldrich)	$\text{C}_7\text{H}_6\text{O}$	106.121 g/mol	
4-hydrazino-7-nitrobenzofurazane (NBDH), (TCI, Tokyo, Japan)	$\text{C}_6\text{H}_5\text{N}_5\text{O}_3 \cdot \text{H}_4\text{N}_2$	227.18 g/mol	

Required Buffers

50 mM potassium-phosphate buffer (KPi) pH 6

123 mL of 50 mM K_2HPO_4 (6.80 g/L)

877 mL of 50 mM KH_2PO_4 (8.71 g/L)

The pH was adjusted to 6 ± 0.1 with 50 mM KH_2PO_4

0.15 M KPi pH 5.72

850 mL of 0.15 M KH_2PO_4 (20.413 g/L)

70 mL of 0.15 M K_2HPO_4 (26.13 g/L)

The pH was adjusted to 5.72 ± 0.1 with 0.15 M K_2HPO_4

10 mM KPi pH 7

390 mL of 10 mM KH_2PO_4 (1.36 g/L)

610 mL of 10 mM K_2HPO_4 (1.74 g/L)

The pH was adjusted to 7 ± 0.1 with 10 mM KH_2PO_4

0.3 M citrate-phosphate buffer (CiPi) pH 2.5

833 mL of 0.3 M citric acid (57.64 g/L)

167 mL of 0.3 M K_2HPO_4 (52.25 g/L)

The pH was adjusted to 2.5 ± 0.1 with 0.3 M K_2HPO_4

Required Solutions

Specifications for screening of one DWP

Lysis buffer (stored for no longer than 5 hours at 4 °C):

31.25 mL 10 mM KPi pH 7

250 μL lysozyme stock (stock solution contained 35 mg/mL lysozyme with 70,000 U/mg dissolved in H_2O and was stored at $-20\text{ }^\circ\text{C}$)

0.5 μL benzonase

62.5 μL Triton-X 100

Substrate solution with NM (prepared immediately before use, no storage!):

10.5 mL 0.15 M KPi pH 5.72

0.7125 mL NM (1.2 mM)

22.37 μL BA (20 mM)

Mixed thoroughly before use

Substrate solution with NE (prepared immediately before use, no storage! Dispose of appropriately):

11 mL 0.15 M KPi pH 5.72

0.3233 mL NE (0.4 mM); nitroethane had to be used in lower excess compared to nitromethane, since its solubility in water respectively buffer is much lower

22.63 μ L BA (20 mM)

Mixed thoroughly before use

4-hydrazino-7-nitrobenzofurazane: stored under exclusion of light for up to 2-3 weeks at 4 °C, do not use anymore if color darkens from bright yellow to brownish orange:

11 mL ethanol

1 mg NBDH

suspended 15 minutes in sonication bath, then centrifuged in an Eppendorf centrifuge 5810 R for 5 min, 2,970 rcf; supernatant used, decant immediately after centrifugation; remaining solid can be resuspended in ethanol and used to prepare more NBDH-solution, but it's important that the solution always is saturated.

Screening Procedure

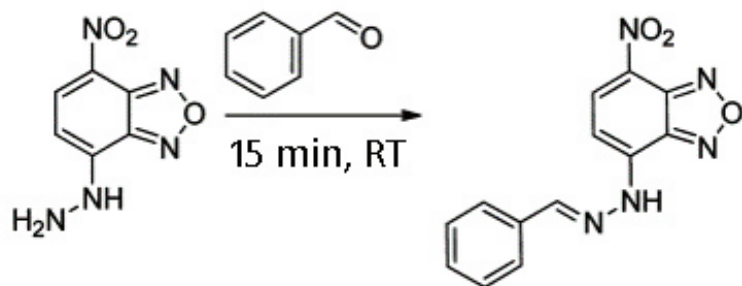
Cell lysis: To each well of the DWPs containing the frozen pellets, 300 μ L lysis buffer were added before the DWPs were incubated at 1,200 rpm for 1 hour at room temperature. Thereafter the DWPs were centrifuged for 20 minutes at 2,970 rcf in an Eppendorf centrifuge 5810 R.

All measurements were carried out at least in duplicate. Generally it is possible to perform five measurements out of the lysate in one DWP.

Enzymatic reaction: In an F-bottom MTP, 50 μ L of cell-free lysate (supernatant after centrifugation) per well were mixed with 50 μ L of the respective substrate solution containing either NE or NM and BA. This reaction mixture was then sealed with aluminum or plastic sealing foil and incubated for 1 h (NM as second substrate) or 90 min (NE as second substrate) at room temperature under exclusion of light.

Each plate also contained a blank, which consisted of 50 μ L of 10 mM KPi pH 7 and 50 μ L of 0.15 M KPi pH 5.72.

Detection: In a different F-bottom MTP, 6 μ L of the reaction mixture were added to 140 μ L 0.3 M CiPi pH 2.5, thus forming the measurement mixture. Thereafter 50 μ L NBDH solution per well were added with an automatic pipette during a time span of no more than 20 seconds in order to keep the bias of the results as low as possible. The reaction scheme of the detection is shown in Scheme 4.



Scheme 4. Detection reaction of NBDH and BA forming the colored product 4-hydrazino-7-nitrobenzofurazane.

Directly after NBDH addition, the current time was noted, so that each measurement plate could be incubated for 15 minutes. Thereafter, the measurement plates were mixed immediately and thoroughly. This was normally carried out using a 96 pin stamp and all wells were stirred simultaneously with the pins. However, stirring had to be done very thoroughly, not only a circling motion, but also an up- and down motion was performed to guarantee reasonably fast and thorough mixing. Also, since the pins were made of metal, it was very important that the contact between pins and inner plastic-surface of the wells was kept as little as possible. Otherwise, small plastic-particles of the MTPs could be abraded and strongly increased the measured absorbance of the respective well. Finally, after incubation, the absorbance of each well was measured at a wavelength of 522 nm.

Additional notes concerning the detection-procedure: The low pH in the measurement mixture is necessary to facilitate the detection-reaction. For BA detection, the BA concentration in the measurement mixture after addition of NBDH (volume of 196 μ L per well) should be no more than 0.15 mM, so that the measured absorbance does not exceed 0.8. Above a BA concentration of 0.15 mM, the relation between measured absorbance and concentration is not linear anymore. This can be seen in Figure 2, where a corresponding calibration curve of BA is shown.

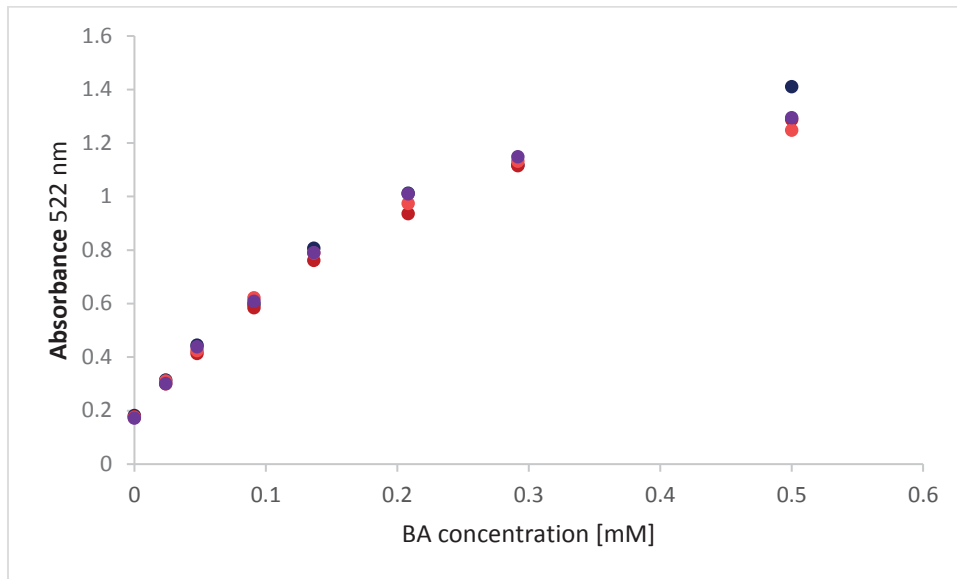


Figure 2. This figure shows the correlation between the BA concentration in the finished measurement mixture of around 200 μ L and the absorbance at 522 nm. The samples for each BA concentration were prepared by diluting the appropriate amount of BA with a 50:50 mixture of 10 mM KPi pH 7 and 0.15 M KPi pH 5.72 to BA concentrations between 0.5 and 10 mM. Thereafter, these samples were handled in the same way as the reaction mixtures of the screening assay, with the exception that 10 μ L of sample were added to 140 μ L of 0.3 M CiPi pH 2.5 instead of 6 μ L. Each concentration was prepared and measured four times.

The absorbance measurement itself had been carried out with two different plate readers:

FLUOstar Omega platereader (BMG Labtech, Ortenberg, Germany), applying the following settings:

Basic settings: Plate mode

Measurement type: Absorbance

Microplate: Greiner 96 F-bottom

Wavelength settings:

- Discrete wavelengths
- No. of wavelengths: 1
- Wavelengths: 522

General settings:

- Positioning delay: 0.1 s
- No. of flashes per well: 30

Eon™ High Performance Microplate Spectrophotometer (BioTek), applying the following settings:

Read Step:

- Wavelengths: 1, 522 nm
- Record Speed: normal

Temperature control: off

If the screening experiment is referred to as standard screening assay, NE was used as second substrate. If NM was used, it will be pointed out specifically. Additionally, the standard assay refers to the assay with cell pellets in DWPs and cell lysis. If cell free lysate obtained with shake flask fermentation (as described in part 4.7) was used, it will be described as 'standard screening assay with cell free lysate'.

All the screening results can be found on the acib server in the project folder 24051 as Microsoft Excel files designated according to the letter code of the 384 well MTPs.

4.4.3. Standard screening assay with cell free lysate

If cell cultivation was not performed in DWPs, the assay could also be carried out with cell-free cell lysates. For that, the lysates were diluted with 10 mM KPi pH 7 to a protein concentration of 1.5 to 5 mg total protein/mL. Most experiments in this work were done with total protein concentrations of 3 mg/mL. Thereafter, the diluted lysates were handled in the same way as the DWP lysates and were used for the enzymatic reaction (see part 4.4.2).

4.4.4. Time dependent screening assay

In general, the time dependent screening assay was based on the standard screening assay with cell free lysate. The aim of this experiment was to show the enzyme catalyzed decrease of BA over several time points. The measurement of all the time points had to be performed simultaneously in one single MTP, because the absorbance results cannot be compared to one another if the incubation time after NBDH addition is not exactly the same. Therefore, the reactions could not be started simultaneously, instead the reactions were started one after another. That means that the sample for 2 hours of enzymatic reaction time had to be started 2 hours before the detection and the sample for 1 hour of enzymatic reaction time had to be started 1 hour after the 2 hours sample and 1 hour before the detection and so on.

For the experiment itself the cell free lysate was diluted to 2 mg total protein/mL with 10 mM KPi pH 7 and aliquots of 50 μ L per time point and variant were prepared in one MTP. Additionally, the second substrate (NE or NM) was dissolved in 0.15 M KPi pH 5.72 to a final volume of 10 mL. For that, either 0.635 mL NM or 0.277 mL NE were used. This solution was used for each time point and stored under exclusion of light at 4 °C during the reaction time. Immediately before each time point, 1 mL of this NE solution was mixed with 2 μ L pure BA in a 1.5 mL reaction tube and thereafter 50 μ L of this substrate solution were added to each of the lysate samples for the respective time point. The corresponding wells of the MTP were then sealed with aluminum sealing foil and stored under exclusion of light at room temperature for the span between each time point. After the last time point,

the one for 'zero minutes', the reaction of all the samples was stopped by adding an aliquot of 6 μL of each reaction sample to 140 μL of 0.3 M CiPi PH 2.5 in a new MTP (in the same way as for the standard screening assay). Thereafter, the procedure was the same as for the standard screening assay described in Part 4.4.2.

4.5.Site specific mutagenesis

The variant AcHNL-T50A was prepared by site specific mutagenesis using QuikChange PCR, transformation of *E. coli* Top10F' and thereafter of *E. coli* BL21 (DE3) Gold, similar to the preparation of the random mutagenesis libraries.

In order to acquire the right template for the PCR, the plasmid comprising the AcHNL wild type variant was isolated using the GeneJet Miniprep kit as described in part 4.2.1. The QuikChange PCR mixture is summarized in Table 8 and the PCR temperature program can be looked up in Table 4, part 4.2.4. The PCR was carried out using a Thermocycler 2720 system. The following forward and reverse primer were used (exchanged nucleotides marked in yellow):

CH2-T50A-for: CCGGGTGCACGTGCCGCATGGCATAACC
 CH2-T50A-rev: GGTATGCCATGCGGCACGTGCACCCGG

Table 8. Summary of the components of the QuikChange PCR mixture for site specific and site saturation mutagenesis.

component	volume
Plasmid template (100 ng/ μL)	1 μL
Forward primer (10 pmol/ μL)	2 μL
Reverse primer (10 pmol/ μL)	2 μL
dNTPs (10 mM stock)	1 μL
<i>Pfu-Ultra</i> Buffer 10 x MgSO_4	5 μL
<i>Pfu-Ultra High Fidelity</i> Polymerase (2.5 U/ μL , Agilent)	1 μL
H_2O	40 μL
Total volume	50 μL

After the QuikChange PCR was finished, the mixture was purified, digested with DpnI and desalted prior to transformation of *E. coli* Top10F', the procedure of those steps can be looked up in Part 4.2.5 and 4.2.6. After 1 hour of regeneration, 100 μL of transformed cell suspension were plated on a

LB-Kan agar plate. The remaining suspension was centrifuged for around 6 seconds in an Eppendorf table top centrifuge before the majority of supernatant was poured off and the pellet was carefully resuspended in the remaining liquid before the whole volume was plated on another LB-Kan agar plate. Both plates were incubated overnight at 37 °C.

Thereafter, 16 colonies of these plates were spotted on a new ‘master plate’ agar plate and incubated for 4.5 hours at 37 °C. These colonies were then used for analysis with colony PCR and agarose gel electrophoresis. The components of a colony PCR mixture can be seen in Table 9, the corresponding temperature program is shown in Table 10 and the PCR itself was carried out using a Thermocycler 2720 system. As template, very small amounts of cell material from the master plate were transferred to the PCR mixture using a toothpick, and the T7 promotor forward primer and the T7 terminator reverse primer were used:

T7prom: TAATACGACTCACTATAGG
T7term: GCTAGTTATTGCTCAGCGG

Table 9. Summary of the components of the colony PCR mixture.

component	volume
Template (whole cells)	picked
Forward primer (10 pmol/μL)	0.5 μL
Reverse primer (10 pmol/μL)	0.5 μL
dNTPs (10 mM stock)	0.5 μL
<i>Dream Taq</i> Buffer	2.5 μL
<i>Dream Taq</i> Polymerase (5 U/μL, Fermentas)	0.125 μL
H ₂ O	20.5 μL
Total volume	25 μL

Table 10. Temperature program of the colony PCR.

step	temperature	time	repetition
Initial denaturation	95 °C	10 min	1
Denaturation	95 °C	30 s	25 cycles
Annealing	55 °C	30 s	
Extension	72 °C	45 s	
Final extension	72 °C	7 min	1
Hold	4 °C	∞	

Subsequently, the colony PCR samples were analysed using analytical agarose gel electrophoresis. The gel was prepared with 1 % agarose dissolved in TAE buffer (TRIS-Acetate-EDTA buffer) supplemented with ~30 µL of 1 mg/mL ethidium bromide for detection. A sample volume of 18 µL colony PCR mix and 7 µL of O'GeneRuler 1kb Plus DNA Ladder, ready-to-use (Thermofisher Scientific) standard were applied and the gel was run for 60 min at 120 V. The resulting bands were then analysed and photographed under UV light.

Upon inspection of the gel, 4 clones comprising the appropriately sized fragment of 594 base pairs were chosen to be sent for sequencing.

Finally, the purified plasmid (using the GeneJET Miniprep kit) of one correctly mutated gene was transformed into *E. coli* BL21 (DE3) Gold cells as has already been described in Part 4.2.8. Only one transformation sample was prepared using 80 µL of electrocompetent cells and 1 µL of plasmid with a concentration of 130 ng/µL.

One successfully transformed clone was then used for shake flask fermentation and production of cell free lysate which could be used for the HPLC measurement.

4.6.Site saturation mutagenesis

The principle of the site saturation mutagenesis which was used to prepare the AchNL-A40H/T50X library was very similar to that of the site specific mutagenesis described above.

The first step, QuikChange PCR was exactly the same as for the production of the AchNL-T50A variant (see part 4.5), only the applied primers were different. In order to possibly introduce all 20 canonical amino acids at position 50, degenerated primer containing the NNK triplet were applied, where N stands for any of the four bases, M is either A or C and K is either G or T:

CH2-T50X-for: CCGGGTGCACGTNNKGCATGGCATAACC

CH2-T50X-rev: GGTATGCCATGCMNNACGTGCACCCGG

All the following steps were similar to the site specific mutagenesis too, but the transformation of *E. coli* Top10F' was carried out two times with 80 µL of cells each instead of only one time as it had been the case for the AchNL-T50A variant. Additionally, the two transformation mixtures were not regenerated for only 1 hour but for 2 hours, while after the first hour of regeneration kanamycin to a final concentration of 40 µg/mL had been added to both transformation mixtures. Thereafter, the two suspensions were pooled and 100 µL were plated on LB-Kan agar plates and incubated overnight at 37 °C. Additionally, 1 mL was used to inoculate 10 mL of LB-Kan broth, as had also been done for the preparation of mixed plasmid preparations of the random mutagenesis libraries. After incubation at

37 °C, 120 rpm overnight, this culture was used to obtain a mixed plasmid preparation in exactly the same way as described in part 4.2.7.

Sixteen of the colonies on the agar plate were used for a colony PCR and sequencing, as had already been described in part 4.5.

The mixed plasmid preparation was then used for transformation of *E. coli* BL21 (DE3) Gold cells as described above. For the transformation, 5 µL purified DNA with 180 ng/µL were applied on 80 µL of cells. The high amount of DNA was necessary because application of less DNA only yielded an insufficient amount of successfully transformed *E. coli* BL21 (DE3) Gold cells. Regeneration was carried out for 1 hour before 100 µL cell suspension and the remaining pellet were plated on LB-Kan agar and incubated at 37 °C overnight.

The colonies on the agar plates were then transferred to five MTPs filled with 100 µL of LB-Kan broth per well using toothpicks. The column 11 was kept sterile and subsequently inoculated with control strains: B11+C11: pET-26b negative control; D11+E11: AcHNL-A40H; F11+G11: AcHNL wild type; A11+H11: sterile.

After the inoculation, the MTPs were equipped with sterile lids and incubated overnight at 37 °C, 500 rpm in an Eppendorf thermomixer device equipped with an adapter for MTPs. The next day, 50 µL of 30 % glycerol were added to each well, the plates were sealed with aluminium sealing foil and stored at -20 °C as cryo-stocks until the screening.

4.7.Preparation of cell free lysate

4.7.1. Shake flask fermentation

For the pre-cultures, ten 15 mL tubes were filled with 6 mL LB-Kan broth and inoculated from cryo-cultures or agar plates using toothpicks. The cultures were then incubated overnight at 37 °C, 120 rpm in a multitron orbital shaker. For the main cultures, baffled shaking flasks were filled with 400 mL LB-Kan broth supplemented with 100 µM MnCl₂ (from 100 mM stock solution) each before the whole pre-culture volume of 6 mL was added to the respective main culture. Then the main cultures were incubated for around 3 hours at 37 °C and 120 rpm until an OD₆₀₀ of 0.8-1 was reached. Thereafter, the cultures were induced with 40 µL 1 M IPTG (sterile filtered) each, yielding a final IPTG concentration of 0.1 mM. These cultures were then again incubated at 25 °C, 120 rpm for around 20 hours, before they were harvested by centrifugation.

4.7.2. Cell harvest and disruption

Cell harvest was performed by transferring the whole cell broth to 500 mL centrifugation tubes before they were centrifuged with a JA-10 rotor for 20 min at 4 °C and 4,400 rcf in an Avanti J-20 XP centrifuge (Beckman Coulter, Brea, CA, USA). The supernatant was discarded and the pellets were weighed and transferred into 50 mL tubes before they were resuspended in 25 mL 50 mM KPi pH 6. Thereafter, the cells were disrupted by sonication using a Branson Sonifier 250 (Emerson, St. Louis, MO, USA) for 6 minutes at 80 % duty cycle and 70 % output control. Afterwards the crude lysate was centrifuged for one hour at 75,600 rcf in a JA 25.50 rotor before the supernatant was filtered through PE 0.45 µm syringe filters (Agilent Technologies, Santa Clara, CA, USA). After an aliquot had been withdrawn for determination of protein concentration and sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) analysis, the remaining lysate was filled into 1.5 mL reaction tubes in portions of 1 mL and stored at -20 °C until further use.

4.7.3. Bradford protein assay

The protein concentration of the lysates was determined using the Bradford assay. The calibration line had been established beforehand using bovine serum albumin with concentrations between 1.5 and 12 µg/mL. Bradford reagent (Bio-rad) was diluted 1:5 with deionized water and the lysate samples were diluted 1:50 with 50 mM KPi pH 6. Then, 50 µL of diluted sample were added to 950 µL of diluted Bradford reagent in a semi-micro cuvette (Greiner Bio-One) and incubated for at least 10 minutes prior to the absorbance measurement at 595 nm. A Cary Series UV-VIS Spectrophotometer (Agilent Technologies) was used for all the measurements. For the calculation of the respective protein concentrations, the following calibration equation was used:

$$\text{Abs} = 2.78987 * \text{Conc} + 0.06217$$

4.7.4. SDS-PAGE analysis

For SDS-PAGE analysis the lysates were diluted to a concentration of 1 mg/mL with deionized H₂O. The pellet samples (the pellet yielded upon cell disruption and centrifugation) were prepared by taking 20-40 mg thereof from the middle part of the pellets surface and adding ten times the volume of 6 M urea. After the pellets were resuspended by vortexing, they were centrifuged in a table top Eppendorf centrifuge for 5 minutes at 13,200 rpm. The supernatant was then diluted 1:2 with deionized H₂O. Of the pellet- and lysate samples, 13.5 µL were mixed with 1.5 µL NuPAGE sample reducing agent (10x) and 5 µL NuPAGE LDS sample buffer (4x). This mixture was then heated to 95 °C for 10 minutes, followed by short centrifugation (~30 seconds, table top centrifuge, 13,200 rpm). Of these mixtures 10 or 8 µL were loaded in 10- (10 µL sample) or 15- (8 µL sample) slot Novex NuPAGE 4-12 % Bis-Tris gels.

As standard 3.5 μ L PageRuler Prestained protein ladder (ThermoFisher Scientific) were used and Novex NuPAGE MES SDS Running Buffer as running buffer. For the run, the predefined NuPAGE gel program in an XCell SureLock™ Mini-Cell Electrophoresis System was applied. All NuPAGE materials including the electrophoresis system were obtained from Life Technologies (Carlsbad, CA, USA).

The gels were then stained for 20 minutes in a solution containing:

- 500 mL ethanol
- 75 mL acetic acid
- 2.5 g Coomassie Brilliant Blue G250
- 425 mL deionized H₂O

Thereafter, destaining was performed for around one hour, until the gels were completely destained. The destaining solution, which was changed every 20 minutes, was composed as follows:

- 200 mL ethanol
- 75 mL acetic acid
- 725 mL deionized H₂O

4.8. Final characterization of hits

4.8.1. Biphasic reaction system

In order to prepare samples for a detailed measurement of conversion and enantioselectivity, the Henry reaction was carried out in a biphasic system consisting of aqueous cell free lysate and TBME (methyl tertiary butyl ether) as organic phase. The method for the Henry reaction in the aqueous-organic system and the HPLC analysis were based on experiments carried out by Fuhshuku *et al.* [28]. The reactions were carried out in two sets, once with the hits and controls of the AcHNL-A40H library and once with the corresponding variants of the GtHNL-A40R library (see Table 11). Each variant, control and time point (2 and 24 hours) was measured in duplicate, which means that also the reaction had been carried out two times. Only the reference sample, which contained 500 μ L of 50 mM KPi pH 6 instead of cell free lysate, was measured three times, since it was used to evaluate the BA conversion of the other samples. These reference samples were prepared and handled in the same way as the other samples (as described below) but they did not contain any enzyme or lysate and were not incubated either, the reaction was stopped by centrifugation immediately after mixing.

Table 11. List of the variants and controls of both sets measured with HPLC.

	AcHNL-A40H library set	GtHNL-A40R library set
variant	AcHNL-A40H/T50A	GtHNL-A40R/Q110L
variant	AcHNL-V23I/V25I/A40H	GtHNL-V23A/A40R
variant	AcHNL-V25I/A40H	GtHNL-A40R/C67S
variant	AcHNL-A40H/T50S	GtHNL-A40R/T50S
variant	AcHNL-A30T/A40H/E119G/D126G	GtHNL-A40R/K93R
control	AcHNL-A40H	GtHNL-A40R
control	pET-26b	pET-26b
reference sample	buffer	buffer

For each set an own substrate solution was prepared so that the samples of a set could be normalized to the internal standard. One reaction mixture consisted of 500 μ L of cell free lysate with a total protein concentration of 5 mg/mL in 50 mM KPi pH 6 and 500 μ L of substrate solution, the composition of which is summarized in Table 12.

Table 12. List of the components of organic substrate solution for a single HPLC sample. However, for each of the two sets of samples an extra, pre-mixed batch of substrate solution had been prepared.

component	volume [μL]	final concentration [mol/L]
Benzaldehyde (Sigma-Aldrich)	2	0.02
Nitroethane (Sigma-Aldrich)	71	1.00
Internal Standard (1,3,5-triisopropylbenzene)	2	
TBME (methyl tertiary butyl ether)	425	

Immediately after the cell free lysate and the substrate solution had been mixed in a 1.5 mL reaction tube, the mixture was placed in an Eppendorf Thermomixer device and shaken for 2 or 24 hours at 30 °C and 1,200 rpm. Thereafter, the reaction was stopped by centrifugation for 5 minutes at 13,200 rpm, 4 °C in an Eppendorf table top centrifuge and subsequent addition of 50 μ L of the organic phase (which was the lighter and thus upper one) to 450 μ L of the HPLC solvent, n-Heptane:2-propanol 9:1 in an HPLC vial. Since both the HPLC solvent and the TBME solution are very volatile, the pipetting had to be performed as fast as possible and the HPLC vial was capped immediately after addition of sample and solvent.

4.8.2. HPLC analysis

The measurement was carried out using an Agilent 1100 Series HPLC system and a CHIRALPAK AD-H column (DIACEL) with the following specifications: column length 250 mm, diameter 4.6 mm and particle size 5 μ m; As solvent, n-heptane: 2-propanol in a ratio of 9:1 was applied and the flow was set

to be isocratic with 0.9 mL/min for 20 min per run. The temperature was set to 20 °C and the injection volume was always 10 µL. The detection of peaks was performed spectrophotometrically at a wavelength of 210 nm.

The evaluation of the HPLC results was done by normalizing the peak areas of the analytes to the respective area of the internal standard for each sample. The respective BA conversion was calculated based on the normalized areas corresponding to the BA peaks of the respective sample and the average of the three reference samples, according to equation 1.

$$BA \text{ conversion } [\%] = 100 * \left(1 - \frac{\frac{\text{peak area } BA_{\text{sample}}}{\text{peak area } IS_{\text{sample}}}}{\frac{\text{peak area } BA_{0' \text{ sample}}}{\text{peak area } IS_{0' \text{ sample}}}} \right) \quad \text{equ. 1}$$

The enantiomeric excess *ee*, diastereomeric excess *de* and the ratio *anti/syn* were calculated according to the following equations (equ. 2-5).

$$ee (R) \text{ anti } [\%] = \left(\frac{\text{peak area}_{1R,2S} - \text{peak area}_{1S,2R}}{\text{peak area}_{1R,2S} + \text{peak area}_{1S,2R}} \right) \cdot 100 \quad \text{equ. 2}$$

$$ee (R) \text{ syn } [\%] = \left(\frac{\text{peak area}_{1R,2R} - \text{peak area}_{1S,2S}}{\text{peak area}_{1R,2R} + \text{peak area}_{1S,2S}} \right) \cdot 100 \quad \text{equ. 3}$$

$$de \text{ anti } (R, S) [\%] = \left[\frac{(\text{peak area}_{1R,2S} + \text{peak area}_{1S,2R}) - (\text{peak area}_{1R,2R} + \text{peak area}_{1S,2S})}{(\text{peak area}_{1R,2S} + \text{peak area}_{1S,2R}) + (\text{peak area}_{1R,2R} + \text{peak area}_{1S,2S})} \right] \cdot 100 \quad \text{equ. 4}$$

$$\text{anti/syn} = \frac{ee (R) \text{ anti}}{ee (R) \text{ syn}} \quad \text{equ. 5}$$

5. Results and Discussion

Out of several different *AcHNL* and *GtHNL* variants comprising nitroaldolase activity, the most promising ones were the *AcHNL*-A40H and the *GtHNL*-A40R. These two enzymes had thus been chosen as basis for further rounds of engineering aimed to identify variants with even better conversion and enantioselectivity. With the corresponding genes as templates, random mutagenesis libraries for the two variants were prepared and screened for increased substrate conversion.

Even though the final characterization of the best variants was to be carried out in an organic/aqueous biphasic system followed by HPLC measurements, the screening itself and the preliminary characterization of the variants was performed in an aqueous system applying a screening assay that had been prepared especially for that purpose (see diploma thesis of Simone Petek and project laboratory report of Silvia Maitz). The use of this screening assay led to a huge increase in throughput compared to HPLC measurements but also to a decrease of sensitivity and the inability to distinguish between the four possible nitro alcohol products.

5.1.Characterization of parent variants in aqueous system

As a preliminary test, enzymatic conversion of BA over time was examined applying the time dependent screening assay. For this experiment, as well as for all the relevant screening measurements, nitroethane and benzaldehyde were used as substrates.

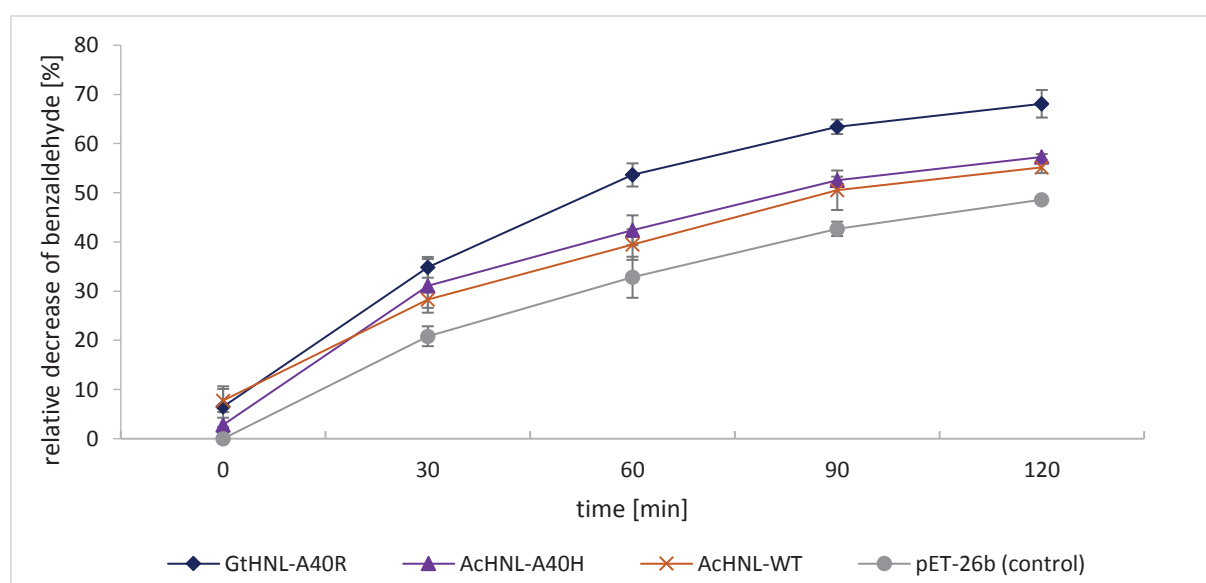


Figure 3. Time dependent relative decrease of BA, applying cell free lysates of the two enzyme variants *GtHNL*-A40R and *AcHNL*-A40H, as well as the *AcHNL* wild type (WT) and the empty vector control pET-26b as controls. The decrease of BA was related to the BA concentration in the 0 minutes sample of pET-26b. Measurements were performed four times for each sample and time point according to the time dependent screening assay with NE.

As can be seen in Figure 3, the best distinction between the BA conversions achieved by the different variants can be made after a reaction time of 60 to 90 minutes, which is the reason why the timespan for the enzymatic reaction applied in the screening assay was set to 80-90 minutes. In general, the variant *GtHNL-A40R* yielded by far the highest conversion of BA, after 90 minutes the amount of BA present in the reaction mixture had been reduced to 37 % of the initial concentration. In comparison, the *AcHNL-A40H* variant only led to a residual BA concentration of 47 % after the same reaction time. The *AcHNL-WT* merely showed a slightly higher residual concentration of 49 %.

The decrease of BA was also strongly influenced by the background reaction which is comparably fast in an aqueous system. After again 90 minutes, even the BA concentration in the empty vector control samples was reduced to 57 % of the initial concentration. Up to a certain degree, this unspecified decrease of substrate is caused by non-catalytic formation of the nitro alcohol products, as can be seen clearly upon examination of the data yielded upon HPLC measurements (see part 5.13). Additionally, in presence of water and oxygen, the oxidation of BA to benzoic acid and the unspecific enzymatic reduction to benzyl alcohol both play a role in the decrease of BA over time (see diploma thesis of Bettina Kothbauer). The sum of both effects and possibly even further side reactions only relevant in aqueous solution leads to the unspecific background conversion of BA, which hampered the identification of potential hits during the screening experiments. These background reactions are not nearly as dominant in the organic/aqueous biphasic system, which is one of the main reasons to conduct the final characterizations of the variants applying enzymatic reactions in the biphasic system followed by HPLC measurements.

Even though the nitro alcohol formation was only slightly accelerated by the *AcHNL-A40H* compared to the WT enzyme, this variant was still chosen as a basis for further engineering because its enantioselectivity was by far the best, as is also shown by the results of the HPLC measurements (see part 5.13). On the other hand, the *GtHNL-A40R* achieved a rather high conversion of BA but the selectivity was distinctly lower compared to the *AcHNL-A40H* [19]. Therefore, both variants had been chosen for further engineering by random mutagenesis, aiming to find enzyme variants that show both characteristics, a high selectivity and a high conversion.

5.2. Library preparation

The random mutagenesis libraries of both genes were prepared using epPCR with *Dream Taq* DNA polymerase and elevated Mn^{2+} as well as Mg^{2+} levels in order to assure an increased error rate during the PCR. As a first step, the epPCR was carried out with various Mn^{2+} concentrations, so that the ideal conditions for a mutation rate of around three nucleotide exchanges per open reading frame (ORF)

could be identified. Manganese concentrations of 0.1, 0.2 and 0.4 mM were tested first with both genes and subsequently up to four clones per concentration-library were sequenced in order to estimate the average mutation rate. However, none of these manganese concentrations led to a satisfactory number of mutations, the exact numbers of which can be seen in Table 25 in the appendix. Therefore, another library was prepared applying 0.15 mM of Mn^{2+} , which yielded a reasonably well result with 3.1 nucleotide exchanges per coding region for the *AcHNL-A40H* library (a graphical representation of the nucleotide exchanges of the 10 analysed genes can be seen in Figure 6, part 5.3). The corresponding *GtHNL-A40R* library yielded 4.7 nucleotide exchanges per coding region which was a comparably high mutation rate. Nevertheless, a pre-screening experiment was carried out with both libraries, in order to examine the influence of the mutation rate on the portion of still active variants remaining among all the clones.

For the pre-screening, the mixed plasmid preparations of the two libraries were transformed into *E. coli* BL21 (DE3) Gold cells and around 45 variants of each library were screened for conversion of BA and NM or NE as substrates, applying the standard screening assay with NM or NE, respectively.

The additional use of the screening assay with NM as second substrate for the pre-screening was beneficial because most variants convert BA and NM much faster than BA and NE. Additionally, the unspecific background reaction is less dominant in reaction mixtures containing NM as second substrate instead of NE. This effect can be seen in Figure 40 in the appendix that shows the time dependent decrease of BA catalysed by *AcHNL-A40H* with NM as second substrate, compared to Figure 3 in part 5.1 for NE as second substrate. Due to this difference in activity, the distinction between active and inactive variants could be effectively facilitated by the use of BA and NM as substrates.

Figure 4 and Figure 5 show the pre-screening results for the two libraries obtained with 0.15 mM Mn^{2+} in the epPCR. Since during the detection procedure the absorbance of the BA-NBDH reaction product is measured, a lower absorbance value at the detection wavelength of 522 nm corresponds to a lower BA concentration and thus to a higher enzyme activity. The comparison of the two libraries immediately shows that for both enzymes the conversion obtained with BA and NM is higher compared to the conversion obtained with NE as second substrate, as was to be expected. Furthermore, the activity of the *AcHNL-A40H* variant with NM and BA is considerably higher than the activity of the *GtHNL-A40R* with the same substrate. This is contrary to the ratio of activities that had been found with BA and NE as substrates, as was shown in part 5.1, Figure 3. This discrepancy once more shows that it is very important to use the relevant substrate, in this case NE, for the screening. Similar preferences were also reported by Bekerle-Bogner *et al*, who investigated the conversion of BA and NM or NE in the biphasic reaction system. After 2 h of reaction time, the *AcHNL-A40H* converted

63 % of BA with NM and 16 % of BA with NE as nucleophile. During the same time, the *GtHNL-A40R* only converted 56 % of BA with NM as nucleophile but 27 % of BA with NE as nucleophile [19].

The pre-screening results also showed that the mutation rate of 4.7 nucleotide exchanges per ORF was clearly too high, only 53.5 % of the variants in the *GtHNL-A40R* library still showed at least residual nitroaldolase activity. On the other hand, exactly 2/3 of all the analysed variants of the *AcHNL-A40H* library still exhibited some activity. Consequently, a new library was prepared for the *GtHNL-A40R* variant, using 0.13 mM Mn^{2+} in the epPCR. Exact mutation rates and further data on this library, can be found in part 5.8, Figure 22.

Other than the *GtHNL-A40R* library, the *AcHNL-A40H* library was suitable to be used for the subsequent screening experiments. Actually, the very first potential hit had already been found in the pre-screening with NE. The variant No. 6 of the *AcHNL-A40H* pre-screening plate clearly showed a lower BA concentration compared to the controls (not shown) and all the other variants on that plate. This variant was thus chosen for the rescreening that was performed after all the variants cultivated for the *AcHNL-A40H* library (around 7,000 clones) had been screened for nitroaldolase activity. Therefore, with a promising variant even before the actual screening experiments had begun, the screening of the *AcHNL-A40H* library was ready to start.

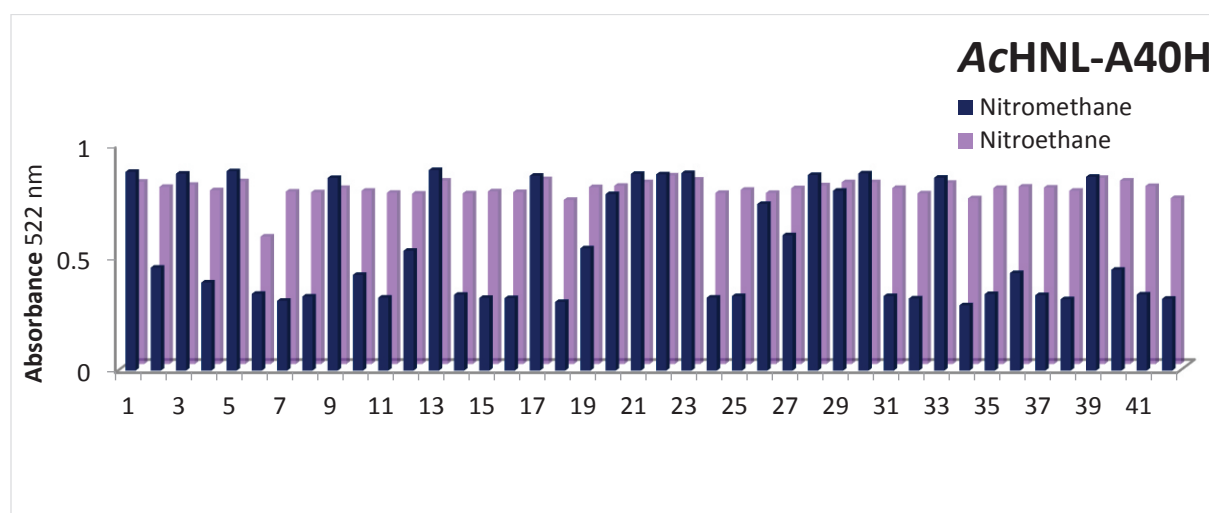


Figure 4. Result of the pre-screening of the *AcHNL-A40H* library with 0.15 mM Mn^{2+} . The variants were screened two times, once according to the screening assay with BA and NM as substrates and once with BA and NE as substrates and the corresponding assay. Absorbance at 522 nm is directly proportional to residual BA concentration.

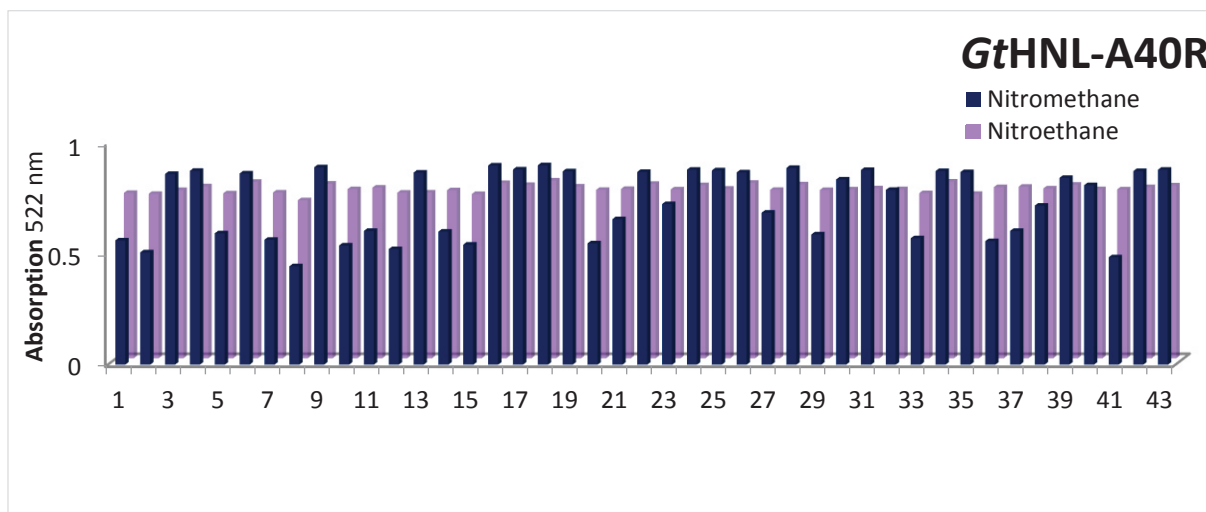


Figure 5. Result of the pre-screening of the *GtHNL-A40R* library with 0.15 mM Mn^{2+} . The variants were screened two times, once according to the screening assay with BA and NM as substrates and once with BA and NE as substrates and the corresponding assay. Absorbance at 522 nm is directly proportional to residual BA concentration.

5.3. *AcHNL-A40H* library screening

As already mentioned in part 5.2, ten arbitrarily chosen clones of the randomly mutagenized *AcHNL-A40H* library in *E. coli* TOP 10F' had been sequenced to estimate the mutation rate prior to the transformation of *E. coli* BL21 (DE3) Gold cells. The number of nucleotides exchanged per sequence can be seen in Figure 6, the average mutation rate was 3.1 nucleotide exchanges per ORF.

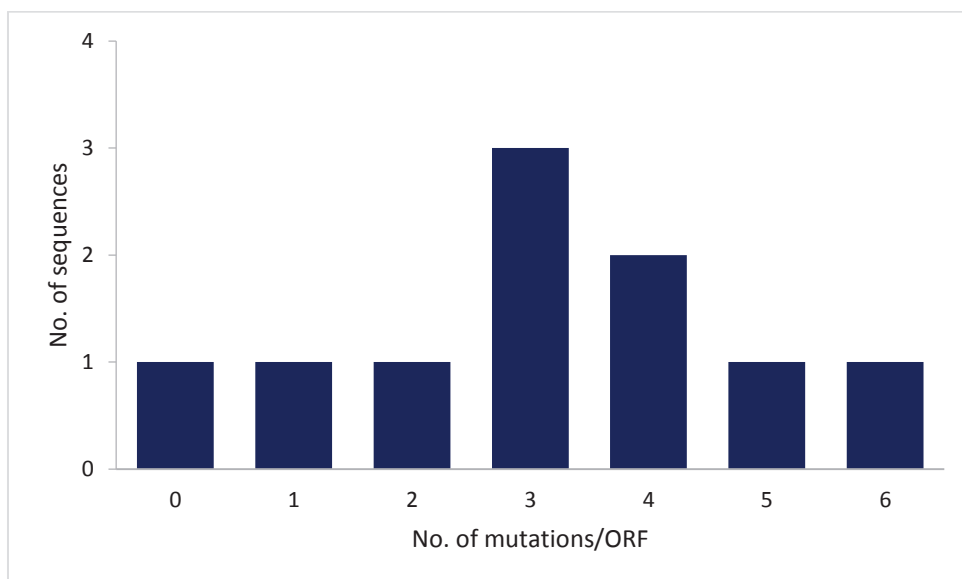


Figure 6. Analysis of nucleotide exchanges in the *AcHNL-A40H* coding region of ten arbitrarily chosen clones of the random mutagenesis library that was prepared with epPCR applying 0.15 mM Mn^{2+} .

The screening of the *AcHNL-A40H* library was carried out according to the standard screening assay with BA and NE as substrates. The following Figures (Figure 7-Figure 11) show a few examples of the screening results. The complete screening results of the *AcHNL-A40H* library and the *GtHNL-A40R*

library as well can be found on the acib server in the project folder 24051. The designation of the screened DWPs was defined with regard to the original 384 well MTPs that had been named by alphabet in arbitrary order. The corresponding DWPs were named after the respective 384 well MTP with the additional code -A1, -A2, -B1 or -B2. A more detailed description of the screening plates' designation can be found in part 4.4.1.

Each screened DWP, containing 88 wells with library clones and 8 wells with controls, was analysed individually, since a comparison of different plates would have been very difficult. The reason for that is the time dependent colour development during the detection procedure: Since the absorbance maximum of the BA-NBDH reaction product shifted to higher wavelengths over time, the measured absorbance at 522 nm increased rather fast during the incubation time. Thus the absorbance of each screening plate would have had to be measured after exactly the same time (plus/minus a few seconds), in this case 15 minutes. Assuring an absolutely constant incubation time would have increased the screening-effort considerably. Therefore the actual incubation time was set to 14-16 minutes, which was comparably easy to comply with. As a consequence, there can also be deviations between the ranges of absorbance values of the duplicate measurements of a single plate, as can be seen clearly in Figure 8.

The identification of potential hits was still comparably difficult since not even the difference between active and completely inactive variants was clear in all cases, due to the high fluctuation of the absorbance measurements and the fast background reaction. However, variants that seemingly showed lower absorbance values and thus higher BA conversions compared to the majority of variants could still be identified, especially upon comparison of the duplicate measurements. In the course of the screenings, note of such variants was taken and they were subsequently included in the rescreening experiment. In addition, several variants that seemingly had low BA concentrations in only one of the two duplicate measurements were still chosen for rescreening, just to be on the safe side and assure that no potential hit was missed. As an example, Figure 9 shows a plate containing such an unlikely potential hit that was still included in the rescreening.

In the case of some of the screened plates, the results of the duplicate measurement differed considerably, a respective example is given in Figure 10. Such deviations can result from several different effects. First of all, the screening of one plate included six different pipetting steps in total, each step contributing to an overall large pipetting error. Additionally, there was a tendency of the absorbance values of the wells at the edges of the MTPs to be lower compared to the inner wells. A likely cause of this effect is a slightly different growth behaviour of cells in the outer DWP-wells, for example because of faster heat transfer and increased evaporation. This might lead to higher cell concentrations or to harvest/induction at growth-phases differing slightly from the remaining cultures. This effect adds to another general feature, namely the faster decrease of BA in *E. coli* cell lysate

(independent of whether the lysate contained any HNLs or not) compared to pure buffer. Thus, the actual total protein concentration in the reaction mixtures also influenced the later absorbance measurement.

Furthermore, the mixing of the measurement mixture, which was an essential step, might not have been perfect in all cases and there was always the possibility that particles from the MTP were abraded during the process. Finally, the absorbance measurement itself added the intrinsic error of the spectrophotometer as well as it increased the error caused by pipetting because already slightly different liquid levels also influenced the measured absorbance according to the Beer-Lambert law.

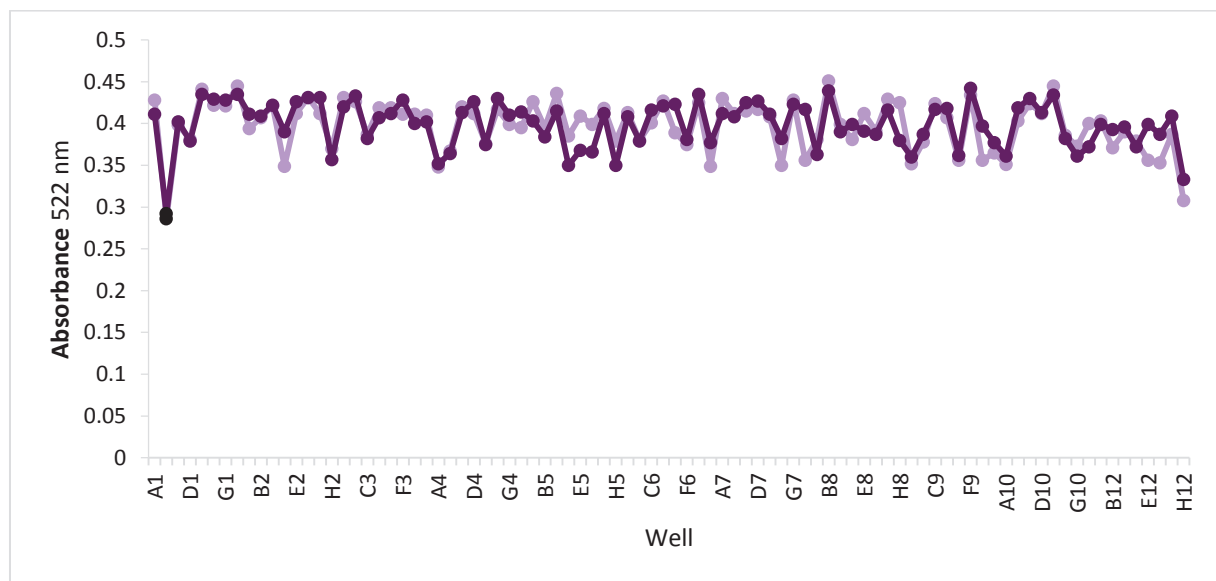


Figure 7. Screening result of DWP C-A2. The screening was performed according to the standard screening assay with NE. Each circle indicates a single variant, the measurement was carried out in duplicate (purple and lilac circles/lines). Black circles indicate a variant that was included in the rescreening and identified as hit “C-C2” (see part 5.4 for further details).

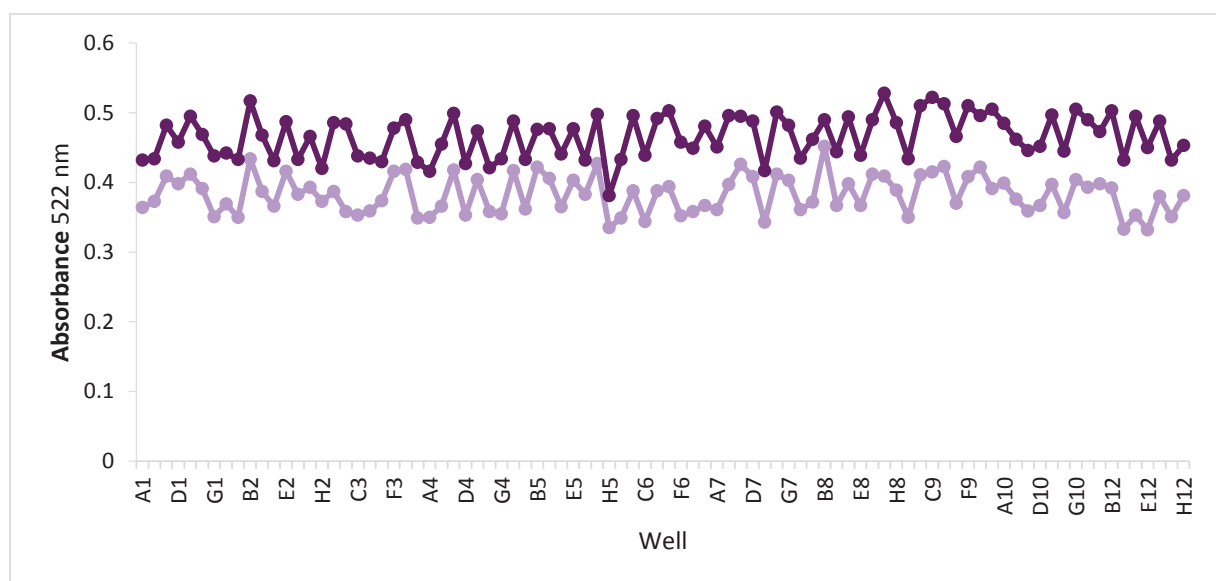


Figure 8. Screening result of DWP L-B2. The screening was performed according to the standard screening assay with NE. Each circle indicates a single variant, the measurement was carried out in duplicate (purple and lilac circles/lines).

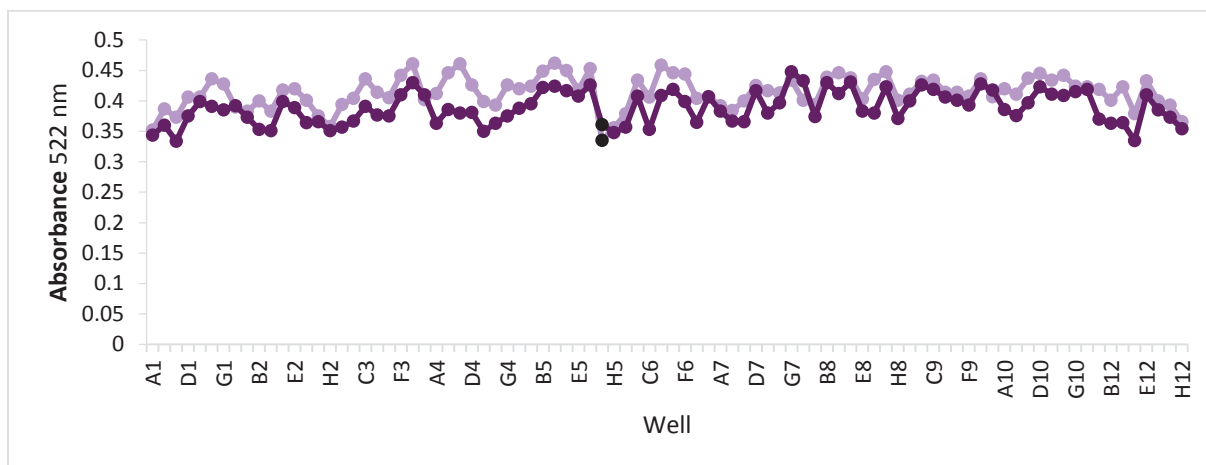


Figure 9. Screening result of DWP I-A2. The screening was performed according to the standard screening assay with NE. Each circle indicates a single variant, the measurement was carried out in duplicate (purple and lilac circles/lines). Black circles indicate a variant that was included in the rescreening and identified as hit “I-M10” (see part 5.4 for further details).

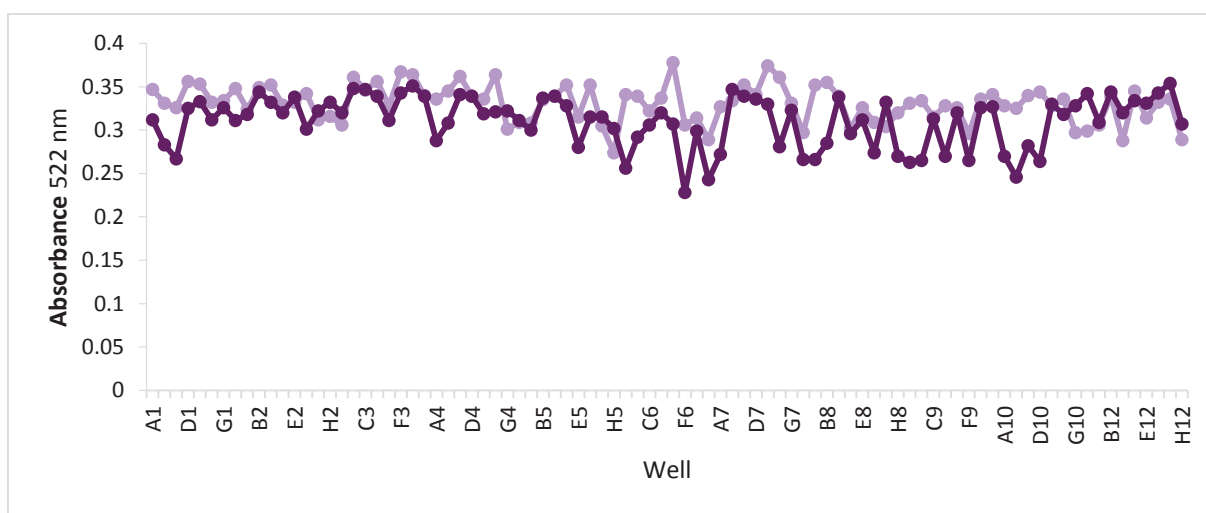


Figure 10. Screening result of DWP F-A2. The screening was performed according to the standard screening assay with NE. Each circle indicates a single variant, the measurement was carried out in duplicate (purple and lilac circles/lines).

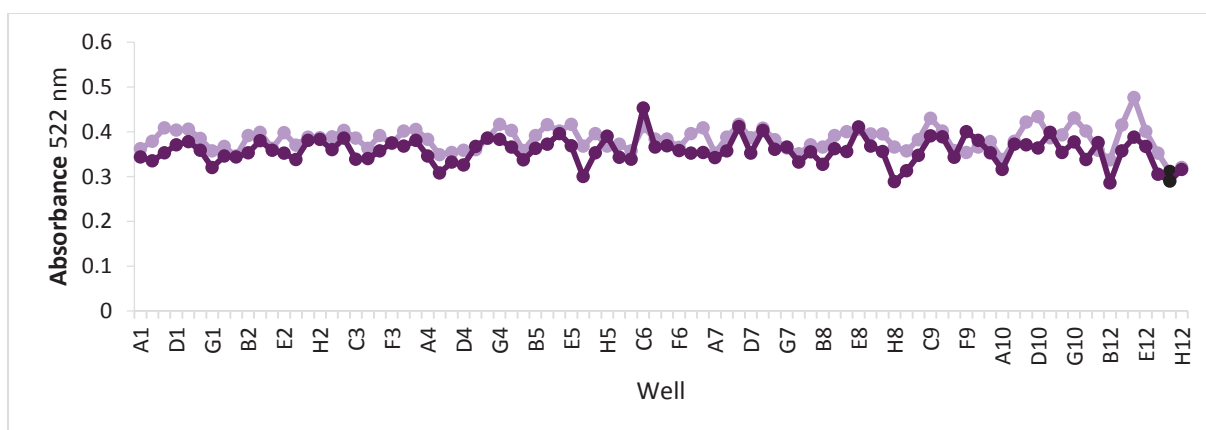


Figure 11. Screening result of DWP K-B1. The screening was performed according to the standard screening assay with NE. Each circle indicates a single variant, the measurement was carried out in duplicate (purple and lilac circles/lines). Black circles indicate a variant that was included in the rescreening and identified as hit “K-N23” (see part 5.4 for further details).

In total, 80 DWPs of the AcHNL-A40H library were screened for nitroaldolase activity. Among these ~7,000 screened clones, 80 potential hits were identified and included in the subsequent rescreening experiment. The black circles in Figure 7-Figure 11 indicate variants that were among the best hits and were thus cultivated in shake flasks and characterized by HPLC measurements later on.

5.4.AcHNL-A40H library rescreening

All the clones chosen for the rescreening experiment can be seen in Table 13, where their positions on the DWPs as well as on the 384 well MTPs are indicated, whereas the latter will be used as designation from this point onwards.

In order to avoid biased results, each clone was included two times on alternating positions on two different rescreening DWPs, and thus screened four times in total during the rescreening. The scheme according to which the clones were distributed on the four rescreening plates (designated RS 1, RS 2, RS 3 and RS 4) can be seen in Table 26-Table 29 in the appendix. The results of the rescreening are shown in the Figure 12-Figure 15.

Table 13. Summary of all the potential hits included in the rescreening of the AchNL-A40H library. Indicated are the positions on the 384 well MTP and on the corresponding DWP. PS functions as an abbreviation for “pre-screening plate”.

384 well MTP: letter-position	DWP: letter-code/position	384 well MTP: letter-position	DWP: letter-code/position
A-O11	A-A1/H6	I-G11	I-A1/D6
A-N14	A-B2/G7	I-M10	I-A2/G5
A-D24	A-B2/B12	I-J5	I-B1/E3
A-C2	A-A2/B1	J-B1	J-B1/A1
A-B5	A-B1/A3	J-M7	J-A1/G4
A-H5	A-B1/D3	J-A11	J-A1/A6
A-J9	A-B1/E5	J-P2	J-B2/H1
A-L8	A-B2/F4	K-C15	K-A1/B8
A-L16	A-B2/F8	K-A11	K-A1/A6
B-B8	B-B2/A4	K-N23	K-B1/G12
B-I15	B-A1/E8	K-C24	K-A2/B12
B-C23	B-A1/B12	L-D5	L-B1/B3
B-A8	B-A2/A4	L-P10	L-B2/H5
B-A16	B-A2/A8	M-I13	M-A1/E7
B-I24	B-A2/E12	M-P3	M-B1/H2
B-L23	B-B1/F12	O-K23	O-A1/F12
C-C2	C-A2/B1	O-G20	O-A2/D10
C-P23	C-B1/H12	O-I20	O-A2/E10
C-O24	C-A2/H12	O-F2	O-B2/C1
C-D7	C-B1/B4	O-D14	O-B2/B7
D-A2	D-A2/A1	P-E10	P-A2/C5
E-L3	E-B1/F2	P-G6	P-A2/D3
E-O19	E-A1/H10	P-J15	P-B1/E8
E-M24	E-A2/G12	P-P3	P-B1/H2
E-P19	E-B1/H10	Q-B17	Q-B1/A9
F-H10	F-B2/D5	Q-C14	Q-A2/B7
F-L23	F-B1/F12	Q-M1	Q-A1/G1
F-L1	F-B1/F1	Q-M4	Q-A2/G2
G-O2	G-A2/H1	Q-D14	Q-B2/B7
G-P3	G-B1/H2	X1-E3	X1-A1/C2
G-P12	G-B2/H6	X2-I12	X2-A2/E6
G-I1	G-A1/E1	X3-A2	X3-A2/A1
H-M7	H-A1/G4	X4-A18	X4-A2/A9
H-M15	H-A1/G8	X4-I9	X4-A1/E5
H-A12	H-A2/A6	X4-E24	X4-A2/C12
H-J7	H-B1/E4	X5-I8	X5-A2/E4
H-P3	H-B1/H2	X5-A10	X5-A2/A5
H-P14	H-B2/H7	X5-P1	X5-B1/H1
I-D20	I-B2/B10	X6-A9	X6-A1/A5
I-P8	I-B2/H4	-	PS D7

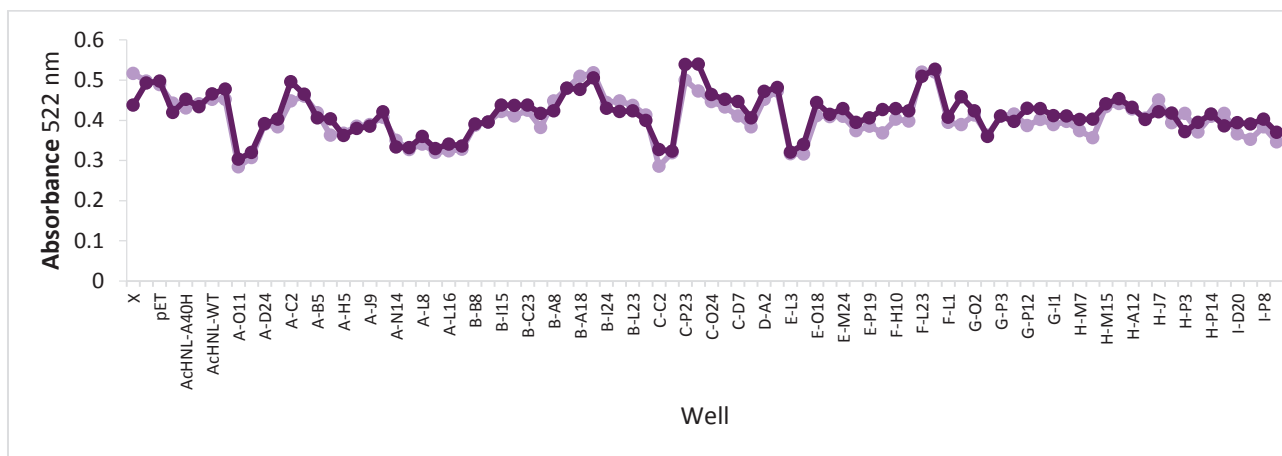


Figure 12. Screening result of rescreening plate 1. The screening was performed according to the standard screening assay with NE. Each circle indicates a single variant, but the potential hits were positioned in pairs of two and the measurement was carried out in duplicate (purple and lilac circles/lines).

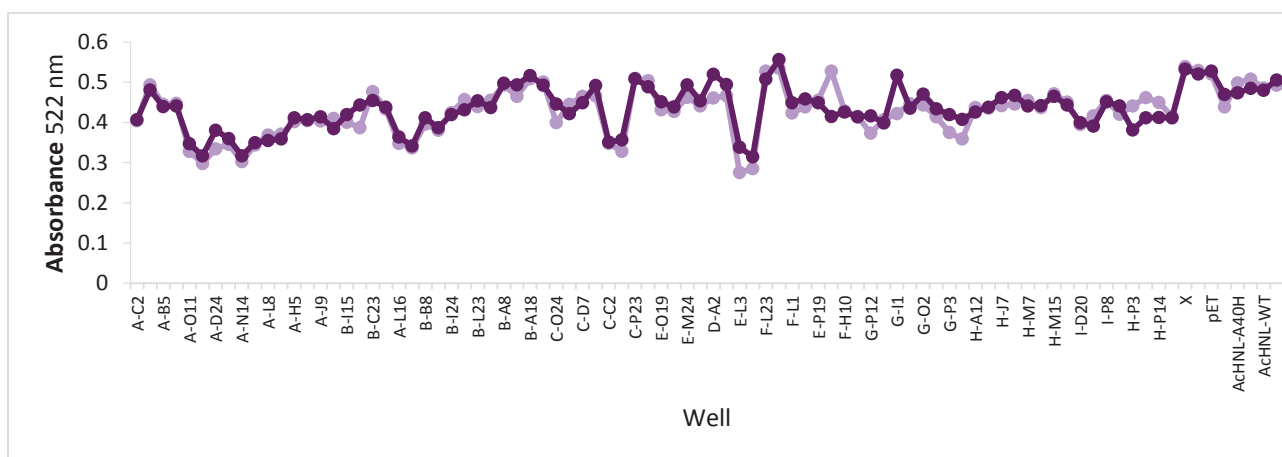


Figure 13. Screening result of rescreening plate 2. The screening was performed according to the standard screening assay with NE. Each circle indicates a single variant, but the potential hits were positioned in pairs of two and the measurement was carried out in duplicate (purple and lilac circles/lines).

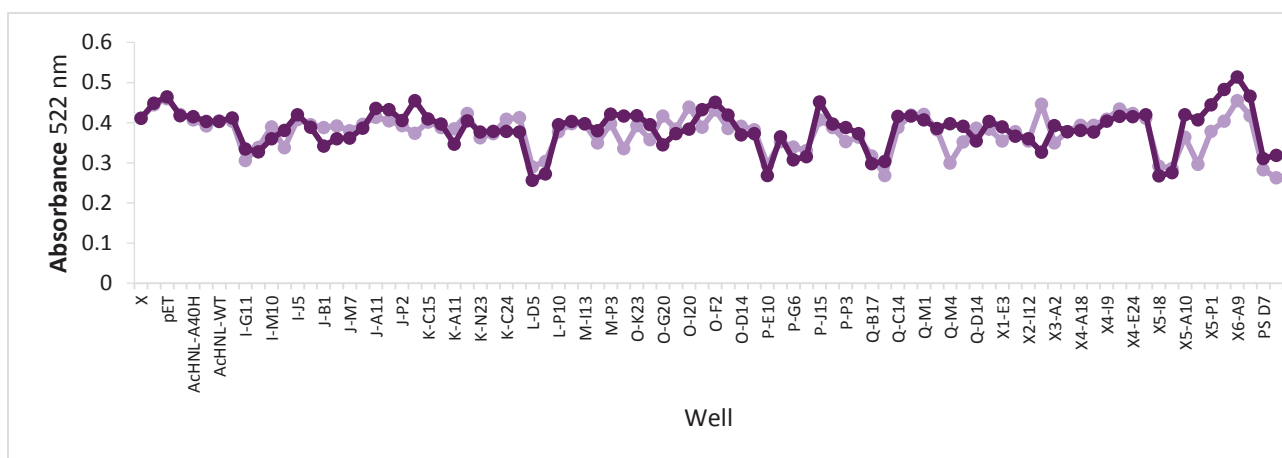


Figure 14. Screening result of rescreening plate 3. The screening was performed according to the standard screening assay with NE. Each circle indicates a single variant, but the potential hits were positioned in pairs of two and the measurement was carried out in duplicate (purple and lilac circles/lines). PS functions as an abbreviation for “prescreening plate”.

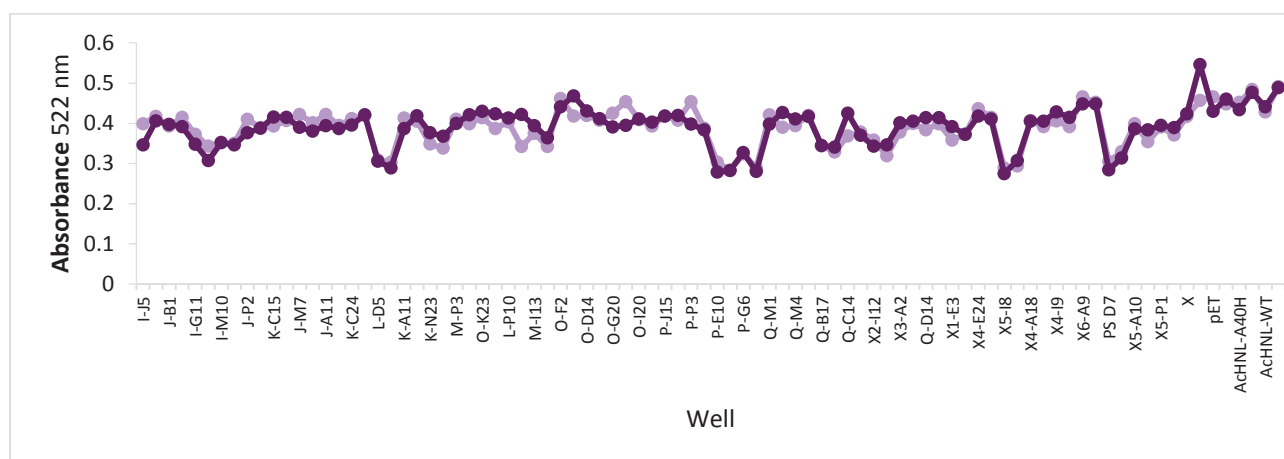


Figure 15. Screening result of rescreening plate 4. The screening was performed according to the standard screening assay with NE. Each circle indicates a single variant, but the potential hits were positioned in pairs of two and the measurement was carried out in duplicate (purple and lilac circles/lines). PS functions as an abbreviation for “prescreening plate”.

The 16 most active clones identified during the rescreening were chosen for sequencing of the respective HNL-gene. The thus found nucleotide and amino acid exchanges are summarized in Table 14. Of these variants, six did not comprise any unique amino acid exchanges. Those variants were not further characterized or cultivated: There were five variants harbouring the AA exchange T50S, of those, only the variant L-D5 was used for further experiments. Two variants (A-N14/P-G6) harboured the AA exchange T50A in addition to a stop codon mutation, leading to the additional translation of a 14 AA long C-terminal sequence that included the His-Tag present on the pET-26b(+) vector. These two variants were not further characterized since the stop codon mutation did not seem to increase the conversion any further compared to the single T50A mutation as present, for example, in variant C-C2.

The ten unique variants, underlined in Table 14, were then cultivated in shake flasks in order to produce larger amounts of cell free enzyme lysate that thereafter could be used for HPLC measurements. Additionally, these lysates were used for experiments where the screening assay with higher statistical significance was applied in order to estimate the enzyme activity and identify the best variants.

Table 14. Summary of the nucleotide mutations and amino acid exchanges found in the AchNL-A40H library’s best hits’ sequences. Variants chosen for further characterization are underlined.

designation	nucleotide mutation	codon mutation	amino acid exchanged
<u>A-O11</u>	A148→G148	ACC→GCC	T50→A50
	A304→T304	ACC→TCC	T102→S102
	A330→T330	CAA→CAT	Q110→H110
	A348→T348	AAA→AAT	K116→N116

<u>A-L16</u>	C29→T29	CCG→CTG	P10→L10
	A148→G148	ACC→GCC	T50→A50
	A347→G347	AAA→AGA	K116→R116
<u>C-C2</u>	A148→G148	ACC→GCC	T50→A50
<u>I-D20</u>	C63→T63	GGC→GGT	V23→I23
	G67→A67	GTT→ATT	V25→I25
<u>I-G11</u>	A148→G148	ACC→GCC	T50→A50
	A329→T329	CAA→CTA	Q110→L110
	T72→A72	CGT→CGA	silent
	A105→G105	GCA→GCG	silent
<u>I-M10</u>	G73→A73	GTT→ATT	V25→I25
	T291→A291	GGT→GGA	silent
P-E10, X5-I8	C149→G149	ACC→AGC	T50→S50
E-L3	A148→T148	ACC→TCC	T50→S50
PS-D7	A148→T148	ACC→TCC	T50→S50
	T39→C39	CGT→CGC	silent
<u>L-D5</u>	A148→T148	ACC→TCC	T50→S50
	C192→T192	ACC→ACT	silent
	A384→G348	AA→AAG	silent
<u>K-N23</u>	G88→A88	GCA→ACA	A30→T30
	A356→G356	GAA→GGA	E119→G119
	A377→G377	GAT→GGT	D126→G126
	T207→C207	CGT→CGC	silent
<u>Q-B17</u>	A64→G64	ACC→GCC	T22→A22
	A148→G148	ACC→GCC	T50→A50
	T72→A72	CGT→CGA	silent
	T231→G231	GTT→GTG	silent
	T243→G234	CGT→CGG	silent
	A384→G384	GAA→GAG	silent
<u>X2-I12</u>	A148→G148	ACC→GCC	T50→A50
	C227→A227	CCG→CAG	P76→Q76
A-N14, P-G6	A148→G148	ACC→GCC	T50→A50;
	T394→C394	TAA→CAA	Stop codon mutation, AA positions 132-145 added: QKLAAALEHHHHHH

Upon examination of the rescreening results and comparison to the respective AA sequences, there is one position that has not only the strongest influence on the conversion but has also been exchanged in 13 of the 16 sequenced hits. This AA is the threonine at position 50 that has been either exchanged for alanine or serine in all 13 variants. In addition to the table above, the exchanged amino acid positions can also be seen in Figure 16, which shows a modelled structure of the AcHNL where the positions exchanged in at least one of the hits are highlighted.

Note that for the exchange of threonine to alanine the replaced codon was always ACC→GCC. This is the only way to change an ACC codon to any of the codons encoding alanine while mutating only one nucleotide position. On the other hand, the codon encoding the new serine was two times AGC and three times TCC, which indicates that there was no strong bias in the library that could have led to an increased amount of certain mutations and thus a lower variability of the gene sequences present in the library. Such a bias could have been generated if the transformation of *E. coli* Top 10 F' cells with the nicked plasmids after the QuikChange PCR had not worked efficiently enough and thus only fewer unique sequences had actually been amplified during cell growth.

However, the inherent tendency of the *Taq* polymerase to preferably exchange A or T bases is reflected quite clearly by the results, of the overall 40 nucleotides that were exchanged, 77 % were either A or T.

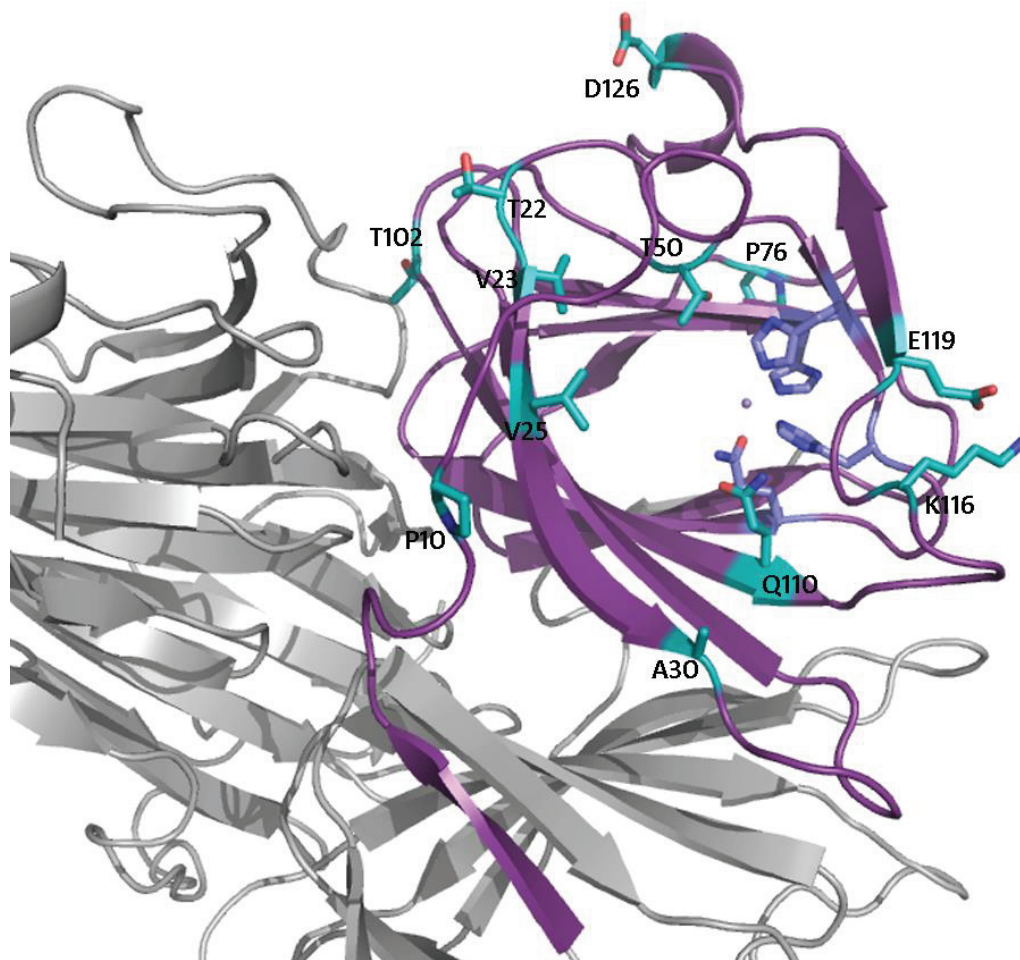


Figure 16. Modelled structure of one monomer of the ACHNL (purple) with the metal-binding site and the manganese ion highlighted in lilac and the amino acid positions that were exchanged in at least one of the hit-variants highlighted in turquoise. The Figure was prepared using the PyMOL Molecular Graphics System and is based on a modelled structure of ACHNL which was computed on basis of the GtHNL crystal structure and the crystal structure of the GtHNL-A40H/V42T/Q110H variant (PDB entry 4bif and 4uxa, respectively).

One way or another, the position 50 seemed to be very important for enzyme activity. Thus, screening of a site saturation mutagenesis library of this position could have yielded other variants with increased BA conversion. The preparation and screening of this ACHNL-A40H/T50X library is described in part 5.6. In addition, a variant ACHNL-T50A that did not harbour the A40H AA exchange was prepared using site specific mutagenesis. Characterization of this variant might have produced valuable information regarding the effect of the T50A AA exchange on the enzyme activity. The preparation of the ACHNL-T50A variant is described in the following chapter and the variant was thereafter included in the final characterization experiment (see part 5.13).

5.5. Preparation and cultivation of ACHNL-T50A and control variants

The variant ACHNL-T50A was prepared applying QuikChange PCR. In order to find variants that definitely contained plasmids with only a single integration of the mutated gene, 15 variants were

examined with colony PCR and agarose gel electrophoresis. Subsequently, four plasmids containing a single copy of the AcHNL-T50A encoding gene were sent for sequencing. The results showed that all four variants harboured the correct sequence. Thereafter, one of these clones was cultivated in shake flasks in addition to the controls that later were used for the HPLC measurements. After cell harvest and disruption, the total protein content of the cell free lysates was determined applying the Bradford protein assay. The results are summarized in Table 15.

Table 15. Total protein concentrations of the cell free lysates of the controls and the variant AcHNL-T50A.

variant	protein concentration lysate [mg/mL]
AcHNL-A40H	13.63
GtHNL-A40R	14.315
pET-26b(+)	10.22
AcHNL-T50A	9.14

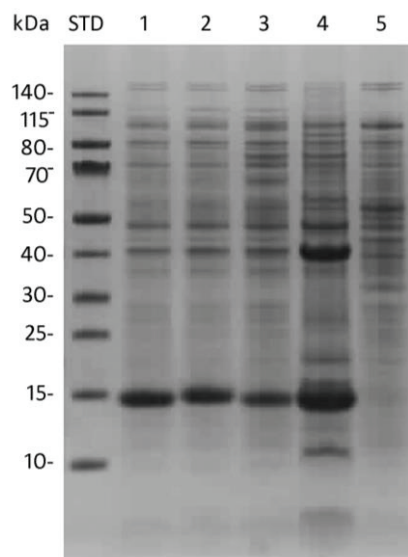


Figure 17. SDS-PAGE gel of cell free lysates (~5 µg total protein per sample) and corresponding pellet fraction of AcHNL-T50A variant and controls. MW AcHNL monomer: ~14.1 kDa, GtHNL monomer: ~15 kDa. Lanes: AcHNL-A40H lysate (1), GtHNL-A40R lysate (2), AcHNL-T50A lysate (3), AcHNL-T50A pellet (4), pET-26b lysate (5); STD: PageRuler Prestained Protein Ladder.

As the SDS-PAGE that was prepared for the variant AcHNL-T50A and the controls shows (see Figure 17), this variant mostly formed inclusion bodies. This was also evident upon examination of the cell pellet yielded after cell disruption. The pellet was larger with a considerable amount of white precipitate (the inclusion bodies) on top of the remaining undisrupted cells. Therefore, it can be concluded that there was a considerable stabilizing interaction between the positions H40 and A50 in the variant AcHNL-A40H/T50A (and likely also the respective T50S variant). Thus, the variant AcHNL-T50A cannot be used effectively without the second AA exchange, which was also confirmed by HPLC measurements (see part 5.13).

5.6. Site saturation mutagenesis: *AcHNL-A40H/T50X*

The *AcHNL-A40H/T50X* library was prepared using QuikChange PCR prior to colony PCR and agarose gel electrophoresis. The agarose gel was used to control the library's quality with regard to the length of the integrated genes and thus the number of integration events. As Figure 18 shows, 13 of the 16 chosen clones show at least weak bands at the right DNA length (approx. 600 bp), while two lanes (lane 7 and lane 16) do not show any band at all, most likely due to application of too much or too little cell material for the colony PCR. Only one clone produced a band that corresponded to a longer DNA fragment (lane 1), maybe originating from some kind of recombination event.

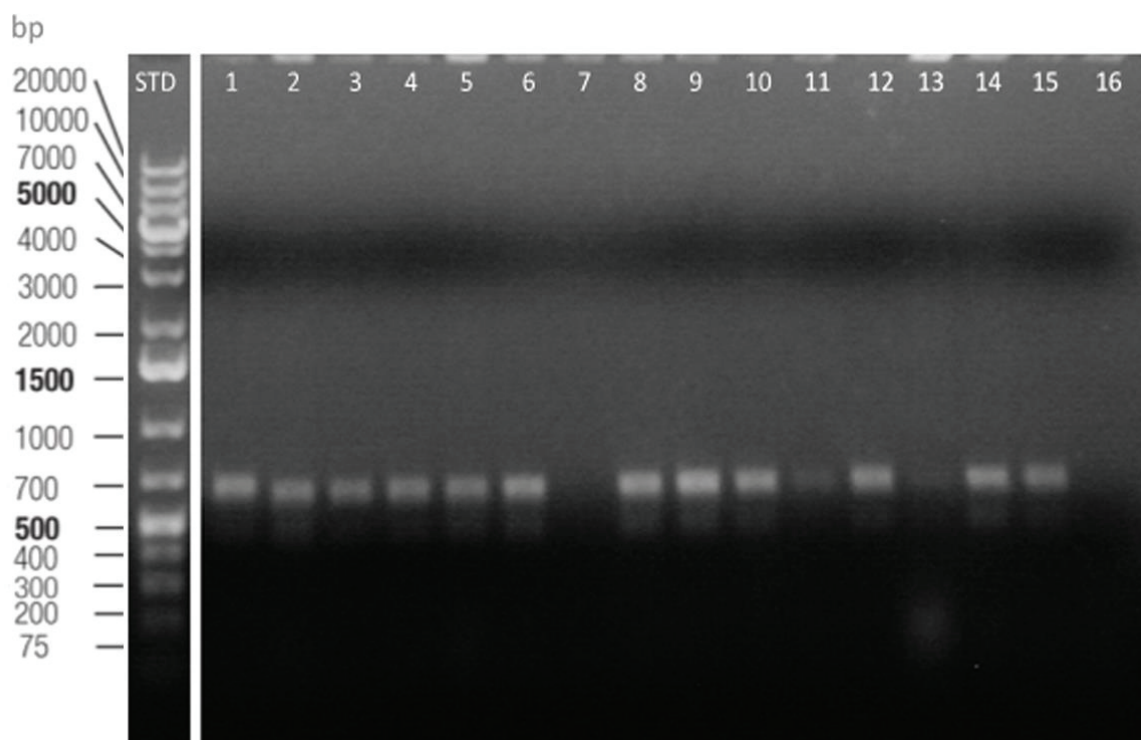


Figure 18. Agarose gel electrophoresis of 16 colony PCR samples of clones of the *AcHNL-A40H/T50X* library. 1-16 correspond to the 16 colony PCR mixtures; STD: O'GeneRuler 1kb Plus DNA Ladder, ready-to-use; Electrophoresis conditions: 1 % agarose, 120 V, 60 min.

Based on the gel, it seemed reasonable to continue the experiment with this library, thus ten arbitrarily chosen clones were used for sequencing of the integrated genes. Of those, five still contained the plasmid with the wild type gene, while the other five sequences comprised the following mutations:

ACC→ACG: silent mutation
 ACC→TCT: T→S
 ACC→CAG: T→Q
 ACC→CCG: T→P
 ACC→GCG: T→A

Even though 50 % of the clones still comprised the wild type sequence, the other 50 % all comprised different codons, thus this library was used for screening.

After the library had been transformed into *E. coli* BL21 (DE3) Gold cells, 440 clones were distributed on five DWPs for cultivation and screening. The high amount of clones was necessary in order to assure that at least one variant with every possible amino acid at position 50 was included on the screening plates, considering that 50 % of the clones still carried the wild type gene. The screening results of two of the five plates can be seen in Figure 19 and Figure 20.

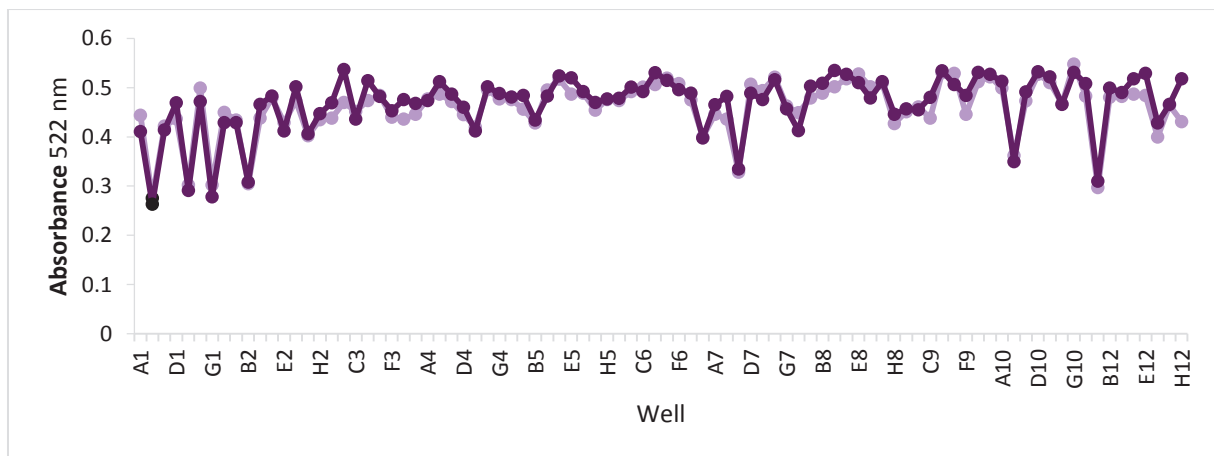


Figure 19. Screening result of AchNL-A40H/T50X plate 3. The screening was performed according to the standard screening assay with NE. Each circle indicates a single variant, the measurement was carried out in duplicate (purple and lilac circles/lines). Black circles indicate a variant that was identified as hit and chosen for sequencing.

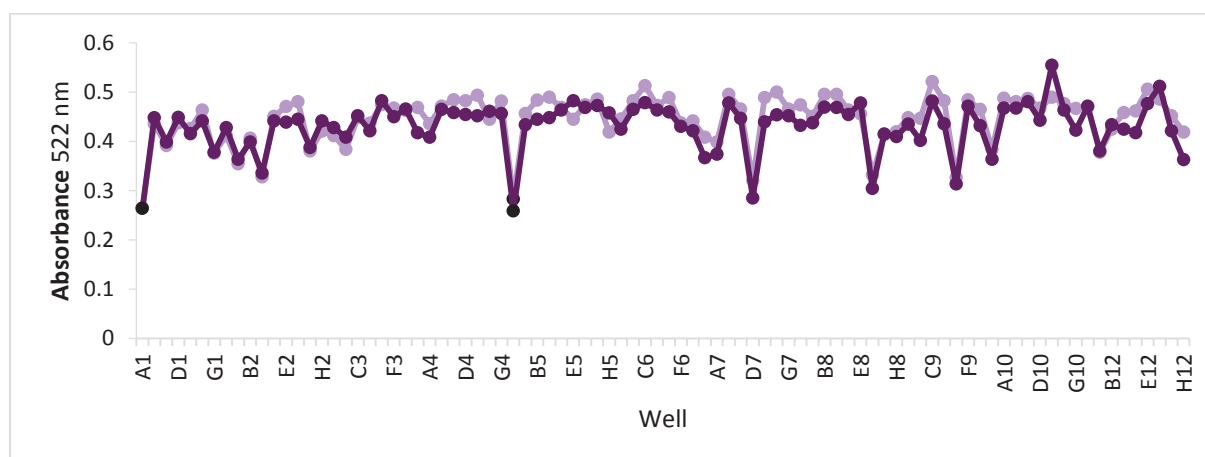


Figure 20. Screening result of AcHNL-A40H/T50X plate 5. The screening was performed according to the standard screening assay with NE. Each circle indicates a single variant, the measurement was carried out in duplicate (purple and lilac circles/lines). Black circles indicate a variant that was identified as hit and chosen for sequencing.

The screening results of the AcHNL-A40H/T50X library show that there are several variants with increased BA conversion compared to the parent variant. However, upon comparison of these results to the rescreening results of the random mutagenesis library (see Figure 12-Figure 15) no variant seems to yield a significantly higher conversion than the variants AcHNL-A40H/T50A or T50S that had already been identified. Nevertheless, ten of the most promising variants were chosen for sequencing of the corresponding gene, the results are summarized in Table 16.

Table 16. Summary of the identified nucleotide mutations and amino acid exchanges of the most promising variants of the AcHNL-A40H/T50X library.

designation: T50X(Plate-No.) position	codon mutation	amino acid exchanged
T50X(1) C4	ACC→AGT	T50→S50
T50X(2) E1	ACC→GCT	T50→A50
T50X(2) G7	ACC→TCG	T50→S50
T50X(2) E2	ACC→GCG	T50→A50
T50X(3) B1	ACC→GCG	T50→A50
T50X(4) B1	ACC→GCG	T50→A50
T50X(4) E4	ACC→AGT	T50→S50
T50X(4) E8	ACC→AGT	T50→S50
T50X(5) A1	ACC→TCG	T50→S50
T50X(5) H4	ACC→TCT	T50→S50

As the sequencing results show, all the chosen variants comprise either the AA exchange T50A or T50S, though with varying codons. On the one hand, this indicates that the activities of these two variants are approximately equal. On the other hand, the only useful information that could be gained by screening of this library was that the incorporation of one of these two amino acids does indeed lead to the highest positive effect on the conversion. Therefore, further varying the amino acids at position 50 did not make much sense and a clear distinction which of the two is actually more beneficial can only be made upon evaluation of the HPLC results (see part 5.13). Thus, the variants AcHNL-A40H/T50A and AcHNL-A40H/T50S as well as the other hits of the AcHNL-A40H random mutagenesis library were used for the following cultivation and the characterization experiments.

5.7.Cultivation of AcHNL-A40H mutants

The ten most promising AcHNL-A40H mutants (see part 5.4) were cultivated in shaking flasks in order to produce larger amounts of cell free lysate. After cell harvest and disruption, the cleared lysates were analyzed through determination of the total protein concentration (see Table 17) and SDS-PAGE.

Table 17. Total protein concentrations of the cell free lysates of the hits identified during the screening of the AcHNL-A40H library.

variant	protein concentration [mg/mL]
AcHNL-A40H/T50A/T102S/Q110H/K116N	9.26
AcHNL-P10L/A40H/T50A/K116R	4.91
AcHNL-A40H/T50A	8.56
AcHNL-V23I/V25I/A40H	11.73
AcHNL-A40H/T50A/Q110L	6.15
AcHNL-V25I/A40H	13.64
AcHNL-A40H/T50S	12.14
AcHNL-A30T/A40H/E119G/D126G	10.87
AcHNL-T22A/A40H/T50A	10.07
AcHNL-A40H/T50A/P76Q	7.89

The subsequent SDS-PAGE analysis yielded the gels shown in Figure 21. The intensity of the bands positioned at approx. 14 kDa shows that most mutants seem to express reasonable amounts of soluble enzyme. However, the variant AcHNL-P10L/A40H/T50A/K116R seemingly forms a high amount of inclusion bodies, the pellet fraction (4) shows an intense band at ~14 kDa on the SDS PAGE. Additionally, the lysate fractions of the variants AcHNL-P10L/A40H/T50A/K116R (3), AcHNL-

A40H/T50A/Q110L (9) and AcHNL-A40H/T50A/P76Q (19) show smaller bands at the respective size and, as Table 17 shows, the total protein content of these lysates is also smaller. This indicates that these three variants actually form inclusion bodies, which is also indicated by the fact that the pellets that remained after cell disruption were comparably large and contained a considerable amount of white precipitate, most likely the denatured enzyme. A possible explanation why the inclusion bodies of variants AcHNL-A40H/T50A/Q110L and AcHNL-A40H/T50A/P76Q are not visible on the SDS gel is that the analysed pellet-fraction only contained material from parts close to the pellet surface and no material from the inclusion bodies that were positioned in between the undisrupted cells and the cell fragments. Only of the AcHNL-P10L/A40H/T50A/K116R variant's pellet the white precipitate was included on purpose in order to test it for inclusion bodies.

The fact that the lysates of these three variants actually contained lesser amounts of enzyme compared to the total protein content had to be considered during further characterization of the variants. Furthermore, these variants are not suited well for industrial application since they cannot be produced in as high amounts as other variants.

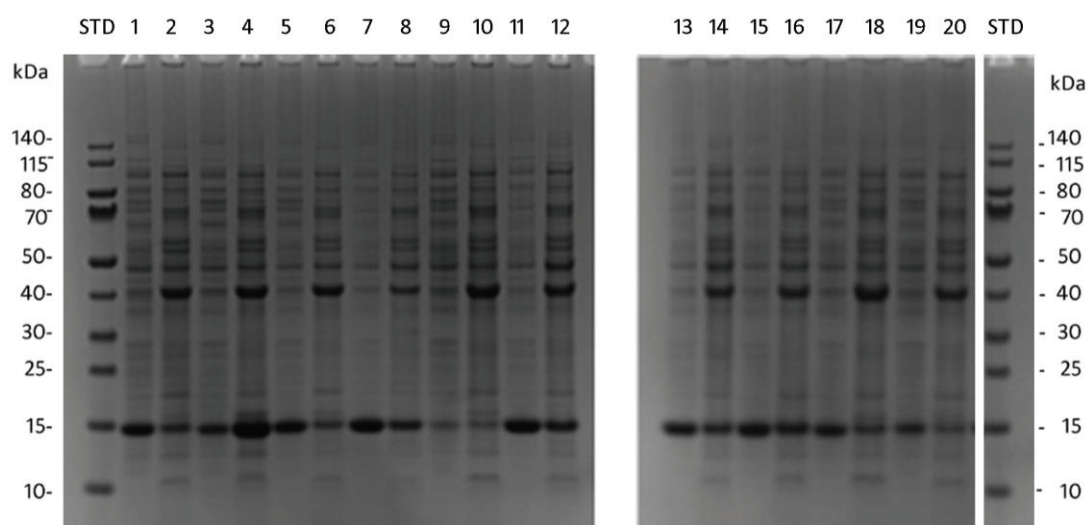


Figure 21. SDS-PAGE gel of cell free lysates (~5 µg total protein per sample) and corresponding pellet fractions of the AcHNL-A40H library's hits. MW AcHNL monomer: ~14.1 kDa. Left gel: AcHNL-A40H/T50A/T102S/Q110H/K116N lysate (1), AcHNL-A40H/T50A/T102S/Q110H/K116N pellet (2), AcHNL-P10L/A40H/T50A/K116R lysate (3), AcHNL-P10L/A40H/T50A/K116R pellet (4), AcHNL-A40H/T50A lysate (5), AcHNL-A40H/T50A pellet (6), AcHNL-V23I/V25I/A40H lysate (7), AcHNL-V23I/V25I/A40H pellet (8), AcHNL-A40H/T50A/Q110L lysate (9), AcHNL-A40H/T50A/Q110L pellet (10), AcHNL-V25I/A40H lysate (11), AcHNL-V25I/A40H pellet (12); Right gel: AcHNL-A40H/T50S lysate (13), AcHNL-A40H/T50S pellet (14), AcHNL-A30T/A40H/E119G/D126G lysate (15), AcHNL-A30T/A40H/E119G/D126G pellet (16), AcHNL-T22A/A40H/T50A lysate (17), AcHNL-T22A/A40H/T50A pellet (18), AcHNL-A40H/T50A/P76Q lysate (19), AcHNL-A40H/T50A/P76Q pellet (20); STD: PageRuler Prestained Protein Ladder.

The cell free lysates of all the hits were then used for conversion experiments in aqueous solution in order to find the most promising variants for subsequent HPLC measurements. The results of these

experiments will be discussed together with the corresponding results for the *GtHNL-A40R* library in part 5.13.

5.8. *GtHNL-A40R* library screening

Since the mutation rates in all initial *GtHNL-A40R* random mutagenesis libraries were not satisfactory, a new library was prepared using 0.13 mM of Mn^{2+} for the epPCR. Of this new library, ten genes were sent for sequencing to evaluate the average mutation rate. The numbers of mutations in the ten sequences can be seen in Figure 22.

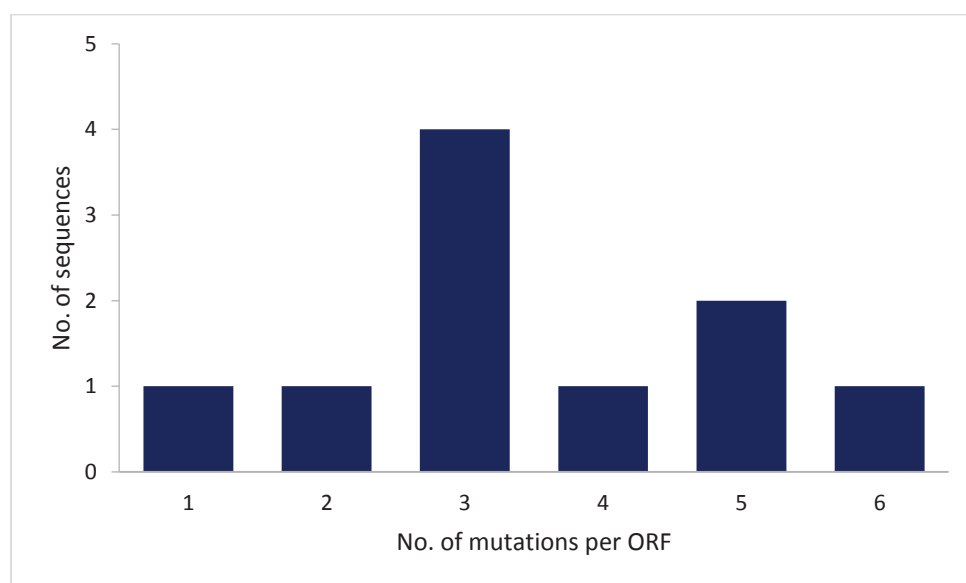


Figure 22. Analysis of nucleotide exchanges in the *GtHNL-A40R* coding region of ten arbitrarily chosen clones of the random mutagenesis library that was prepared with epPCR applying 0.13 mM Mn^{2+} .

The average mutation rate was 3.5 nucleotide exchanges per ORF, which is still rather high. However, the difference between the mutation rate of the *AcHNL-A40H* and this *GtHNL-A40R* library was only 0.4, and the *AcHNL-A40H* library had yielded several promising hits after all. Therefore, the *GtHNL-A40R* library prepared with 0.13 mM Mn was used for the screening.

Of this library, ~6,800 clones were screened in total and two examples of screened DWPs are shown in Figure 23 and Figure 24.

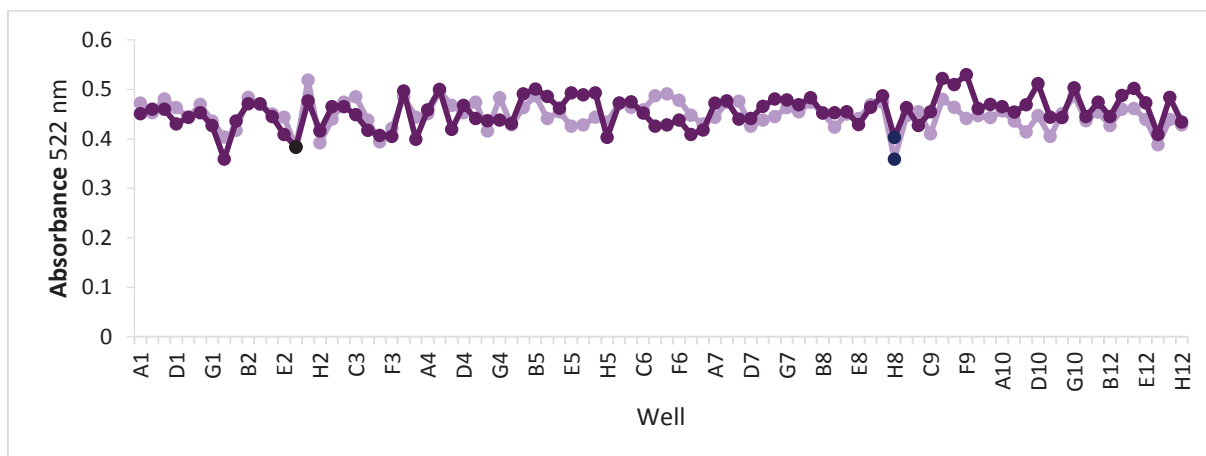


Figure 23. Screening result of DWP N-B1. The screening was performed according to the standard screening assay with NE. Each circle indicates a single variant, the measurement was carried out in duplicate (purple and lilac circles/lines). Position F2 (black circles) corresponds to a variant that was included in the rescreening and identified as hit “N-L3” (see part 5.9 for further details). The position H8 (blue circles) was also included in the rescreening but on the rescreening plates it did not achieve a higher BA conversion than the controls.

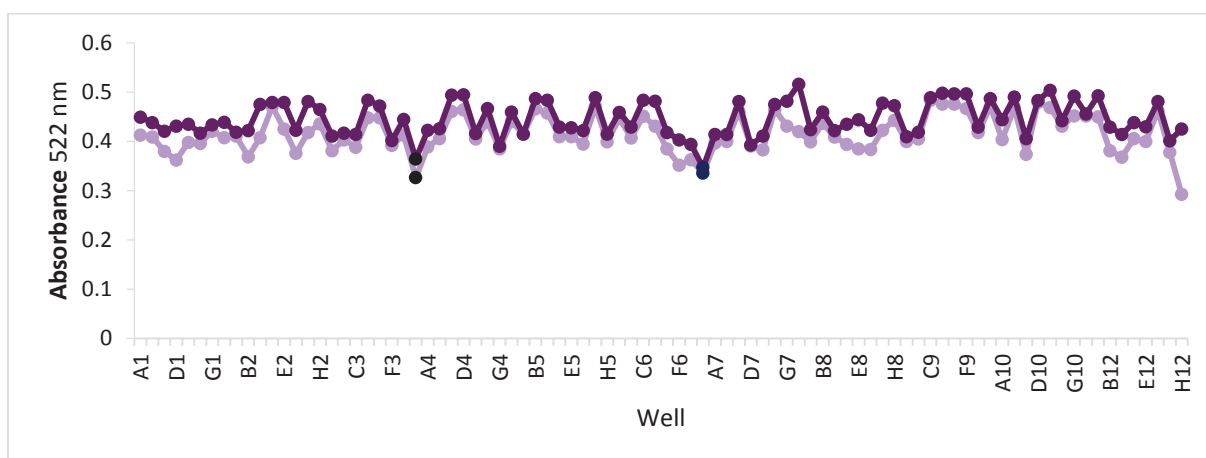


Figure 24. Screening result of DWP R-B2. The screening was performed according to the standard screening assay with NE. Each circle indicates a single variant, the measurement was carried out in duplicate (purple and lilac circles/lines). Position H3 (black circles) corresponds to a variant that was included in the rescreening and identified as hit “R-P6” (see part 5.9 for further details). The position H6 (blue circles) was also included in the rescreening but on the rescreening plates it did not achieve a higher BA conversion than the controls.

Compared to the AChNL-A40H library, this library did not yield as many significantly better variants during the screening, even though there were several positions that seemed likely to correspond to variants with a slightly higher activity. The variant marked as a potential hit on the DWP in Figure 23, position F2, for example, was later identified as the most active variant found in the GtHNL-A40R library. This, though, could hardly be deduced from the results of this first round of screening. The mostly indistinct results led to the selection of a large number of positions for rescreening, even though most of them seemed rather unlikely to harbour variants with increased activity.

5.9. *Gt*HNL-A40R library rescreening

Overall, 137 variants were chosen for rescreening, their positions on the 384 well MTPs and on the respective DWPs are summarized in Table 18.

Table 18. Summary of all the potential hits included in the rescreening of the *Gt*HNL-A40R library. Indicated are the positions on the 384 well MTP and on the corresponding DWP.

384-well MTP: letter-position	DWP: letter-code/position	384-well MTP: letter-position	DWP: letter-code/position
A-A1/E10	A-I19	L-B1/G6	L-N11
A-A1/D8	A-G15	L-B2/H1	L-P2
A-A1/H6	A-O11	L-B2/C9	L-F18
A-A2/A6	A-A12	L-B2/H6	L-P12
A-A2/E7	A-I14	M-A1/A1	M-A1
A-A2/D9	A-G18	M-A1/E3	M-I5
A-A2/F12	A-K24	M-A1/F6	M-K11
A-B1/E6	A-J11	M-A1/G4	M-M7
A-B2/A2	A-B4	M-A2/C9	M-E18
A-B2/E6	A-J12	M-B1/B4	M-D7
B-A2/B5	B-C10	M-B1/G9	M-N17
B-A2/H12	B-O24	M-B2/G3	M-N6
B-B1/G6	B-N11	N-A1/H6	N-O11
B-B1/E2	B-J3	N-A1/E5	N-I9
C-A1/D7	C-G13	N-B1/H8	N-P15
C-A1/H2	C-O3	N-B1/F2	N-L3
C-A2/E5	C-I10	O-A1/B4	O-C7
C-B1/G4	C-N7	O-A1/E3	O-I5
C-B1/H6	C-P11	O-A2/A5	O-A10
C-B2/D5	C-H10	O-B1/D2	O-H3
C-B2/B6	C-D12	O-B1/E3	O-J5
C-B2/H5	C-P10	O-B1/H8	O-P15
D-A2/H6	D-O12	O-B2/E5	O-J10
D-A2/B2	D-C4	P-A1/A6	P-A11
E-A2/D4	E-G8	P-A1/C9	P-E17
E-B1/E5	E-J9	P-A1/E2	P-I3
E-B1/E6	E-J11	P-A1/G6	P-M11
E-B2/A2	E-B4	P-A1/H12	P-O23
E-B2/H6	E-P12	P-A2/H12	P-O24
F-A1/H4	F-O7	P-B2/F9	P-L18
F-A1/D6	F-G11	P-B2/B8	P-D16
F-A2/F2	F-K4	P-B2/A2	P-B4
F-A2/C4	F-E8	Q-A1/D10	Q-G19
F-B1/C6	F-F11	Q-A1/H6	Q-O11
G-A2/E2	G-I4	Q-A1/A3	Q-A5
G-A2/D10	G-G20	Q-A2/E4	Q-I8
G-A2/B6	G-C12	Q-A2/A8	Q-A16
G-B1/A5	G-B9	Q-A2/F6	Q-K12
G-B2/A8	G-B16	Q-A2/G7	Q-M14
G-B2/G12	G-N24	Q-B1/B7	Q-D13
H-A1/E4	H-I7	Q-B1/A2	Q-B3
H-A2/B2	H-C4	Q-B1/G12	Q-N23
H-B1/A3	H-B5	Q-B2/H2	Q-P4
H-B1/C12	C-F23	Q-B2/G10	Q-N20
H-B2/H1	H-P2	Q-B2/E5	Q-J10
H-B2/H12	H-P24	R-A1/G3	R-M5
I-A1/H8	I-O15	R-A1/H4	R-O7
I-A1/A3	I-A5	R-A2/E3	R-I6
I-B1/F6	I-L11	R-B1/B5	R-BD9
I-B2/G10	I-N20	R-B1/A4	R-B7
I-B2/F4	I-L8	R-B2/H3	R-P6

I-B2/A3	I-B6	R-B2/H6	R-P12
J-A1/F12	J-K23	S-A1/E6	S-I11
J-A1/E5	J-I9	S-A1/C6	S-E11
J-A1/B8	J-C15	S-A2/D3	S-G6
J-A2/E4	J-I8	S-A2/B7	S-C14
J-A2/A6	J-A12	S-B1/F12	S-L23
J-B1/D2	J-H3	S-B1/B2	S-D3
J-B1/G7	J-N13	S-B1/E5	S-J9
K-A1/D4	K-G7	S-B2/C5	S-F10
K-A1/A3	K-A5	X2-A1/F6	X2-K11
K-A1/F5	K-K9	X6-A1/G7	X6-M13
K-B1/E4	K-J7	X6-A2/G4	X6-M8
K-B1/C9	K-F17	X6-A2/H1	X6-O2
K-B2/H9	KP18	X6-B1/E7	X6-J13
K-B2/A5	K-B10	X6-B1/G1	X6-N1
L-A1/F12	L-K23	X6-B2/A2	X6-B4
L-A1/F4	L-K7	X6-B2/B3	X6-D6
L-B1/E2	L-J3		

For the rescreening, each clone was included two times on alternating positions in two different rescreening DWPs, and thus screened four times in total. Only the variants on the RS 7 plate were included four times in a single plate. The scheme according to which the clones were distributed on the seven rescreening plates (designated RS 1-RS 7) can be seen in Table 30-Table 36 in the appendix. The results of the rescreening are shown in Figure 25-Figure 31.

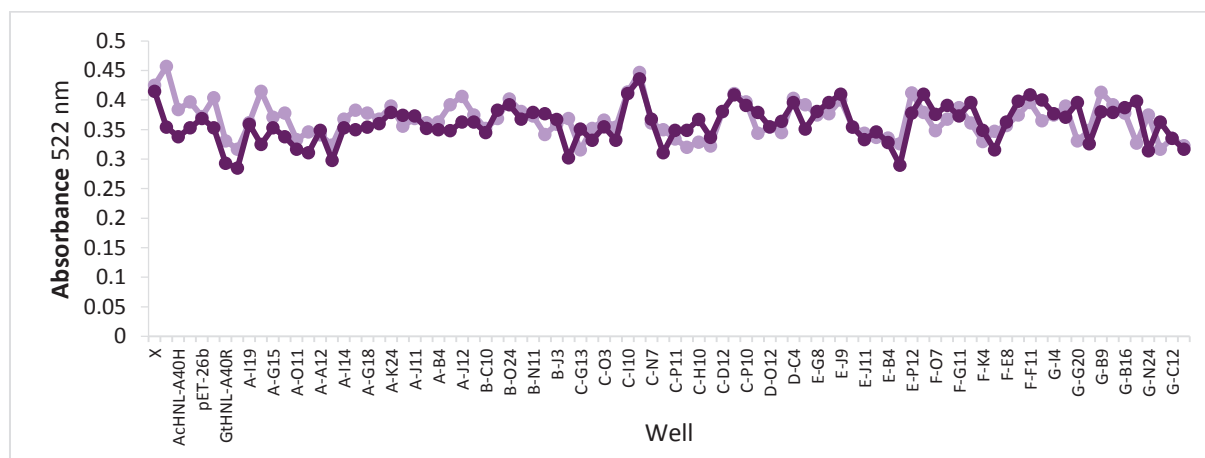


Figure 25. Screening result of *GtHNL-A40R* rescreening plate 1. The screening was performed according to the standard screening assay with NE. Each circle indicates a single variant, but the potential hits were positioned in pairs of two and the measurement was carried out in duplicate (purple and lilac circles/lines).

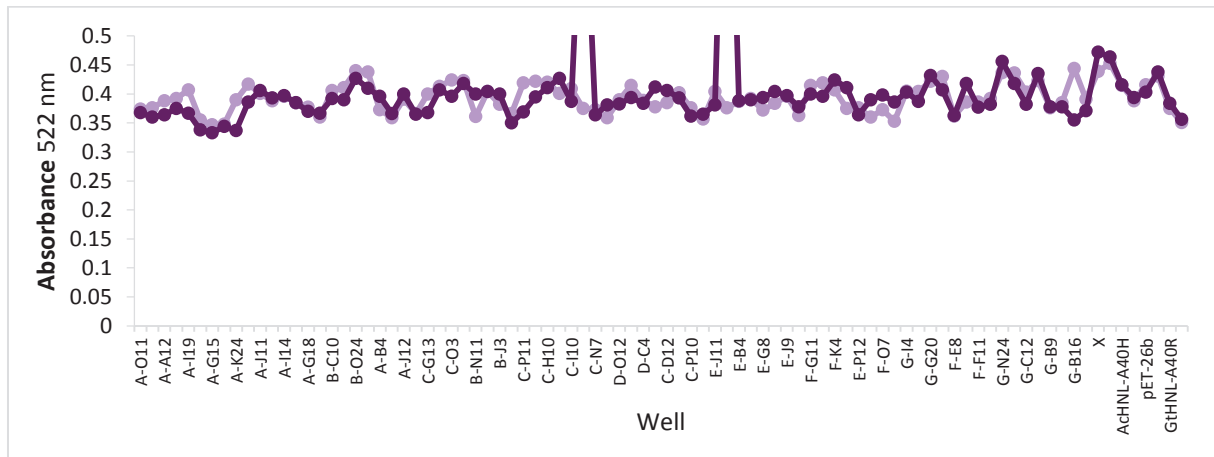


Figure 26. Screening result of *GtHNL-A40R* rescreening plate 2. The screening was performed according to the standard screening assay with NE. Each circle indicates a single variant, but the potential hits were positioned in pairs of two and the measurement was carried out in duplicate (purple and lilac circles/lines). Two measurement points of one of the duplicate data (purple) are outliers, most likely due to particles in the measured wells.

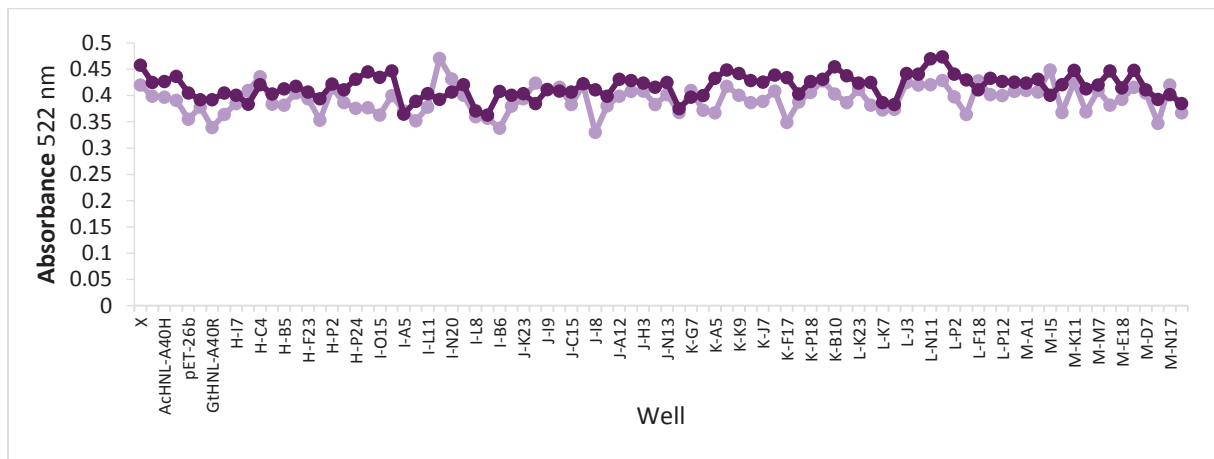


Figure 27. Screening result of *GtHNL-A40R* rescreening plate 3. The screening was performed according to the standard screening assay with NE. Each circle indicates a single variant, but the potential hits were positioned in pairs of two and the measurement was carried out in duplicate (purple and lilac circles/lines).

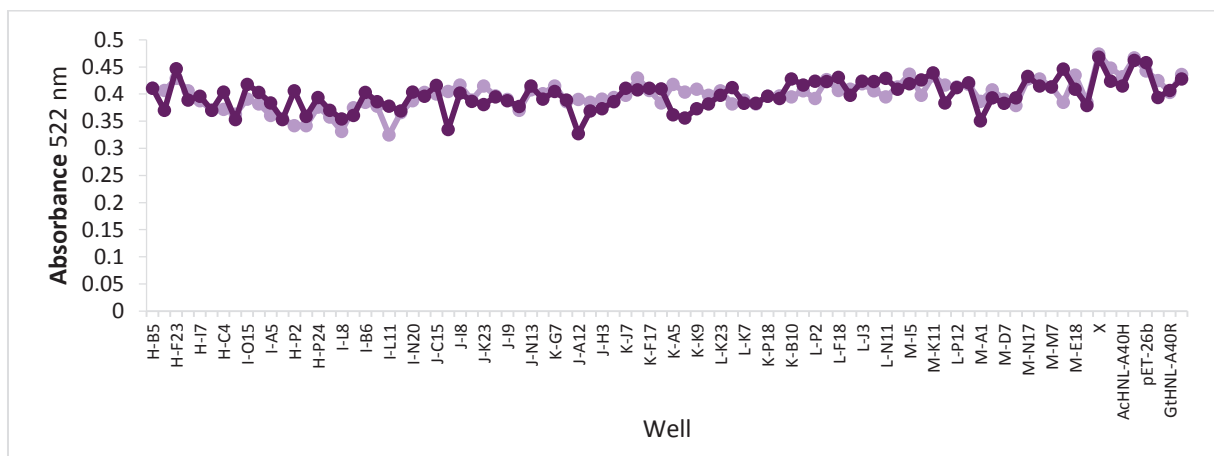


Figure 28. Screening result of *GtHNL-A40R* rescreening plate 4. The screening was performed according to the standard screening assay with NE. Each circle indicates a single variant, but the potential hits were positioned in pairs of two and the measurement was carried out in duplicate (purple and lilac circles/lines).

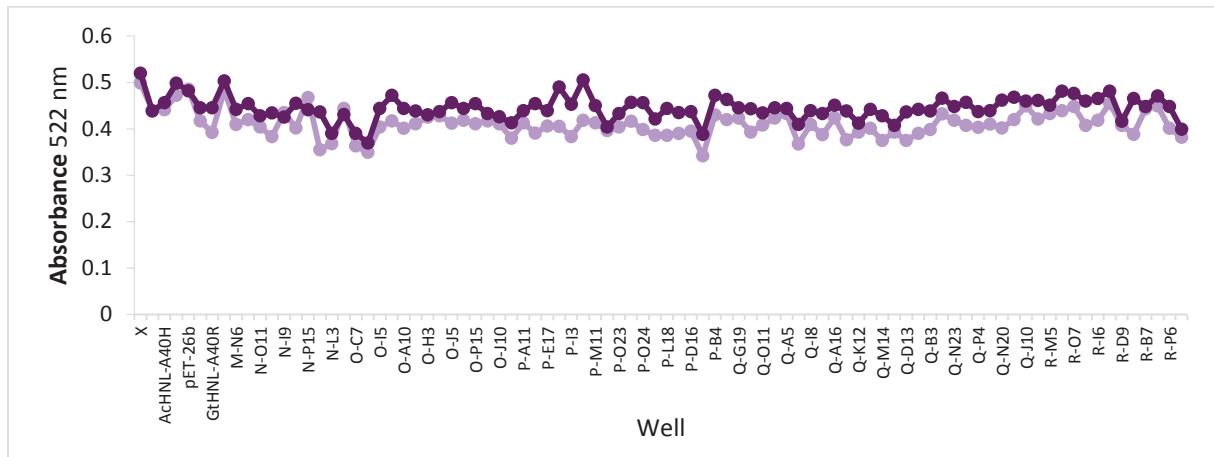


Figure 29. Screening result of *GtHNL-A40R* rescreening plate 5. The screening was performed according to the standard screening assay with NE. Each circle indicates a single variant, but the potential hits were positioned in pairs of two and the measurement was carried out in duplicate (purple and lilac circles/lines).

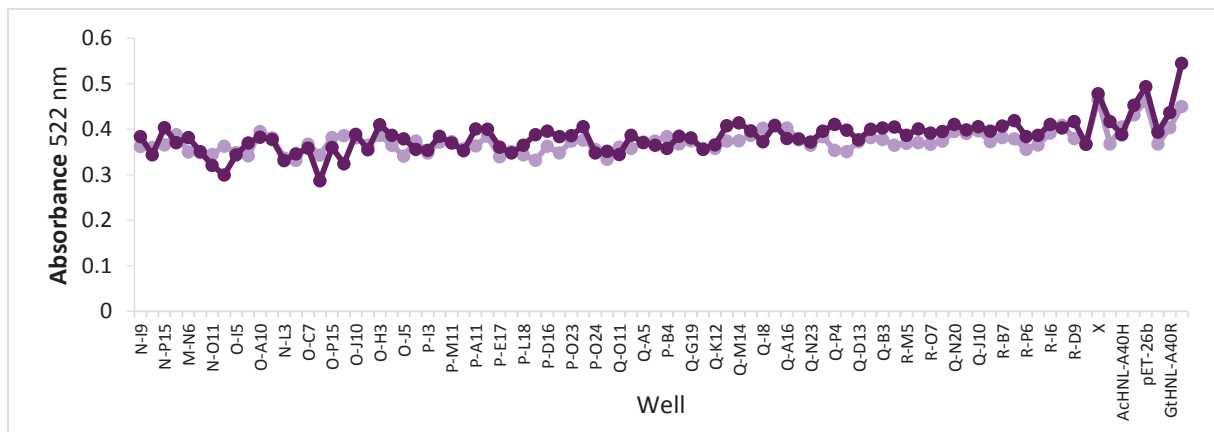


Figure 30. Screening result of *GtHNL-A40R* rescreening plate 6. The screening was performed according to the standard screening assay with NE. Each circle indicates a single variant, but the potential hits were positioned in pairs of two and the measurement was carried out in duplicate (purple and lilac circles/lines).

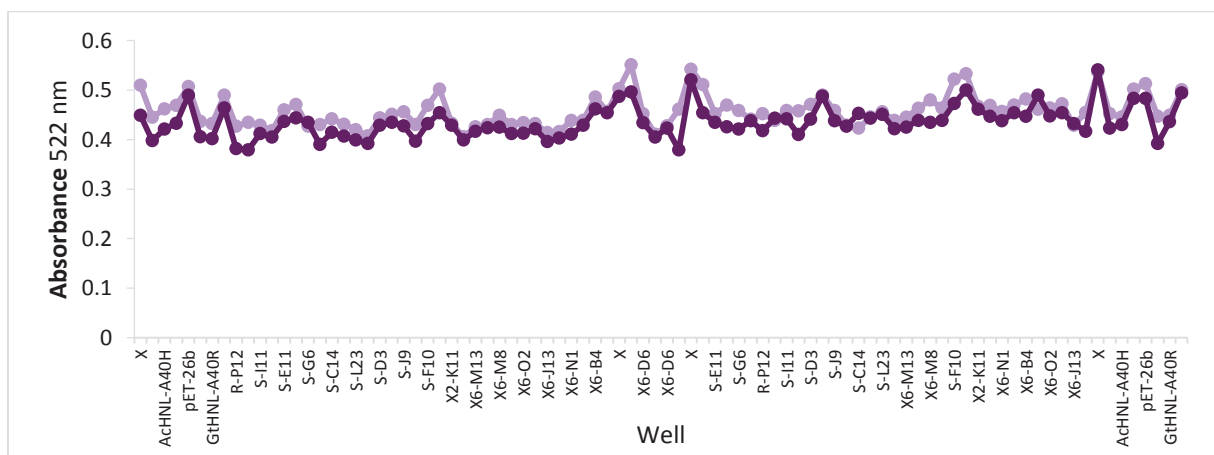


Figure 31. Screening result of *GtHNL-A40R* rescreening plate 7. The screening was performed according to the standard screening assay with NE. Each circle indicates a single variant, but the potential hits were positioned in pairs of two and the measurement was carried out in duplicate (purple and lilac circles/lines).

As it was already the case for the first screening data, the results of the rescreening are not very clear. It seems that there are no variants with significantly increased activity. Even so, the ten seemingly best clones were chosen for sequencing of the respective HNL gene. The sequencing results are summarized in Table 19.

Table 19. Summary of the nucleotide mutations and amino acid exchanges found in the *GtHNL-A40R* library's best hits' sequences. Variants chosen for further characterization are underlined.

designation	nucleotide mutation	codon mutation	amino acid exchanged
<u>I-L8</u>	A148→T148	ACC→TCC	T50→S50
	C60→A60	ACC→ACA	silent
	A195→G195	GCA→GCG	silent
	A210→T210	GCA→GCT	silent
<u>N-L3</u>	A329→T329	CAG→CTG	Q110→L110
	T390→G390	CGT→CGG	silent
<u>X6-J13</u>	T68→C68	GTT→GCT	V23→A23
<u>L-K7</u>	G295→A295	GCA→ACA	A99→T99
	A329→T329	CAG→CTG	Q110→L110
	A228→T228	GCA→GCT	silent
<u>R-P6</u>	G91→A91	GCA→ACA	A31→T31
	T199→A199	TGT→AGT	C67→S67
<u>P-L18</u>	A278→G278	AAA→AGA	K93→R93
	T15→C15	ATT→ATC	silent
	T258→A258	GTT→GTA	silent
A-O11	G168→A168	CGG→CGA	silent
I-A5	T204→C204	GGT→GGC	Silent
N-O11	T57→C57	TTT→TTC	silent
S-C14	-	-	wild type

In accordance with the indistinct nature of the library's screening results, 40 % of the variants chosen for sequencing comprise the wild type amino acid sequence. However, except for one variant, all of them carry silent mutations which might have led to higher expression levels and thus to the false identification as hits. Nevertheless, only the six variants with AA exchanges were cultivated in shake flasks in order to prepare cell free lysate for the final characterization. Figure 32 shows the 3D structure of the *GtHNL* where all the mutated positions are highlighted.

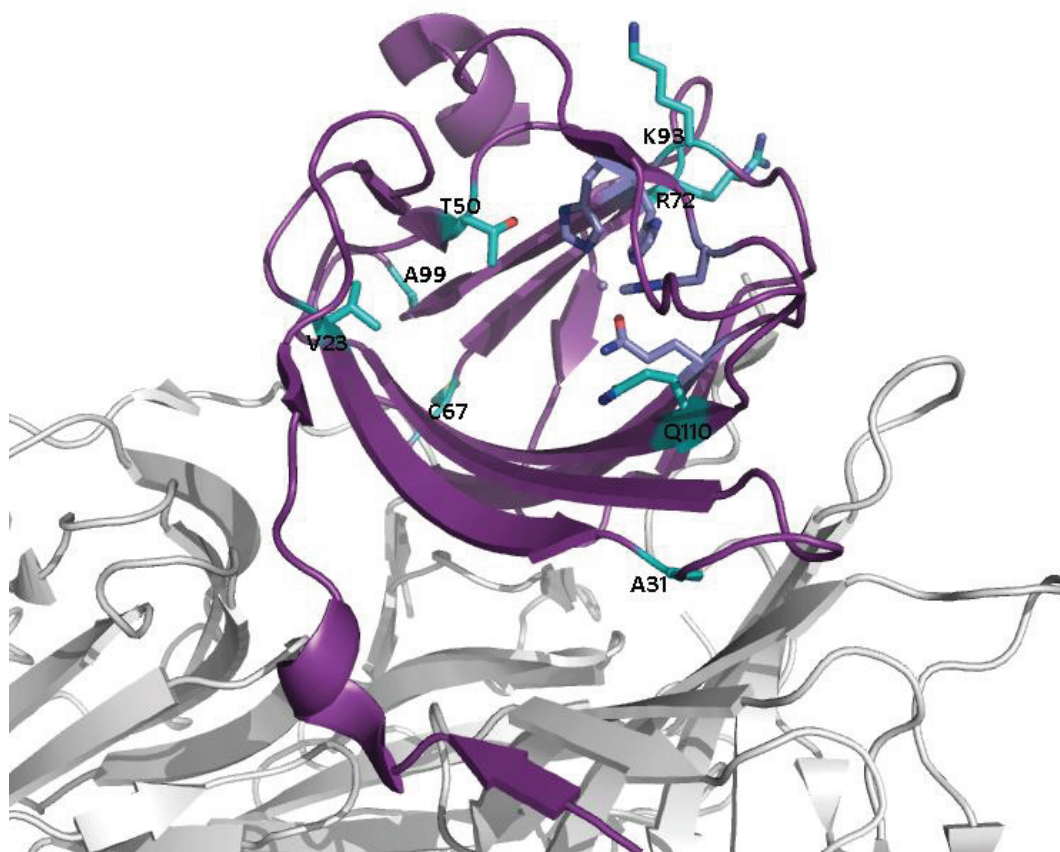


Figure 32. Structure of one monomer of the *GtHNL* (purple) with the metal-binding site and the manganese ion highlighted in lilac and the amino acid positions that were exchanged in at least one of the hit-variants highlighted in turquoise. The Figure was prepared using the PyMOL Molecular Graphics System and is based on the PDB entry 4bif.

5.10. Cultivation of *GtHNL*-A40R variants

As for the *AcHNL*-A40H library's hits, the chosen variants of the *GtHNL*-A40R library were cultivated and the thus obtained cell free lysates were used for the Bradford assay and SDS-PAGE. The results are summarized in Table 20 and Figure 33.

Table 20. Total protein concentrations of the cell free lysates of the hits identified during the screening of the *GtHNL*-A40R library.

variant	protein concentration [mg/mL]
<i>GtHNL</i> -A40R/Q110L	15.56
<i>GtHNL</i> -A40R/A99T/Q110L	14.14
<i>GtHNL</i> -V23A/A40R	14.49
<i>GtHNL</i> -A31T/A40R/C67S	12.98
<i>GtHNL</i> -A40R/T50S	12.90
<i>GtHNL</i> -A40R/K93R	14.15

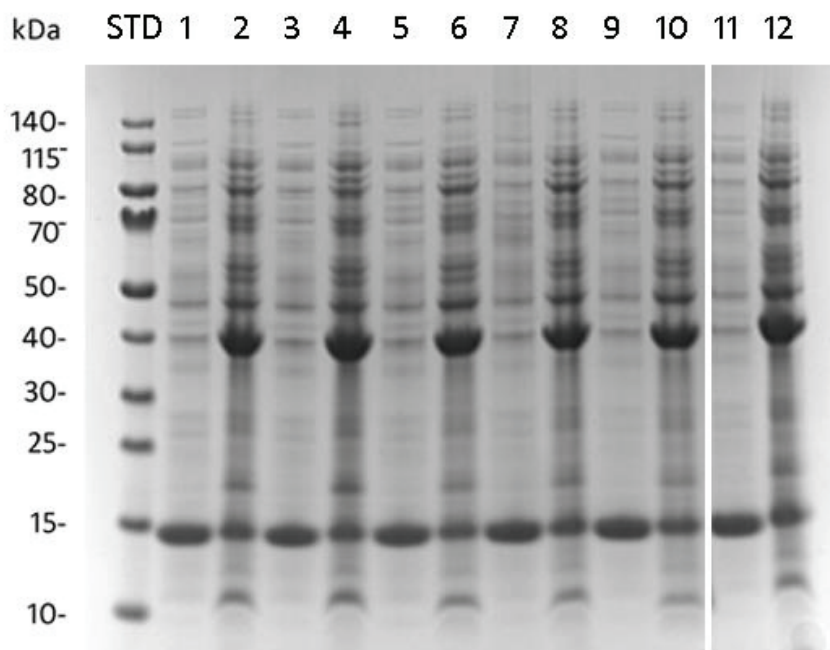


Figure 33. SDS-PAGE gel of cell free lysates (~5 µg total protein per sample) and corresponding pellet fractions of the *GtHNL-A40R* library's hits. MW *GtHNL* monomer: ~15 kDa. Lanes: *GtHNL-A40R/Q110L* lysate (1), *GtHNL-A40R/Q110L* pellet (2), *GtHNL-A40R/A99T/Q110L* lysate (3), *GtHNL-A40R/A99T/Q110L* pellet (4), *GtHNL-A31T/A40R/C67S* lysate (5), *GtHNL-A31T/A40R/C67S* pellet (6), *GtHNL-V23A/A40R* lysate (7), *GtHNL-V23A/A40R* pellet (8), *GtHNL-A40R/T50S* lysate (9), *GtHNL-A40R/T50S* pellet (10), *GtHNL-A40R/K93R* lysate (11), *GtHNL-A40R/K93R* pellet (12); STD: PageRuler Prestained Protein Ladder.

As the SDS gel shows, each hit clearly exhibited the overexpression of the respective *GtHNL* variant and the expression levels did not vary among the different clones. It also seemed like there were no variants that formed a significant amount of inclusion bodies.

The cell free lysates were therefore ready to be used for the final characterization. However, in order to reduce the number of lysates that had to be measured with HPLC, the conversions achieved by the different variants were first analysed applying the screening assay with cell free lysate and a higher number of samples per variant in order to increase the statistical significance of the results.

5.11. *AchNL-A40H* variants: conversion analysis in aqueous system

The BA conversion achieved by the *AchNL* variants under the conditions also present during the screening assay was analysed using the cell free lysates and six samples per variant. This experiment should assure that the variants measured with HPLC exhibited an activity that was at least in the same range as the parent variants activity. The respective decrease of BA was analysed by comparing the average absorbance measured after enzymatic reaction of each variant to the measured absorbance of the pET-26b control lysate. The resulting relative conversions can be seen in Figure 34.

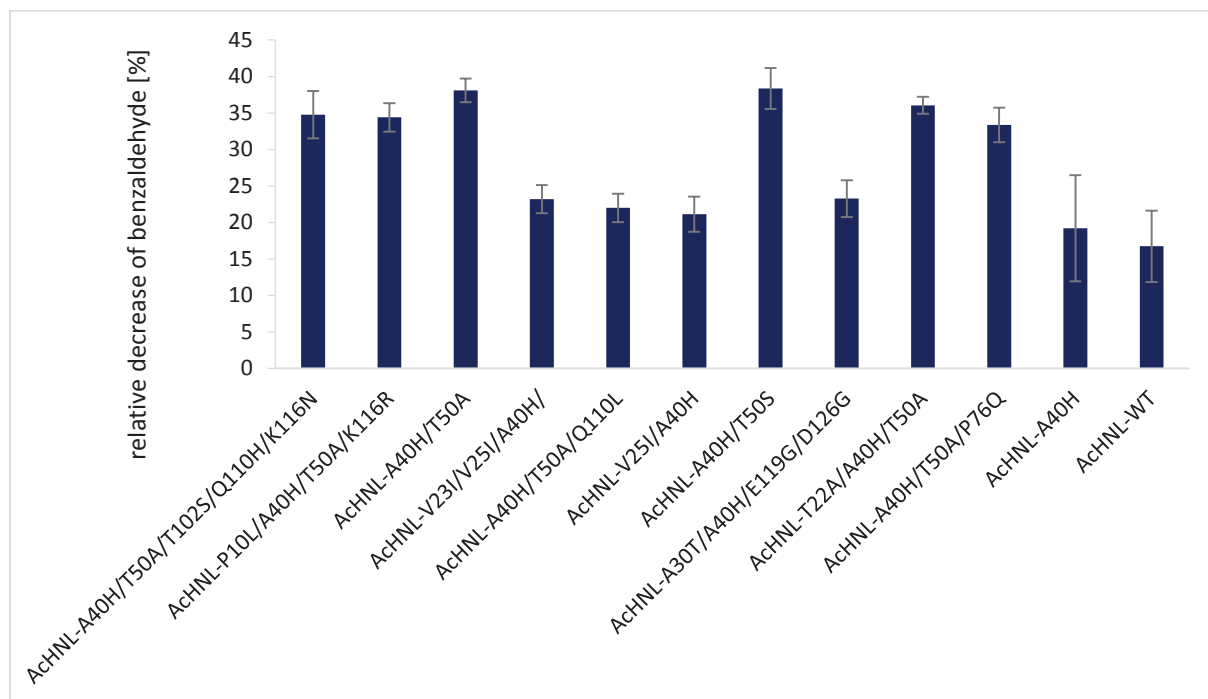


Figure 34. Relative decrease of BA, applying cell free lysates of the hits identified during the screening of the AChNL-A40H library. The percentage of BA converted was related to the remaining amount of BA in the pET-26b control sample. Measurements were performed six times for each sample according to the screening assay with cell free lysate and NE.

As Figure 34 shows, each of the measured variants showed at least a slight increase of conversion compared to the AChNL-A40H variant, even if the difference was not significant in some cases. However, since the evaluation of conversions measured in aqueous system is subject to significant fluctuations and measurement errors, the variants with conversions in the range of the parent variant were still included in the HPLC measurements just to be on the safe side.

Amongst the hits found in the AChNL-A40H library, there are five variants that harbour the AA exchange T50A in addition to one or more other AA exchanges. As the conversion analysis shows, none of these variants yielded a higher BA conversion than the variant AChNL-A40H/T50A. Thus it was concluded that the increased conversion yielded by these five variants was only due to the AA exchange T50A and the additional exchanges either did not influence the activity or, as it is clearly the case for the variant AChNL-A40H/T50A/Q110L, even decreased the resulting activity. Therefore, only the variants AChNL-A40H/T50A, AChNL-V23I/V25I/A40H, AChNL-V25I/A40H, AChNL-A40H/T50S and AChNL-A30T/A40H/E119G/D126G were chosen for the subsequent HPLC measurements.

5.12. GtHNL-A40R variants: conversion analysis in aqueous system

In the same way as the hits found in the AChNL-A40H library, the hits chosen from the GtHNL-A40R library were also measured in the aqueous system using the cell free lysates. The obtained conversions

can be seen in Figure 35. With regard to these measurements, there seems to be no variant with a big increase of BA conversion compared to the parent variant, but except for the variants *GtHNL-A31T/A40R/C67S* and *GtHNL-V23A/A40R* all the variants showed at least slightly increased conversions. The variant with the seemingly lowest conversion, *GtHNL-A31T/A40R/C67S*, yielded a BA decrease similar to the parent variant in an experiment carried out with lower total protein concentration (1.5 mg/mL instead of 3 mg/mL, data not shown) thus the lower conversion achieved in the experiment shown here might be due to a mistake made during the dilution or the measurement itself. However, since the difference was not too significant in any of these cases, all the variants were included in the HPLC measurement after all.

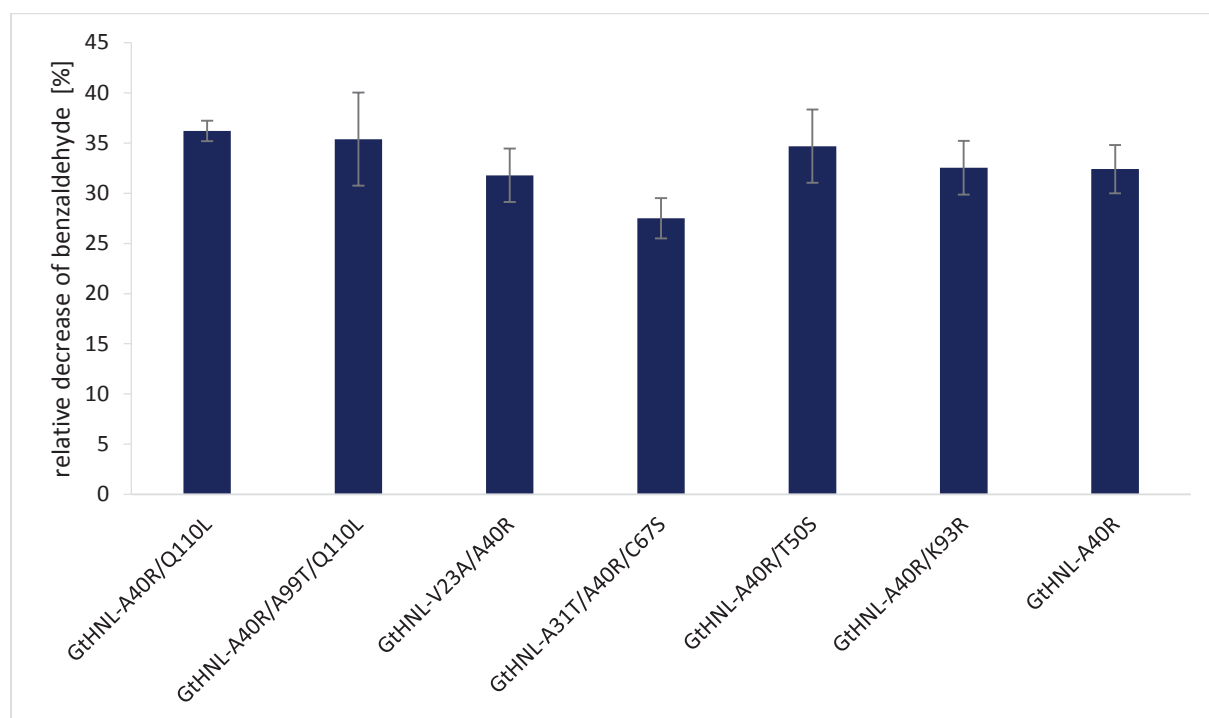


Figure 35. Relative decrease of BA, applying cell free lysates of the hits identified during the screening of the *GtHNL-A40R* library. The percentage of BA converted was related to the remaining amount of BA in the pET-26b control sample. Measurements were performed six times for each sample according to the screening assay with NE and cell free lysate.

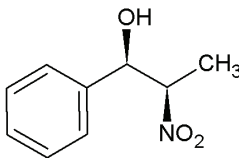
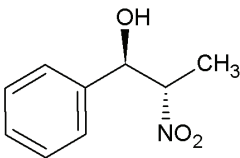
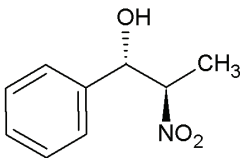
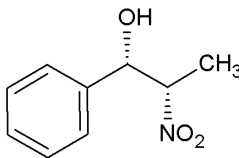
5.13. Final characterization in biphasic system and HPLC analysis

All the variants chosen for the final HPLC measurements can be looked up in Table 11 in part 4.8. The big advantage of this method is that the enantioselectivities of the various variants can be analysed in detail. Additionally, the unspecific background reaction is suppressed in the biphasic system and the HPLC analysis itself is, due to the use of an internal standard, much more precise compared to the screening assay.

After the reaction in the biphasic system had been started, each sample was measured in duplicate after 2 hours and after 24 hours. The resulting peak areas were normalized based on the internal standard and the conversion was calculated according to the BA peak area of the reference samples (see equation 1 in part 4.8.2). The samples and controls of the *AcHNL-A40H* library were measured separately from those of the *GtHNL-A40R* library and thus each library had extra reference samples. The exact resulting peak areas and all the chromatograms can be found on the acib server in the project folder 24051. Additionally, the chromatogram in Figure 41 in the appendix, which represents the pET-26b sample after 24 hours, relates all the eluted peaks to the corresponding substances. The chromatograms of the controls *AcHNL-A40H* and *GtHNL-A40R* after 2 hours of reaction are also shown in the appendix, see Figure 42 and Figure 43, respectively. A graphical representation of the average conversions can be seen in Figure 36, the corresponding data can be found in Table 21.

For each variant, the enantiomeric and the diastereomeric excess of the products were calculated, the results can also be viewed in Table 21. Additionally, for all of the hits a graphical representation of the four possible products' peak areas is shown in Figure 37 and Figure 38.

Table 21. Summary of the BA conversions and stereochemical properties of the reaction products yielded upon enzymatic conversion with the chosen hits of both libraries.

						
	(1R,2R)	(1R,2S)	(1S,2R)	(1S,2S)		
	reaction time [h]	conversion BA [%]	<i>ee</i> (R) anti [%]	<i>ee</i> (R) syn [%]	<i>de anti</i> (R,S) [%]	<i>anti/syn</i>
pET-26b (<i>AcHNL</i> library)	2 h	4.8	0.5	0.01	7.8	-0.90
	24 h	23	-0.4	0.28	-14	-1.6
<i>AcHNL-A40H</i>	2 h	7.2	69	45	37	1.6
	24 h	54.4	80	58	28	1.4
<i>AcHNL-A40H/T50A</i>	2 h	35	82	90	-13	0.91
	24 h	69	80	87	-24	0.92
<i>AcHNL-V23I/V25I/A40H</i>	2 h	18	61	53	24	1.2
	24 h	64	76	72	5.2	1.1
<i>AcHNL-V25I/A40H</i>	2 h	11	63	45	32	1.4
	24 h	60	81	65	22	1.2
<i>AcHNL-A40H/T50S</i>	2 h	41	65	85	-25	0.76
	24 h	68	79	86	-24	0.92
<i>AcHNL-A30T/A40H/E119G/D126G</i>	2 h	19	50	28	32	1.8
	24 h	42	56	26	22	2.1

	reaction time [h]	conversion BA [%]	<i>ee</i> (<i>R</i>) anti [%]	<i>ee</i> (<i>R</i>) syn [%]	<i>de anti</i> (<i>R,S</i>) [%]	<i>anti/syn</i>
pET-26b (<i>GtHNL</i> library)	2 h	3.7	0.9	-0.2	-4.9	6.9
	24 h	24	-1.2	0.3	-16	-4.2
<i>GtHNL</i>-A40R	2 h	14	81	80	16	1.0
	24 h	72	74	77	0.5	0.97
<i>GtHNL</i>-A40R/Q110L	2 h	32	88	86	25	1.0
	24 h	74	72	83	-18	0.87
<i>GtHNL</i>-V23A/A40R	2 h	21	60	69	0.2	0.87
	24 h	73	78	80	-13	0.98
<i>GtHNL</i>-A31T/A40R/C67S	2 h	19	70	72	14	0.97
	24 h	71	69	74	-1.9	0.93
<i>GtHNL</i>-A40R/T50S	2 h	27	68	80	-5.4	0.86
	24 h	72	81	86	-20	0.94
<i>GtHNL</i>-A40R/K93R	2 h	20	64	68	13	0.94
	24 h	71	68	76	-5.1	0.88

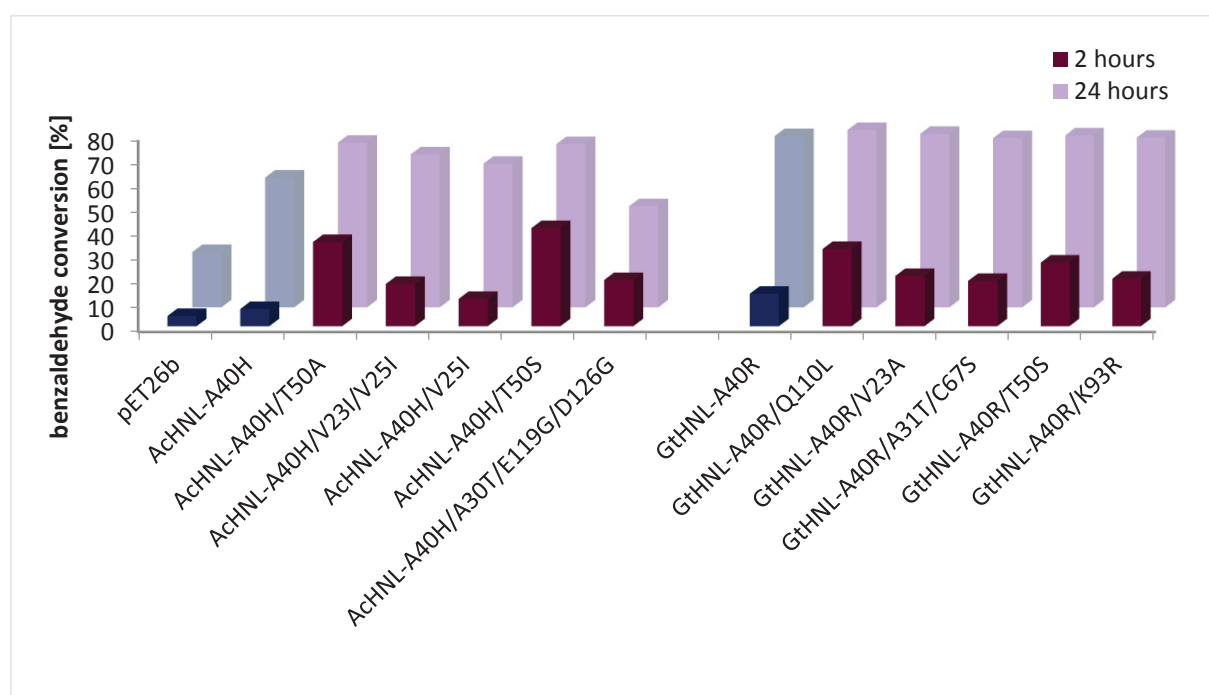


Figure 36. BA conversion achieved by the hits measured with HPLC after 2 and 24 hours of reaction time. Blue bars indicate the controls. Shown are the average conversions of the duplicate measurements.

The results of the HPLC measurement showed that all the chosen hits of both libraries achieved higher BA conversions than the respective controls, at least in the biphasic system. Of the AcHNL-A40H library, the variant AcHNL-A40H/T50S with 41 % conversion after 2 hours even yielded nearly six times the conversion of the control, which was 7 % after the same time. The variant AcHNL-V25I/A40H which had the comparatively lowest conversion could still convert 11 % BA after 2 hours, 4 % more than the parent variant. The same applies for the hits of the *GtHNL*-A40R library. Here, the best variant *GtHNL*-A40R/Q110L yielded with 32 % conversion after 2 hours 2.4-fold the conversion of the parent

variant with 14 %. Compared to that some of the other variants, for example *GtHNL-A31T/A40R/C67S* with 19 %, showed a comparatively slight increase of BA conversion.

Amongst all the variants of both libraries, the *AcHNL-A40H/T50S* variant achieved by far the highest conversion, followed by *AcHNL-A40H/T50A* and *GtHNL-A40R/Q110L*. Also the variant *GtHNL-A40R/T50S* achieved with 27 % a rather high conversion, however it is significantly lower than the conversion achieved with the corresponding *AcHNL* variant. This once more indicates that exchanges at the AA positions 40 and 50 cannot influence the enzyme activity independently from one another, even though a direct interaction of the two positions is, due to the long distance, not possible. Accordingly, the variant *AcHNL-T50A* that was also measured with HPLC did not achieve a higher conversion than the variant *AcHNL-A40H* (both variants converted 7 % of BA, data not shown), which is most likely due to a lower fraction of HNL in the soluble protein fraction because of inclusion body formation. Thus, the actual activity of purified *AcHNL-T50A* most likely is at least a little higher than that of purified *AcHNL-A40H* but any further specification would be just guesswork.

An additional effect that can also be seen in the aqueous system (see part 5.11 and 5.12) is that the BA conversion hardly ever exceeds 70 %. Independent of their respective conversion after 2 hours, none of the analysed variants yields a conversion higher than 74 % after 24 hours of reaction time. All measured variants of the *GtHNL-A40R* library, including the parent variant, achieve more or less the same conversion of 71-74 % after 24 hours. Also the two most active variants of the *AcHNL* library only achieve conversions of 68 and 69 % after 24 hours. Thus it can be concluded, that the reaction reaches an equilibrium at a BA conversion of around 70 %.

The negative control pET-26b only converts 4 % of BA after 2 hours compared to 49 % after the same time in the aqueous system. Although the conversion reached 23 % after 24 hours it is still obvious that the biphasic system strongly suppresses the unspecific background reaction and the corresponding chromatogram shows that the only relevant products formed after 24 hours are the four β -nitro alcohols (see Figure 41 in the appendix).

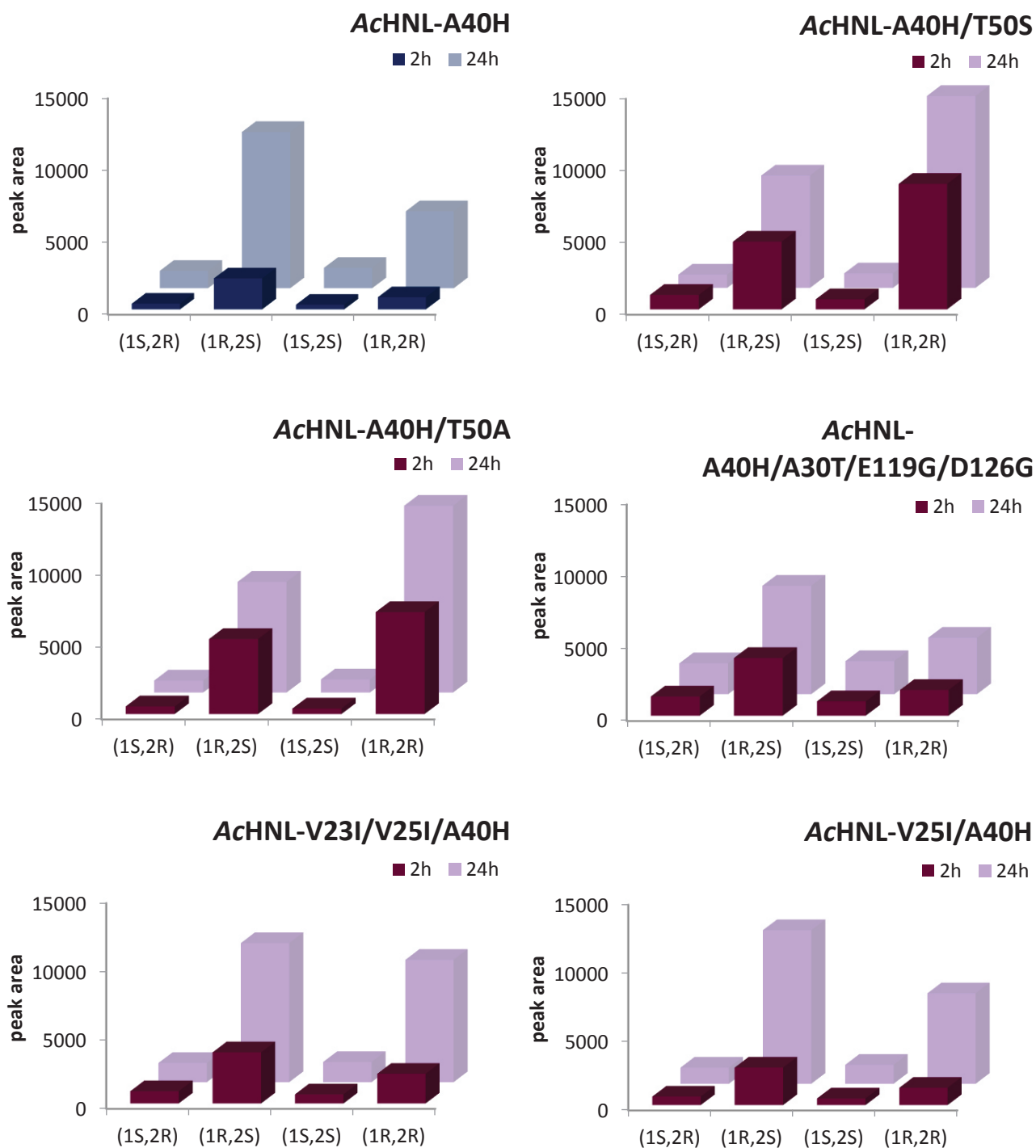


Figure 37. Representation of the formation of the four different products upon enzymatic conversion with AcHNL-A40H variants after 2 or 24 hours of reaction in the biphasic system. The peak areas correspond to the respective product concentration. (1S,2R) corresponds to product (1S,2R) 2-nitro-1-phenyl-1-propanol, (1R,2S) to product (1R,2S) 2-nitro-1-phenyl-1-propanol, (1S,2S) to product (1S,2S) 2-nitro-1-phenyl-1-propanol and (1R,2R) to product (1R,2R) 2-nitro-1-phenyl-1-propanol.

As promising as the analysis of conversions seemed, the analysis of enantioselectivities towards the four different products yielded very disappointing results. Of all the AcHNL-A40H variants, none

yielded a higher overall ratio of product 1*R*,2*S* after 2 hours than the parent variant AcHNL-A40H. The ratios of the 1*R*,2*S* products' peak areas to the sum of all products' peak areas can be seen in Table 22. First of all, after 2 h of reaction time, the variants AcHNL-A40H/T50A and AcHNL-A40H/T50S produced 1.4 and 1.9 times the amount of the 1*R*,2*R* product compared to the desired 1*R*,2*S* product, respectively. The production of the two 1*S* products was suppressed though, the *ee* (*R*) *anti* and *ee* (*R*) *syn* of the AcHNL-A40H/T50A products lie between 80 and 90 % and are thus even higher than the corresponding values of the AcHNL-A40H variant. Furthermore, the *ee* values of all variants remain more or less constant over the reaction time, indicating that the enzymatic reaction outperforms the unspecific background reaction.

The variant AcHNL-V23I/V25I/A40H also showed less preference for the 1*R*,2*S* product compared to the parent variant, even though not to the same extent as the two T50 variants. On the other hand, the variant AcHNL-A30T/A40H/E119G/D126G, which produced 2.2 times as much 1*R*,2*S* product than 1*R*,2*R* after two hours, showed a tendency to produce much more of the two 1*S* products, the *ee anti* was only around 50 % and the *ee syn* was with 26 % after 2 hours the lowest of all the measured variants.

Finally, the variant AcHNL-V25I/A40H showed only a slightly lower selectivity towards the 1*R*,2*S* product compared to the parent variant, especially after 24 hours, while for all the other variants the selectivity decreased with prolonged reaction time. However, the ratio of 1*R*,2*S* product (54 % after 2 h) of this variant is still lower than of the parent variant (58 % after 2 h) and additionally the overall BA conversion is only slightly higher (11 % compared to 7 % after 2 hours). The *de anti* is with 32 % after 2 h and 22 % after 24 h as good as the one of the AcHNL-A30T/A40H/E119G/D126G variant.

All in all, not a single variant exhibited a clear improvement of properties compared to the parent variant AcHNL-A40H.

Table 22. Summary of the relative production of the desired product, (1*R*,2*S*) 2-nitro-1-phenyl-1-propanol. The percentage of (1*R*,2*S*) was calculated as the ratio of the (1*R*,2*S*) product's peak area to the sum of all the four products' peak areas. The total conversion to (1*R*,2*S*) was the percentage of BA converted to the (1*R*,2*S*) product.

	% (1 <i>R</i> ,2 <i>S</i>) among all products		Total conversion to (1 <i>R</i> ,2 <i>S</i>) [%]	
	2h	24 h	2 h	24 h
pET-26b	27	21	1.3	4.9
AcHNL-A40H	58	58	4.2	31
AcHNL-A40H/T50A	40	34	14	24
AcHNL-V23I/V25I/A40H	50	46	8.8	30
AcHNL-V25I/A40H	54	55	6.1	33
AcHNL-A40H/T50S	31	34	13	23
AcHNL-A30T/A40H/E119G/D126G	49	47	9.5	20

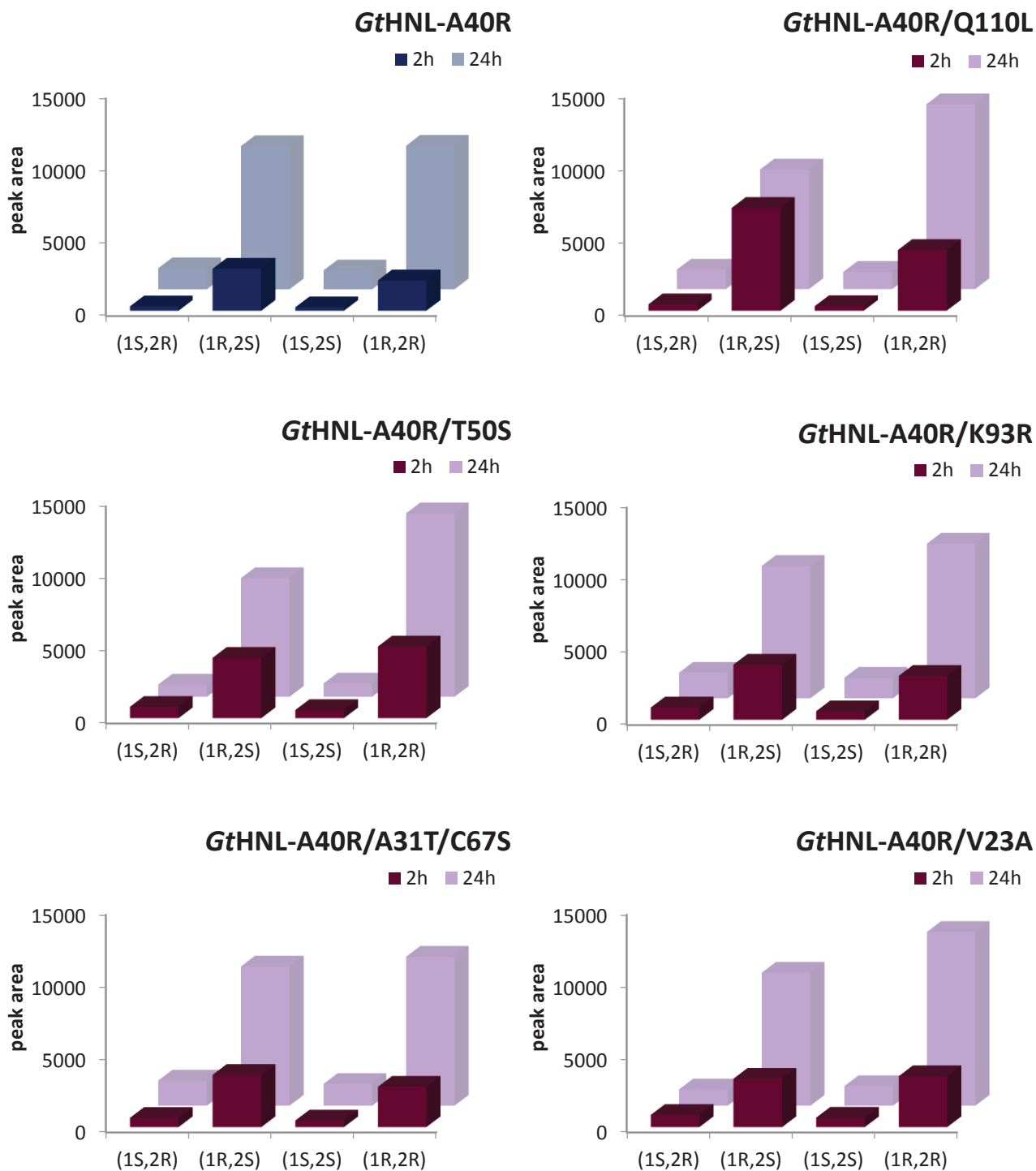


Figure 38. Representation of the formation of the four different products upon enzymatic conversion with GtHNL-A40R variants after 2 or 24 hours of reaction in the biphasic system. The peak areas correspond to the respective product concentration. (1S,2R) corresponds to product (1S,2R) 2-nitro-1-phenyl-1-propanol, (1R,2S) to product (1R,2S) 2-nitro-1-phenyl-1-propanol, (1S,2S) to product (1S,2S) 2-nitro-1-phenyl-1-propanol and (1R,2R) to product (1R,2R) 2-nitro-1-phenyl-1-propanol.

Table 23. Summary of the relative production of the desired product, (1*R*,2*S*) 2-nitro-1-phenyl-1-propanol. The percentage of (1*R*,2*S*) was calculated as the ratio of the (1*R*,2*S*) product's peak area to the sum of all the four products' peak areas. The total conversion to (1*R*,2*S*) was the percentage of BA converted to the (1*R*,2*S*) product.

	% (1 <i>R</i> ,2 <i>S</i>) among all products		Total conversion to (1 <i>R</i> ,2 <i>S</i>) [%]	
	2h	24 h	2 h	24 h
pET-26b	24	21	0.89	4.9
GtHNL-A40R	52	44	7.1	31
GtHNL-A40R/Q110L	59	35	19	26
GtHNL-V23A/A40R	40	39	8.4	28
GtHNL-A31T/A40R/C67S	48	42	9.2	29
GtHNL-A40R/T50S	40	36	11	26
GtHNL-A40R/K93R	46	40	9.2	28

As Figure 38 shows, the GtHNL-A40R variant produces more of the 1*R*,2*R* product compared to the AChNL-A40H right from the beginning. The *de anti* (*R,S*) of the AChNL-A40H is 37 % after 2 h compared to only 16 % for the GtHNL-A40R. The *ee* (*R*) *anti* and *ee* (*R*) *syn* values are with 81 and 80 % better than the corresponding values of the AChNL-A40H, which are 69 % for the *ee* (*R*) *anti* and 45 % for the *ee* (*R*) *syn*.

After 2 h the GtHNL-A40R produces still 1.4 times more of the 1*R*,2*S* product than of the 1*R*,2*R*, but after 24 hours these two products are present in roughly the same concentration. After 24 hours, there is not a single variant among the GtHNL-A40R hits that produces more 1*R*,2*S* product than 1*R*,2*R* product. However, after 2 hours of reaction time all variants, except for GtHNL-V23/A40R and GtHNL-A40R/T50S, produce at least a little more of the 1*R*,2*S* product than of the 1*R*,2*R* product. The *ee* (*R*) values, that lie mostly between 70 and 90 %, are on average higher than the corresponding values of the AChNL-A40H library hits.

Only one of the variants produces a higher ratio of 1*R*,2*S* product than the parent enzyme (52 %), namely the variant GtHNL-A40R/Q110L with 59 % (see Table 23). This variant, that also yielded the highest conversion of all the GtHNL variants, is definitely the most promising of all the enzymes identified during the screening of the two libraries. Overall 58 % of 1*R*,2*S* product after two hours and a BA conversion of 32 % is even better than the variant AChNL-A40H produced. Both *ee* values are with 86 % *ee syn* and 88 % *ee anti* after 2 h the very best of both libraries. However, after 24 hours of reaction time, not only the *ee* values had decreased, also the portion of 1*R*,2*R* product strongly increased and surpassed the amount of 1*R*,2*S* by far. Between 2 h and 24 h of reaction, the *de anti* (*R,S*) shifted from 25 % to -18 %. Therefore, this variant could only be used with shorter reaction times and thus also lower conversions, which is a major disadvantage.

Interestingly, a variant AChNL-A40H/T50A/Q110L had been identified during the screening of the AChNL-A40H library. However, the corresponding BA conversion measured in the aqueous system was

with 22 % considerably lower than that of the *AcHNL-A40H/T50A* variant with 38 % and hence this variant was not chosen for the final characterization. Other than the variant *GtHNL-A40R/Q110L*, it formed inclusion bodies during the expression. This once again indicates that interactions of other amino acid positions with the position T50 are essential for stability of at least the *AcHNL*.

As already mentioned in part 3.2.1, Bekerle-Bogner et al also studied the conversion of benzaldehyde with nitroethane as nucleophile and reaction in the biphasic system. They used purified preparations of *AcHNL-A40H* and *GtHNL-A40R* and applied 2.5 mg enzyme per mL. The conversions they achieved were generally around twice as high (16 % BA conversion with *AcHNL-A40H* and 27 % BA conversion with *GtHNL-A40R*) as was found for this thesis, but since the relative activities of the two variants were similar, the higher conversion was most likely due to higher overall concentration of enzymes. With the *AcHNL-A40H*, Bekerle-Bogner and coworkers observed a proportion of (1*R*,2*S*) product of 70 % compared to 58 % found during this study. A possible reason for the lower proportion could be a slower BA conversion and thus a higher contribution of the unspecific enzymatic background reaction.

6. Conclusion

Both random mutagenesis libraries yielded variants with significantly higher BA conversion compared to the respective parent variants. However, the HPLC analysis also showed that the selectivity towards the 1*R*,2*S* product decreased for most tested variants. The AcHNL-A40H library yielded the variant AcHNL-A40H/T50S which, although it yielded the highest BA conversion of all the variants, also produced nearly twice as much of the undesired product 1*R*,2*R* than of the targeted 1*R*,2*S* product. The best variant of the GtHNL-A40R library was the GtHNL-A40R/Q110L which yielded a considerably higher conversion than the parent variant and was also the only variant that actually exhibited a better selectivity towards the 1*R*,2*S* product than the parent variant, but only for a short reaction time (two hours). However, the respective *de anti* and *anti/syn* ratios were lower than those of the AcHNL-A40H variant which still showed the best selectivity of all the variants. With regard to that, construction of a new variant AcHNL-A40H/Q110L could possibly yield an enzyme with better activity and better - or at least preserved - selectivity.

7. List of Literature

- [1] Lanfranchi, E., Steiner, K., Glieder, A., Hajnal, I., Sheldon, R. A., van Pelt, S., Winkler, M. Mini-Review: Recent Developments in Hydroxynitrile Lyases for Industrial Biotechnology. *Recent Patents on Biotechnology*, 2013, 7, 197-206
- [2] Wiedner, R., Kothbauer, B., Pavkov-Keller, T., Gruber-Khadjawi, M., Gruber, K., Schwab, H., Steiner, K. Improving the Properties of Bacterial R-Selective Hydroxynitrile Lyases for Industrial Applications. *ChemCatChem*. 2015, 7, 325–332
- [3] Dadashipour, M., Asano, Y. Hydroxynitrile Lyases: Insights into Biochemistry, Discovery, and Engineering. *ACS Catal.* 2011, 1, 1121–1149
- [4] Purkarthofer, T., Skranc, W., Schuster, C., Griengl, H. Potential and capabilities of hydroxynitrile lyases as biocatalysts in the chemical industry. *Appl. Microbiol. Biot.* 2007, 76, 309-320.
- [5] Sukumaran, J., Hanefeld, U. Enantioselective C-C bond synthesis catalysed by enzymes. *Chem. Soc. Rev.* 2005, 34, 530-542.
- [6] Wohler, F., Liebig, J. Über die Bildung des Bittermandelöls. *Annalen der Pharmacie* 1837, 22:1-24.
- [7] Rosenthaler, L. Durch Enzyme bewirkte asymmetrische Synthesen. *Biochem. Z.* 1908, 14, 238-253.
- [8] Faber, K. Biotransformations in Organic Chemistry. Springer-Verlag Berlin Heidelberg, 2011
- [9] Poulton J. E. Cyanogenesis in Plants. *Plant Physiology* 1990, 94, 401-405
- [10] Nahrstedt, A. Cyanogenic Compounds as Protecting Agents for Organisms. *Plant Syst. Evol.* 1985, 150, 35-47.
- [11] Chen, FX, Feng, X. Asymmetric synthesis of cyanohydrins. *Curr. Org. Synth.* 2006, 3, 77-97.
- [12] Poechlauer, P., Skranc, W., Wubbolts, M. The large-scale biocatalytic synthesis of enantiopure cyanohydrins. *Asymmetric Catalysis on Industrial Scale: Challenges, Approaches and Solutions*. Wiley VCH Verlag GmbH & Co. KGaA, 2004, 327-340
- [13] Dausmann, T., Rosen, T. C., Dunkelmann, P. Oxidoreductases and hydroxynitrilase lyases: Complementary enzymatic technologies for chiral alcohols. *Eng. Life Sci.* 2006, 6, 125-129.
- [14] Hussain, Z., Wiedner, R., Steiner, K., Hajek, T., Avi, M., Hecher, B., Sessitsch, A., Schwab, H. Characterization of Two Bacterial Hydroxynitrile Lyases with High Similarity to Cupin Superfamily Proteins. *Appl. Environ. Microbiol.* 2012, 78, 2053–2055
- [15] Wiedner, R., Gruber-Khadjawi, M., Schwab, H. and Steiner, K. Discovery of a novel (R)-selective bacterial hydroxynitrile lyase from *Acidobacterium capsulatum*. *Comput. Struct. Biotechnol. J.* 2014, 10, 58–62
- [16] Hajnal, I., Lyskowski, A., Hanefeld, U., Gruber, K., Schwab, H., Steiner, K. Biochemical and structural characterization of a novel bacterial manganese-dependent hydroxynitrile lyase. *FEBS J.* 2013, 280, 5815-5828
- [17] Dunwell J. M., Purvis A., Khuri S. Cupins: the most functionally diverse protein superfamily? *Phytochemistry* 2004, 65, 7–17
- [18] Chavez F. A., Banerjee A., Sljivic B. Modeling the metal binding site in cupin proteins. Pramatarova L, editor. *On biomimetics*. In tech; 2011

- [19] Bekerle-Bogner, M., Gruber-Khadjawi, M., Wiltsche, H., Wiedner, R., Schwab, H. and Steiner, K. (R)-Selective Nitroaldol Reaction Catalyzed by Metal-Dependent Bacterial Hydroxynitrile Lyases. *ChemCatChem*. 2016, 8, 2214-2216
- [20] Henry, L., Nitro-alcohols. *Compt. Rend.* 1895, 120, 1265–1268.
- [21] Milner, S. E., Moody, T. S., Maguire, A. R. Biocatalytic Approaches to the Henry (Nitroaldol) Reaction. *Eur. J. Org. Chem.* 2012, 3059–3067
- [22] Sasai, H. Tokunaga, T. Watanabe, S. Suzuki, T. Itoh N., Shibasaki M. Efficient Diastereoselective and Enantioselective Nitroaldol Reactions from Prochiral Starting Materials: Utilization of La-Li-6,6'-Disubstituted BINOL Complexes as Asymmetric Catalysts. *J. Org. Chem.* 1995, 60, 7388– 7389
- [23] Marcelli, T., van der Haas R. N. S., van Maarseveen, J., Hiemstra, H. Asymmetric Organocatalytic Henry Reaction. *Angew. Chem. Int. Ed.* 2006, 45, 929 –931.
- [24] Arunachalam, R., Aswathi, C. S., Das, A., Kureshy, R. I., Subramanian, P. S. Diastereoselective Nitroaldol Reaction Catalyzed by Binuclear Copper(II) Complexes in Aqueous Medium. *ChemPlusChem*. 2015, 80, 209–216
- [25] Mei, H., Xiao, X., Zhao, X., Fang, B., Liu, X., Lin, L., Feng, X. Catalytic Asymmetric Henry Reaction of Nitroalkanes and Aldehydes Catalyzed by a Chiral N,N'-Dioxide/Cu(I) Complex. *J. Org. Chem.* 2015, 80, 2272–2280
- [26] Purkarthofer, T., Gruber, K., Gruber-Khadjawi, M., Waich, K., Skranc, W., Mink, D., Griengl, H. A Biocatalytic Henry Reaction—The Hydroxynitrile Lyase from *Hevea brasiliensis*. *Angew. Chem. Int. Ed.* 2006, 45, 3454 –3456
- [27] Gruber-Khadjawi, M., Purkarthofer, T., Skranc, W., Griengl, H. Hydroxynitrile Lyase-Catalyzed Enzymatic Nitroaldol (Henry) Reaction. *Adv. Synth. Catal.* 2007, 349, 1445 – 1450
- [28] Fuhshuku, K., Asano Y. Synthesis of (R)-nitro alcohols catalyzed by R-selective hydroxynitrile lyase from *Arabidopsis thaliana* in the aqueous–organic biphasic system. *J. Biotechnol.* 2011, 153 153–159
- [29] Buchholz, K., Kasche, V., Bornscheuer, U. T. Biocatalysts and Enzyme Technology. Wiley-VCH Verlag & Co. KGaA, 2012
- [30] Chica, RA., Doucet, N., Pelletier, J. N. Semi-rational approaches to engineering enzyme activity: combining the benefits of directed evolution and rational design. *Curr. Opin. Biotech.* 2005, 16, 378–384
- [31] Cadwell, R. C., Joyce, G. F. Randomization of genes by PCR mutagenesis. *PCR Methods Appl.*, 1992, 2, 28-33
- [32] Gubitz, G., Wintersteiger, R., Frei, R. W. Fluorogenic labeling of carbonyl compounds with 7-hydrazino-4-nitrobenzo-2-oxa-1,3,-diazole (NBD-H). *J. Liq. Chromatogr.* 1984, 7, 839–854.
- [33] Bornscheuer, U. T., Altenbrunner, J., and Meyer, H. Directed evolution of an esterase for the stereoselective resolution of a key intermediate in the synthesis of epothilones. *Biotechnol. Bioeng.*, 1998, 58, 554-559
- [34] Wong, T. S., Tee, K. L., Hauer, B., Schwaneberg, U. Sequence saturation mutagenesis (SeSaM): a novel method for directed evolution *Nucleic Acids Research*, 2004, 32, 3 e26
- [35] Gerlach, J., Reisinger, C., Pohn, B., Schwab, H., Klimant, I. 2-D solid-state assay platform: a tool for screening aldehyde-releasing enzyme activity in colonies. *Microchim Acta* 2007, 156, 209–218

- [36] Gerlach, J., Pohn, B., Karl, W., Scheideler, M., Uray, M., Bischof, H., Schwab, H., Klimant, I.** Planar optical sensors: A tool for screening enzyme activity in high density cell arrays. *Sensor. Actuat. B* 2006, 114, 984–994
- [37] Key, J., Li, C., Cairo, C.** Detection of Cellular Sialic Acid Content Using Nitrobenzoxadiazole Carbonyl-Reactive Chromophores. *Bioconjugate Chem.* 2012, 23, 363–371
- [38] Pscheidt, B., Liu, Z., Gaisberger, R., Avi, M., Skranc, W., Gruber, K., Griengl, H. and Glieder, A.** Efficient Biocatalytic Synthesis of (R)-Pantolactone. *Adv. Synth. Catal.* 2008. 350, 1943–1948

8. Appendix

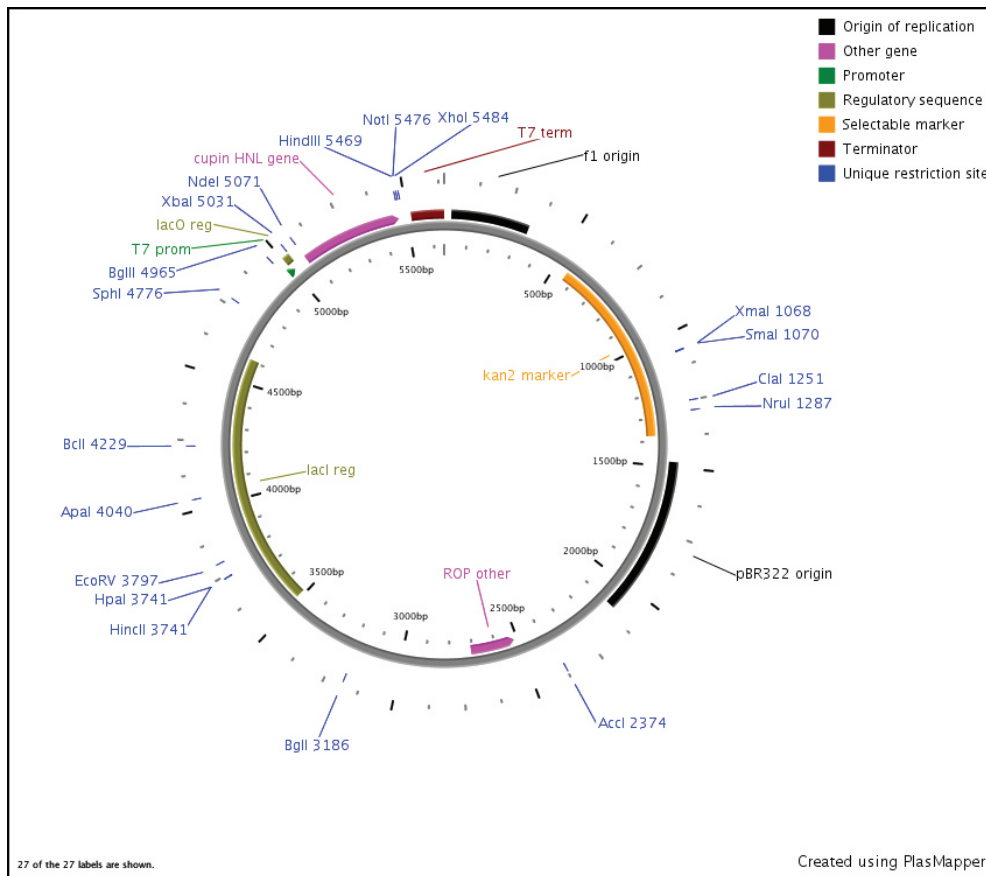


Figure 39. Plasmid map of the expression vector including the cupin-fold HNL gene, either *AcHNL-A40H* or *GtHNL-A40R*.

Table 24. Amino acid- and nucleotide sequences of the *GtHNL-A40R* and *AcHNL-A40H* enzymes, nucleotide sequences are both codon-optimized for *E. coli*.

<i>GtHNL-A40R</i>	amino acid sequence	MEIKRVGSQASGKGPADWFTGTVRIDPLFQAPDPALVAGRSVTFEPGARTAWHT HPLGQTLIVTAGCGWAQREGGAVEEIHPGDVVWFSPGKHWGHAAPTTAMTH LAIQERLDGKAVDWMEHVTDEQYRR
	nucleotide sequence	ATGGAAATTAACGTGTTGGTAGCCAGGCAAGCGGTAAAGGTCCGGCAGATT GGTTTACCGGCACCGTTCGTATTGATCCGCTGTTTCAGGCACCGGATCCGGCA CTGGTTGCCGGTAGGAGCGTTACCTTTGAACCGGTGCACGTACCGCATGGCA TACCCATCCGCTGGGTGAGACCCTGATTGTTACCGCAGGTTGTGGTTGGGCAC AGCGTGAAGGTGGTGCAGTTGAAGAAATTCACCGGGTGATGTTGTTTGGTTT AGTCCGGGTGAAAAACATTGGCATGGTGCAGCACCGACCACCGCAATGACCC ATCTGGCAATTCAGGAACGTCTGGACGGTAAAGCAGTTGATTGGATGGAACA TGTTACCGATGAACAGTATCGTCGC
<i>AcHNL-A40H</i>	amino acid sequence	MQITRNGSQPSGRGPAEYFTGTVRVDPLFAAPEPARVAGHSVTFEPGARTAWHT HPLGQTLIVTSGCGRVQREGGPVEEIRPGDVVWFPTGERHWHGASPSTAMTHIA IQEKLDGKVVEWLEHVTDAEYAG
	nucleotide sequence	ATGCAGATTACCCGTAATGGTAGCCAGCCGAGCGGTCGTGGTCCGGCAGAAT ATTCACCGGCACCGTTCGTGTTGATCCGCTGTTTGCAACCGGAACCGGCA CGTGTTCGCCGGTCATAGCGTTACCTTTGAACCGGTGCACGTACCGCATGGCA TACCCATCCGCTGGGTGAGACCCTGATTGTTACCAGCGGTTGTGGTTCGTGTTCA GCGTGAAGGTGGTCCGGTTGAAGAAATTCGTCCGGGTGATGTTGTTTGGTTTA CACCGGTGAACGTCAATGGCATGGTGAAGCCCGAGCACCGCAATGACCCA TATTGCAATTCAGAAAACTGGATGGCAAAGTTGTTGAATGGCTGGAACATG TTACCGATGCAGAATATGCAGGT

Table 25. Mutation rates of the various random mutagenesis libraries.

library		sequencing data	
parent variant	Mn ²⁺ concentration [mM]	No. of nucleotides exchanged per ORF	average mutation rate [nucleotides exchanged/ORF]
AcHNL-A40H	0.1	0	0.6
GtHNL-A40R	0.1	1, 1, 1, 0	
AcHNL-A40H	0.2	5, 5, 0	4.14
GtHNL-A40R	0.2	7, 6, 5, 1	
AcHNL-A40H	0.4	18, 9, 8	11.67
GtHNL-A40R	0.4	NA	
AcHNL-A40H	0.15	6, 5, 4, 4, 3, 3, 3, 2, 1, 0	3
GtHNL-A40R	0.15	8, 6, 6, 5, 5, 5, 4, 3, 3, 2	4.7

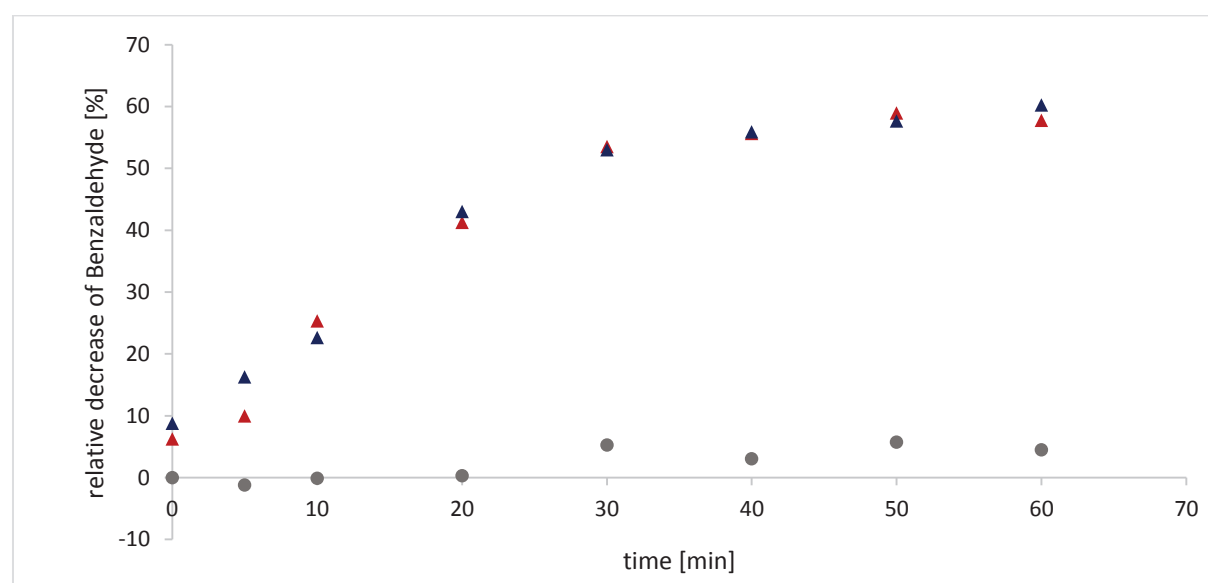


Figure 40. Time dependent relative decrease of BA, applying cell free lysates of AcHNL-A40H as well as the empty vector lysate pET-26b as background control. The decrease of BA was related to the BA concentration in the 0 minutes sample of pET-26b. Measurements were performed in duplicate for the AcHNL-A40H cell free lysate: red and blue triangles; grey circles correspond to the pET-26b controls. Measurements were performed applying the time dependent screening assay with NM.

Table 26. Scheme according to which the potential hits of the AchNL-A40H library were distributed on the rescreening plate 1 (RS 1). At positions E7 and F7 the clone E-O18 was accidentally used instead of E-O19 as had actually been planned. So the correct variant was only present in duplicate on RS 2. X indicates sterile wells, pET-26b indicates the empty vector control; AchNL-A40H and AchNL-WT indicate the parent variant and the wild type, respectively.

	1	2	3	4	5	6	7	8	9	10	11	12
A	X	A-O11	A-H5	A-L16	B-A8	C-C2	D-A2	E-P19	G-O2	H-M7	X	H-P3
B	pET-26b	A-O11	A-H5	A-L16	B-A8	C-C2	D-A2	E-P19	G-O2	H-M7	pET-26b	H-P3
C	pET-26b	A-D24	A-J9	B-B8	B-A18	C-P23	E-L3	F-H10	G-P3	H-M15	pET-26b	H-P14
D	AchNL-A40H	A-D24	A-J9	B-B8	B-A18	C-P23	E-L3	F-H10	G-P3	H-M15	AchNL-A40H	H-P14
E	AchNL-A40H	A-C2	A-N14	B-I15	B-I24	C-O24	E-O18	F-L23	G-P12	H-A12	AchNL-A40H	I-D20
F	AchNL-WT	A-C2	A-N14	B-I15	B-I24	C-O24	E-O18	F-L23	G-P12	H-A12	AchNL-WT	I-D20
G	AchNL-WT	A-B5	A-L8	B-C23	B-L23	C-D7	E-M24	F-L1	G-I1	H-J7	AchNL-WT	I-P8
H	X	A-B5	A-L8	B-C23	B-L23	C-D7	E-M24	F-L1	G-I1	H-J7	X	I-P8

Table 27. Scheme according to which the potential hits of the AchNL-A40H library were distributed on the rescreening plate 2 (RS 2). X indicates sterile wells, pET-26b indicates the empty vector control; AchNL-A40H and AchNL-WT indicate the parent variant and the wild type, respectively.

	1	2	3	4	5	6	7	8	9	10	11	12
A	A-C2	A-N14	B-I15	B-I24	C-O24	E-O19	F-L23	G-P12	H-A12	I-D20	X	X
B	A-C2	A-N14	B-I15	B-I24	C-O24	E-O19	F-L23	G-P12	H-A12	I-D20	pET-26b	pET-26b
C	A-B5	A-L8	B-C23	B-L23	C-D7	E-M24	F-L1	G-I1	H-J7	I-P8	pET-26b	pET-26b
D	A-B5	A-L8	B-C23	B-L23	C-D7	E-M24	F-L1	G-I1	H-J7	I-P8	AchNL-A40H	AchNL-A40H
E	A-O11	A-H5	A-L16	B-A8	C-C2	D-A2	E-P19	G-O2	H-M7	H-P3	AchNL-A40H	AchNL-A40H
F	A-O11	A-H5	A-L16	B-A8	C-C2	D-A2	E-P19	G-O2	H-M7	H-P3	AchNL-WT	AchNL-WT
G	A-D24	A-J9	B-B8	B-A18	C-P23	E-L3	F-H10	G-P3	H-M15	H-P14	AchNL-WT	AchNL-WT
H	A-D24	A-J9	B-B8	B-A18	C-P23	E-L3	F-H10	G-P3	H-M15	H-P14	X	X

Table 28. Scheme according to which the potential hits of the AchNL-A40H library were distributed on the rescreening plate 3 (RS 3). X indicates sterile wells, pET-26b indicates the empty vector control; AchNL-A40H and AchNL-WT indicate the parent variant and the wild type, respectively.

	1	2	3	4	5	6	7	8	9	10	11	12
A	X	I-G11	J-M7	K-A11	L-P10	O-G20	P-E10	Q-B17	Q-D14	X4-A18	X	X5-A10
B	pET-26b	I-G11	J-M7	K-A11	L-P10	O-G20	P-E10	Q-B17	Q-D14	X4-A18	pET-26b	X5-A10
C	pET-26b	I-M10	J-A11	K-N23	M-I13	O-I20	P-G6	Q-C14	X1-E3	X4-I9	pET-26b	X5-P1
D	AchNL-A40H	I-M10	J-A11	K-N23	M-I13	O-I20	P-G6	Q-C14	X1-E3	X4-I9	AchNL-A40H	X5-P1
E	AchNL-A40H	I-J5	J-P2	K-C24	M-P3	O-F2	P-J15	Q-M1	X2-I12	X4-E24	AchNL-A40H	X6-A9
F	AchNL-WT	I-J5	J-P2	K-C24	M-P3	O-F2	P-J15	Q-M1	X2-I12	X4-E24	AchNL-WT	X6-A9
G	AchNL-WT	J-B1	K-C15	L-D5	O-K23	O-D14	P-P3	Q-M4	X3-A2	X5-I8	AchNL-WT	PS D7
H	X	J-B1	K-C15	L-D5	O-K23	O-D14	P-P3	Q-M4	X3-A2	X5-I8	X	PS D7

Table 29. Scheme according to which the potential hits of the AChNL-A40H library were distributed on the rescreening plate 4 (RS 4). X indicates sterile wells, pET-26b indicates the empty vector control; AChNL-A40H and AChNL-WT indicate the parent variant and the wild type, respectively.

	1	2	3	4	5	6	7	8	9	10	11	12
A	I-J5	J-P2	K-C24	M-P3	O-F2	P-J15	Q-M1	X2-I12	X4-E24	X6-A9	X	X
B	I-J5	J-P2	K-C24	M-P3	O-F2	P-J15	Q-M1	X2-I12	X4-E24	X6-A9	pET-26b	pET-26b
C	J-B1	K-C15	L-D5	O-K23	O-D14	P-P3	Q-M4	X3-A2	X5-I8	PS D7	pET-26b	pET-26b
D	J-B1	K-C15	L-D5	O-K23	O-D14	P-P3	Q-M4	X3-A2	X5-I8	PS D7	AChNL-A40H	AChNL-A40H
E	I-G11	J-M7	K-A11	L-P10	O-G20	P-E10	Q-B17	Q-D14	X4-A18	X5-A10	AChNL-A40H	AChNL-A40H
F	I-G11	J-M7	K-A11	L-P10	O-G20	P-E10	Q-B17	Q-D14	X4-A18	X5-A10	AChNL-WT	AChNL-WT
G	I-M10	J-A11	K-N23	M-I13	O-I20	P-G6	Q-C14	X1-E3	X4-I9	X5-P1	AChNL-WT	AChNL-WT
H	I-M10	J-A11	K-N23	M-I13	O-I20	P-G6	Q-C14	X1-E3	X4-I9	X5-P1	X	X

Table 30. Scheme according to which the potential hits of the GtHNL-A40R library were distributed on the rescreening plate 1 (RS 1). X indicates sterile wells, pET-26b indicates the empty vector control; AChNL-A40H and GtHNL-A40R indicate the positive controls.

	1	2	3	4	5	6	7	8	9	10	11	12
A	X	A-I19	A-I14	A-B4	B-N11	B-I10	C-D12	E-G8	E-P12	F-E8	X	G-B9
B	X	A-I19	A-I14	A-B4	B-N11	C-I10	C-D12	E-G8	E-P12	F-E8	X	G-B9
C	GtHNL-A40R	A-G15	A-G18	A-J12	B-J3	C-N7	C-P10	E-J9	F-O7	F-F11	GtHNL-A40R	G-B16
D	GtHNL-A40R	A-G15	A-G18	A-J12	B-J3	C-N7	C-P10	E-J9	F-O7	F-F11	GtHNL-A40R	G-B16
E	pET-26b	A-O11	A-K24	B-C10	C-G13	C-P11	D-O12	E-J11	F-G11	G-I4	pET-26b	G-N24
F	pET-26b	A-O11	A-K24	B-C10	C-G13	C-P11	D-O12	E-J11	F-G11	G-I4	pET-26b	G-N24
G	AChNL-A40H	A-A12	A-J11	B-O24	C-O3	C-H10	D-C4	E-B4	F-K4	G-G20	AChNL-A40H	G-C12
H	AChNL-A40H	A-A12	A-J11	B-O24	C-O3	C-H10	D-C4	E-B4	F-K4	G-G20	AChNL-A40H	G-C12

Table 31. Scheme according to which the potential hits of the GtHNL-A40R library were distributed on the rescreening plate 2 (RS 2). X indicates sterile wells, pET-26b indicates the empty vector control; AChNL-A40H and GtHNL-A40R indicate the positive controls.

	1	2	3	4	5	6	7	8	9	10	11	12
A	A-O11	A-K24	B-C10	C-G13	C-P11	D-O12	E-J11	F-G11	G-I4	G-N24	X	X
B	A-O11	A-K24	B-C10	C-G13	C-P11	D-O12	E-J11	F-G11	G-I4	G-N24	X	X
C	A-A12	A-J11	B-O24	C-O3	C-H10	D-C4	E-B4	F-K4	G-G20	G-C12	AChNL-A40H	AChNL-A40H
D	A-A12	A-J11	B-O24	C-O3	C-H10	D-C4	E-B4	F-K4	G-G20	G-C12	AChNL-A40H	AChNL-A40H
E	A-I19	A-I14	A-B4	B-N11	C-I10	C-D12	E-G8	E-P12	F-E8	G-B9	pET-26b	pET-26b
F	A-I19	A-I14	A-B4	B-N11	C-I10	C-D12	E-G8	E-P12	F-E8	G-B9	pET-26b	pET-26b
G	A-G15	A-G18	A-J12	B-J3	C-N7	C-P10	E-J9	F-O7	F-F11	G-B16	GtHNL-A40R	GtHNL-A40R
H	A-G15	A-G18	A-J12	B-J3	C-N7	C-P10	E-J9	F-O7	F-F11	G-B16	GtHNL-A40R	GtHNL-A40R

Table 32. Scheme according to which the potential hits of the *GtHNL-A40R* library were distributed on the rescreening plate 3 (RS 3). X indicates sterile wells, pET-26b indicates the empty vector control; *AcHNL-A40H* and *GtHNL-A40R* indicate the positive controls.

	1	2	3	4	5	6	7	8	9	10	11	12
A	X	H-I7	H-P2	I-L11	J-K23	J-A12	K-A5	K-P18	L-J3	L-P12	X	M-M7
B	<i>AcHNL-A40H</i>	H-I7	H-P2	I-L11	J-K23	J-A12	K-A5	K-P18	L-J3	L-P12	<i>AcHNL-A40H</i>	M-M7
C	<i>AcHNL-A40H</i>	H-C4	H-P24	I-N20	J-I9	J-H3	K-K9	K-B10	L-N11	M-A1	<i>AcHNL-A40H</i>	M-E18
D	pET-26b	H-C4	H-P24	I-N20	J-I9	J-H3	K-K9	K-B10	L-N11	M-A1	pET-26b	M-E18
E	pET-26b	H-B5	I-O15	I-L8	J-C15	J-N13	K-J7	L-K23	L-P2	M-I5	pET-26b	M-D7
F	<i>GtHNL-A40R</i>	H-B5	I-O15	I-L8	J-C15	J-N13	K-J7	L-K23	L-P2	M-I5	<i>GtHNL-A40R</i>	M-D7
G	<i>GtHNL-A40R</i>	H-F23	I-A5	I-B6	J-I8	K-G7	K-F17	L-K7	L-F18	M-K11	<i>GtHNL-A40R</i>	M-N17
H	X	H-F23	I-A5	I-B6	J-I8	K-G7	K-F17	L-K7	L-F18	M-K11	X	M-N17

Table 33. Scheme according to which the potential hits of the *GtHNL-A40R* library were distributed on the rescreening plate 4 (RS 4). X indicates sterile wells, pET-26b indicates the empty vector control; *AcHNL-A40H* and *GtHNL-A40R* indicate the positive controls.

	1	2	3	4	5	6	7	8	9	10	11	12
A	H-B5	I-O15	I-L8	J-C15	J-N13	K-J7	L-K23	L-P2	M-I5	M-D7	X	X
B	H-B5	I-O15	I-L8	J-C15	J-N13	K-J7	L-K23	L-P2	M-I5	M-D7	<i>AcHNL-A40H</i>	<i>AcHNL-A40H</i>
C	H-F23	I-A5	I-B6	J-I8	K-G7	K-F17	L-K7	L-F18	M-K11	M-N17	<i>AcHNL-A40H</i>	<i>AcHNL-A40H</i>
D	H-F23	I-A5	I-B6	J-I8	K-G7	K-F17	L-K7	L-F18	M-K11	M-N17	pET-26b	pET-26b
E	H-I7	H-P2	I-L11	J-K23	J-A12	K-A5	K-P18	L-J3	L-P12	M-M7	pET-26b	pET-26b
F	H-I7	H-P2	I-L11	J-K23	J-A12	K-A5	K-P18	L-J3	L-P12	M-M7	<i>GtHNL-A40R</i>	<i>GtHNL-A40R</i>
G	H-C4	H-P24	I-N20	J-I9	J-H3	K-K9	K-B10	L-N11	M-A1	M-E18	<i>GtHNL-A40R</i>	<i>GtHNL-A40R</i>
H	H-C4	H-P24	I-N20	J-I9	J-H3	K-K9	K-B10	L-N11	M-A1	M-E18	X	X

Table 34. Scheme according to which the potential hits of the *GtHNL-A40R* library were distributed on the rescreening plate 5 (RS 5). X indicates sterile wells, pET-26b indicates the empty vector control; *AcHNL-A40H* and *GtHNL-A40R* indicate the positive controls.

	1	2	3	4	5	6	7	8	9	10	11	12
A	X	M-N6	N-L3	O-H3	P-A11	P-O23	P-B4	Q-I8	Q-D13	Q-N20	X	R-I6
B	<i>AcHNL-A40H</i>	M-N6	N-L3	O-H3	P-A11	P-O23	P-B4	Q-I8	Q-D13	Q-N20	<i>AcHNL-A40H</i>	R-I6
C	<i>AcHNL-A40H</i>	N-O11	O-C7	O-J5	P-E17	P-O24	Q-G19	Q-A16	Q-B3	Q-J10	<i>AcHNL-A40H</i>	R-D9
D	pET-26b	N-O11	O-C7	O-J5	P-E17	P-O24	Q-G19	Q-A16	Q-B3	Q-J10	pET-26b	R-D9
E	pET-26b	N-I9	O-I5	O-P15	P-I3	P-L18	Q-O11	Q-K12	Q-N23	R-M5	pET-26b	R-B7
F	<i>GtHNL-A40R</i>	N-I9	O-I5	O-P15	P-I3	P-L18	Q-O11	Q-K12	Q-N23	R-M5	<i>GtHNL-A40R</i>	R-B7
G	<i>GtHNL-A40R</i>	N-P15	O-A10	O-J10	P-M11	P-D16	Q-A5	Q-M14	Q-P4	R-O7	<i>GtHNL-A40R</i>	R-P6
H	X	N-P15	O-A10	O-J10	P-M11	P-D16	Q-A5	Q-M14	Q-P4	R-O7	X	R-P6

Table 35. Scheme according to which the potential hits of the *GtHNL-A40R* library were distributed on the rescreening plate 6 (RS 6). X indicates sterile wells, pET-26b indicates the empty vector control; *AcHNL-A40H* and *GtHNL-A40R* indicate the positive controls.

	1	2	3	4	5	6	7	8	9	10	11	12
A	N-I9	O-I5	O-P15	P-I3	P-L18	Q-O11	Q-K12	Q-N23	R-M5	R-B7	X	X
B	N-I9	O-I5	O-P15	P-I3	P-L18	Q-O11	Q-K12	Q-N23	R-M5	R-B7	AcHNL-A40H	AcHNL-A40H
C	N-P15	O-A10	O-J10	P-M11	P-D16	Q-A5	Q-M14	Q-P4	R-O7	R-P6	AcHNL-A40H	AcHNL-A40H
D	N-P15	O-A10	O-J10	P-M11	P-D16	Q-A5	Q-M14	Q-P4	R-O7	R-P6	pET-26b	pET-26b
E	M-N6	N-L3	O-H3	P-A11	P-O23	P-B4	Q-I8	Q-D13	Q-N20	R-I6	pET-26b	pET-26b
F	M-N6	N-L3	O-H3	P-A11	P-O23	P-B4	Q-I8	Q-D13	Q-N20	R-I6	GtHNL-A40R	GtHNL-A40R
G	N-O11	O-C7	O-J5	P-E17	P-O24	Q-G19	Q-A16	Q-B3	Q-J10	R-D9	GtHNL-A40R	GtHNL-A40R
H	N-O11	O-C7	O-J5	P-E17	P-O24	Q-G19	Q-A16	Q-B3	Q-J10	R-D9	X	X

Table 36. Scheme according to which the potential hits of the *GtHNL-A40R* library were distributed on the rescreening plate 7 (RS 7). X indicates sterile wells, pET-26b indicates the empty vector control; *AcHNL-A40H* and *GtHNL-A40R* indicate the positive controls.

	1	2	3	4	5	6	7	8	9	10	11	12
A	X	R-P12	S-C14	S-F10	X6-O2	X	S-E11	S-D3	X6-M13	X6-N1	X	X
B	AcHNL-A40H	R-P12	S-C14	S-F10	X6-O2	X	S-E11	S-D3	X6-M13	X6-N1	X	AcHNL-A40H
C	AcHNL-A40H	S-I11	S-L23	X2-K11	X6-J13	X6-D6	S-G6	S-J9	X6-M8	X6-B4	X	AcHNL-A40H
D	pET-26b	S-I11	S-L23	X2-K11	X6-J13	X6-D6	S-G6	S-J9	X6-M8	X6-B4	X	pET-26b
E	pET-26b	S-E11	S-D3	X6-M13	X6-N1	X6-D6	R-P12	S-C14	S-F10	X6-O2	X	pET-26b
F	GtHNL-A40R	S-E11	S-D3	X6-M13	X6-N1	X6-D6	R-P12	S-C14	S-F10	X6-O2	X	GtHNL-A40R
G	GtHNL-A40R	S-G6	S-J9	X6-M8	X6-B4	X	S-I11	S-L23	X2-K11	X6-J13	X	GtHNL-A40R
H	X	S-G6	S-J9	X6-M8	X6-B4	X	S-I11	S-L23	X2-K11	X6-J13	X	X

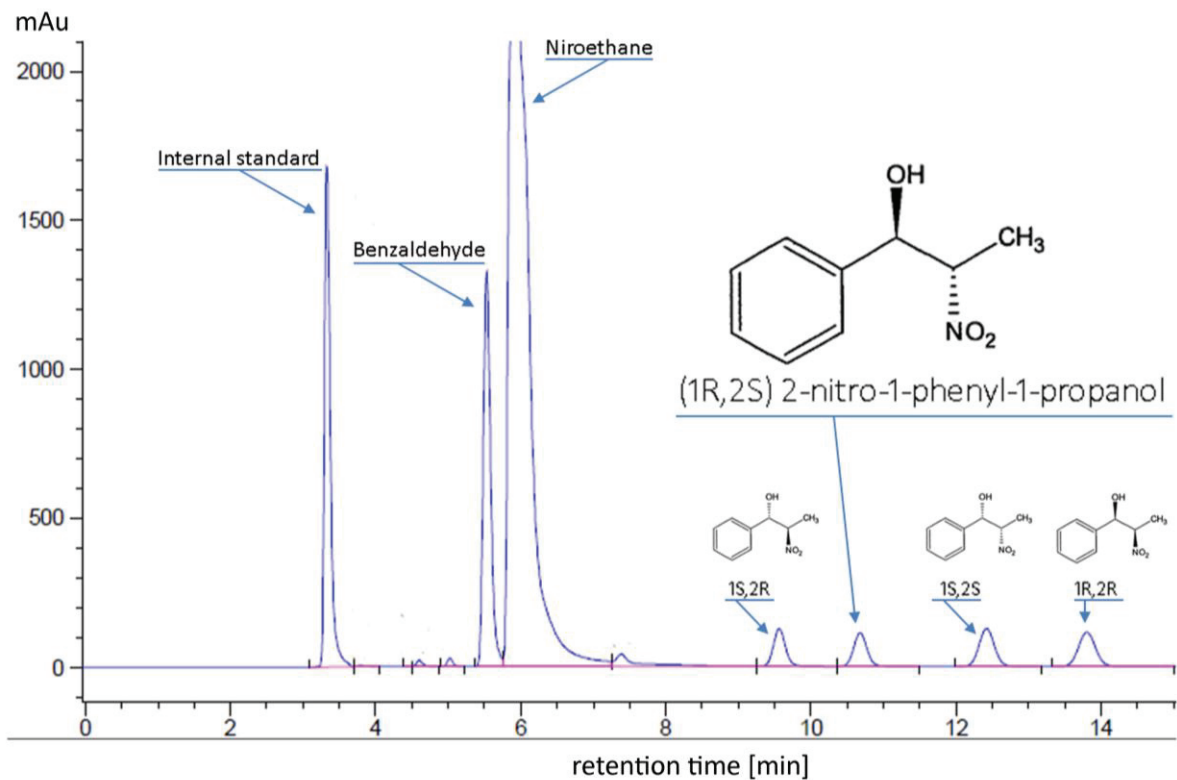


Figure 41. HPLC chromatogram of the control sample pET-26b, 24 h. Peaks were detected spectrophotometrically at 210 nm. Each major peak is assigned to the corresponding substance.

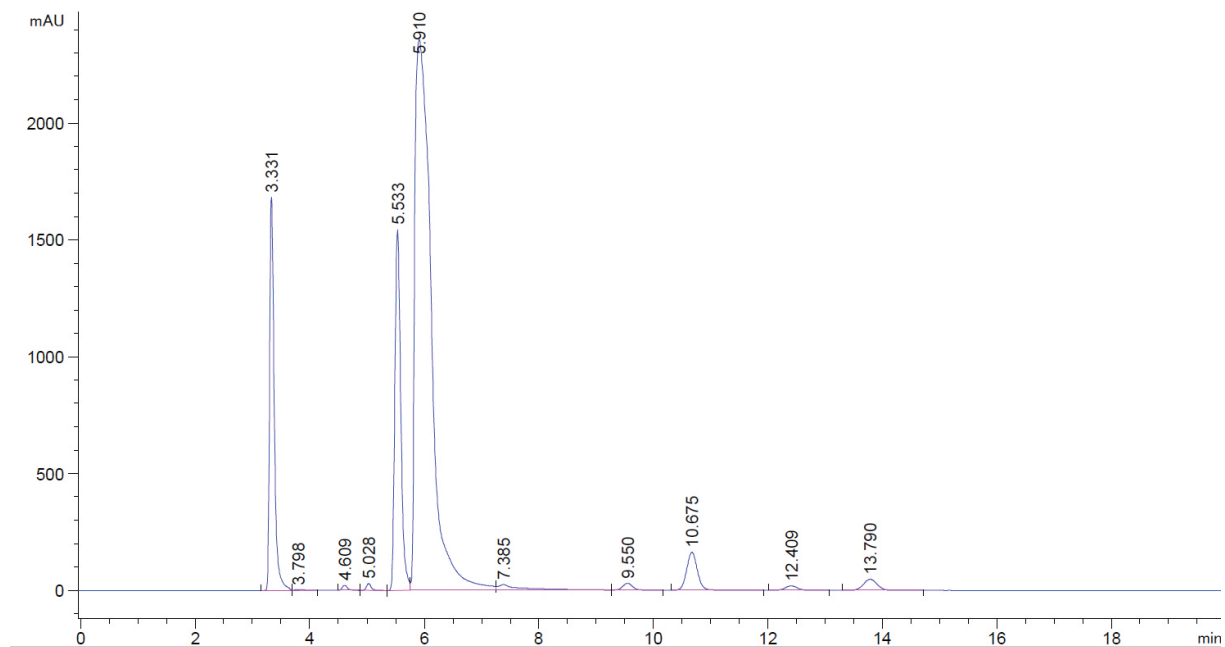


Figure 42. HPLC chromatogram of the control sample AchNL-A40H, 2 h. Peaks were detected spectrophotometrically at 210 nm.

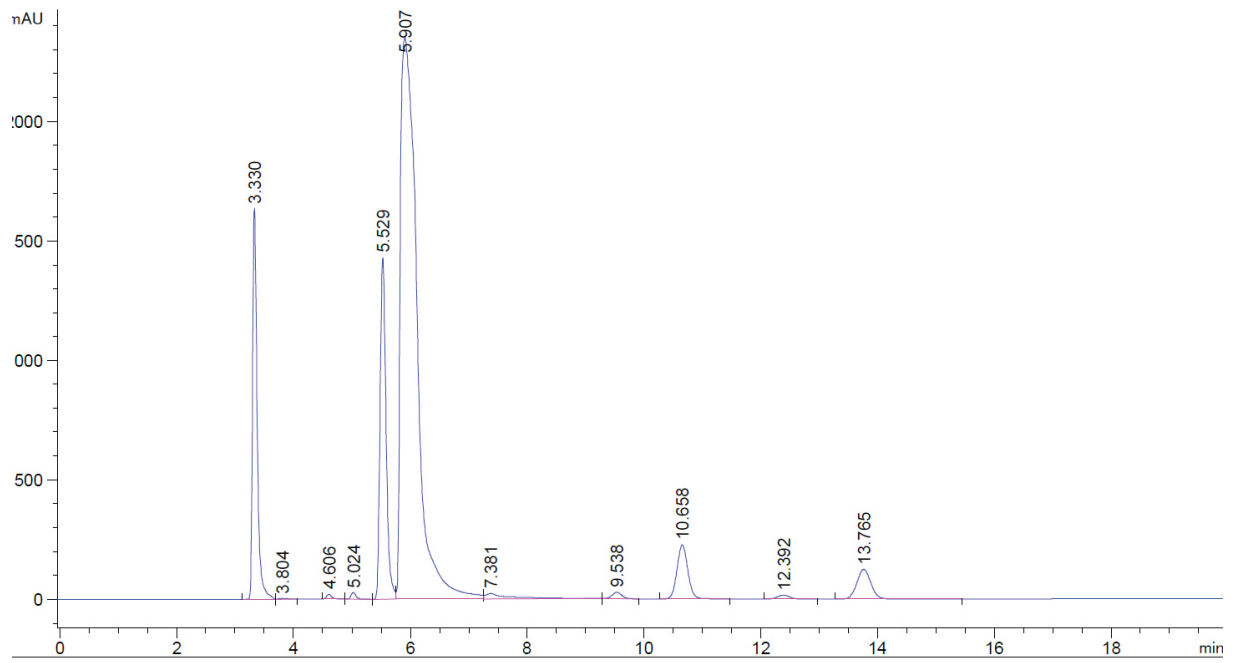


Figure 43. HPLC chromatogram of the control sample *GtHNL-A40R*, 2 h. Peaks were detected spectrophotometrically at 210 nm.

9. Table of Figures

Figure 1. Graphical representation of the 3-dimensional structure of two homotetramers of GtHNL. One monomer is depicted in red, the four corresponding metal binding amino acids and the manganese ion are highlighted in purple. The Figure was prepared using the PyMOL Molecular Graphics System and is based on the PDB entry 4bif.	2
Figure 2. This figure shows the correlation between the BA concentration in the finished measurement mixture of around 200 μ L and the absorbance at 522 nm. The samples for each BA concentration were prepared by diluting the appropriate amount of BA with a 50:50 mixture of 10 mM KPi pH 7 and 0.15 M KPi pH 5.72 to BA concentrations between 0.5 and 10 mM. Thereafter, these samples were handled in the same way as the reaction mixtures of the screening assay, with the exception that 10 μ L of sample were added to 140 μ L of 0.3 M CiPi pH 2.5 instead of 6 μ L. Each concentration was prepared and measured four times.	22
Figure 3. Time dependent relative decrease of BA, applying cell free lysates of the two enzyme variants GtHNL-A40R and AchNL-A40H, as well as the AchNL wild type (WT) and the empty vector control pET-26b as controls. The decrease of BA was related to the BA concentration in the 0 minutes sample of pET-26b. Measurements were performed four times for each sample and time point according to the time dependent screening assay with NE.	32
Figure 4. Result of the pre-screening of the AchNL-A40H library with 0.15 mM Mn^{2+} . The variants were screened two times, once according to the screening assay with BA and NM as substrates and once with BA and NE as substrates and the corresponding assay. Absorbance at 522 nm is directly proportional to residual BA concentration.	35
Figure 5. Result of the pre-screening of the GtHNL-A40R library with 0.15 mM Mn^{2+} . The variants were screened two times, once according to the screening assay with BA and NM as substrates and once with BA and NE as substrates and the corresponding assay. Absorbance at 522 nm is directly proportional to residual BA concentration.	36
Figure 6. Analysis of nucleotide exchanges in the AchNL-A40H coding region of ten arbitrarily chosen clones of the random mutagenesis library that was prepared with epPCR applying 0.15 mM Mn^{2+} ...	36
Figure 7. Screening result of DWP C-A2. The screening was performed according to the standard screening assay with NE. Each circle indicates a single variant, the measurement was carried out in duplicate (purple and lilac circles/lines). Black circles indicate a variant that was included in the rescreening and identified as hit "C-C2" (see part 6.4 for further details).	38
Figure 8. Screening result of DWP L-B2. The screening was performed according to the standard screening assay with NE. Each circle indicates a single variant, the measurement was carried out in duplicate (purple and lilac circles/lines).....	38
Figure 9. Screening result of DWP I-A2. The screening was performed according to the standard screening assay with NE. Each circle indicates a single variant, the measurement was carried out in duplicate (purple and lilac circles/lines). Black circles indicate a variant that was included in the rescreening and identified as hit "I-M10" (see part 6.4 for further details).	39
Figure 10. Screening result of DWP F-A2. The screening was performed according to the standard screening assay with NE. Each circle indicates a single variant, the measurement was carried out in duplicate (purple and lilac circles/lines).....	39
Figure 11. Screening result of DWP K-B1. The screening was performed according to the standard screening assay with NE. Each circle indicates a single variant, the measurement was carried out in duplicate (purple and lilac circles/lines). Black circles indicate a variant that was included in the rescreening and identified as hit "K-N23" (see part 6.4 for further details).....	39
Figure 12. Screening result of rescreening plate 1. The screening was performed according to the standard screening assay with NE. Each circle indicates a single variant, but the potential hits were	

positioned in pairs of two and the measurement was carried out in duplicate (purple and lilac circles/lines). 42

Figure 13. Screening result of rescreening plate 2. The screening was performed according to the standard screening assay with NE. Each circle indicates a single variant, but the potential hits were positioned in pairs of two and the measurement was carried out in duplicate (purple and lilac circles/lines). 42

Figure 14. Screening result of rescreening plate 3. The screening was performed according to the standard screening assay with NE. Each circle indicates a single variant, but the potential hits were positioned in pairs of two and the measurement was carried out in duplicate (purple and lilac circles/lines). PS functions as an abbreviation for “prescreening plate”. 42

Figure 15. Screening result of rescreening plate 4. The screening was performed according to the standard screening assay with NE. Each circle indicates a single variant, but the potential hits were positioned in pairs of two and the measurement was carried out in duplicate (purple and lilac circles/lines). PS functions as an abbreviation for “prescreening plate”. 43

Figure 16. Modelled structure of one monomer of the AchNL (purple) with the metal-binding site and the manganese ion highlighted in lilac and the amino acid positions that were exchanged in at least one of the hit-variants highlighted in turquoise. The Figure was prepared using the PyMOL Molecular Graphics System and is based on a modelled structure of AchNL which was computed on basis of the GtHNL crystal structure and the crystal structure of the GtHNL-A40H/V42T/Q110H variant (PDB entry 4bif and 4uxa, respectively). 46

Figure 17. SDS-PAGE gel of cell free lysates (~5 µg total protein per sample) and corresponding pellet fraction of AchNL-T50A variant and controls. MW AchNL monomer: ~14.1 kDa, GtHNL monomer: ~15 kDa. Lanes: AchNL-A40H lysate (1), GtHNL-A40R lysate (2), AchNL-T50A lysate (3), AchNL-T50A pellet (4), pET-26b lysate (5); STD: PageRuler Prestained Protein Ladder..... 47

Figure 18. Agarose gel electrophoresis of 16 clonoy PCR samples of clones of the AchNL-A40H/T50X library. 1-16 correspond to the 16 colony PCR mixtures; STD: O’GeneRuler 1kb Plus DNA Ladder, ready-to-use; Electrophoresis conditions: 1 % agarose, 120 V, 60 min..... 48

Figure 19. Screening result of AchNL-A40H/T50X plate 3. The screening was performed according to the standard screening assay with NE. Each circle indicates a single variant, the measurement was carried out in duplicate (purple and lilac circles/lines). Black circles indicate a variant that was identified as hit and chosen for sequencing. 49

Figure 20. Screening result of AchNL-A40H/T50X plate 5. The screening was performed according to the standard screening assay with NE. Each circle indicates a single variant, the measurement was carried out in duplicate (purple and lilac circles/lines). Black circles indicate a variant that was identified as hit and chosen for sequencing. 50

Figure 21. SDS-PAGE gel of cell free lysates (~5 µg total protein per sample) and corresponding pellet fractions of the AchNL-A40H library’s hits. MW AchNL monomer: ~14.1 kDa. Left gel: AchNL-A40H/T50A/T102S/Q110H/K116N lysate (1), AchNL-A40H/T50A/T102S/Q110H/K116N pellet (2), AchNL-P10L/A40H/T50A/K116R lysate (3), AchNL-P10L/A40H/T50A/K116R pellet (4), AchNL-A40H/T50A lysate (5), AchNL-A40H/T50A pellet (6), AchNL-V23I/V25I/A40H lysate (7), AchNL-V23I/V25I/A40H pellet (8), AchNL-A40H/T50A/Q110L lysate (9), AchNL-A40H/T50A/Q110L pellet (10), AchNL-V25I/A40H lysate (11), AchNL-V25I/A40H pellet (12); Right gel: AchNL-A40H/T50S lysate (13), AchNL-A40H/T50S pellet (14), AchNL-A30T/A40H/E119G/D126G lysate (15), AchNL-A30T/A40H/E119G/D126G pellet (16), AchNL-T22A/A40H/T50A lysate (17), AchNL-T22A/A40H/T50A pellet (18), AchNL-A40H/T50A/P76Q lysate (19), AchNL-A40H/T50A/P76Q pellet (20); STD: PageRuler Prestained Protein Ladder..... 52

Figure 22. Analysis of nucleotide exchanges in the GtHNL-A40R coding region of ten arbitrarily chosen clones of the random mutagenesis library that was prepared with epPCR applying 0.13 mM Mn ²⁺	53
Figure 23. Screening result of DWP N-B1. The screening was performed according to the standard screening assay with NE. Each circle indicates a single variant, the measurement was carried out in duplicate (purple and lilac circles/lines). Position F2 (black circles) corresponds to a variant that was included in the rescreening and identified as hit “N-L3” (see part 6.9 for further details).The position H8 (blue circles) was also included in the rescreening but on the rescreening plates it did not achieve a higher BA conversion than the controls.	54
Figure 24. Screening result of DWP R-B2. The screening was performed according to the standard screening assay with NE. Each circle indicates a single variant, the measurement was carried out in duplicate (purple and lilac circles/lines). Position H3 (black circles) corresponds to a variant that was included in the rescreening and identified as hit “R-P6” (see part 6.9 for further details).The position H6 (blue circles) was also included in the rescreening but on the rescreening plates it did not achieve a higher BA conversion than the controls.	54
Figure 25. Screening result of GtHNL-A40R rescreening plate 1. The screening was performed according to the standard screening assay with NE. Each circle indicates a single variant, but the potential hits were positioned in pairs of two and the measurement was carried out in duplicate (purple and lilac circles/lines).....	56
Figure 26. Screening result of GtHNL-A40R rescreening plate 2. The screening was performed according to the standard screening assay with NE. Each circle indicates a single variant, but the potential hits were positioned in pairs of two and the measurement was carried out in duplicate (purple and lilac circles/lines). Two measurement points of one of the duplicate data (purple) are outliers, most likely due to particles in the measured wells.....	57
Figure 27. Screening result of GtHNL-A40R rescreening plate 3. The screening was performed according to the standard screening assay with NE. Each circle indicates a single variant, but the potential hits were positioned in pairs of two and the measurement was carried out in duplicate (purple and lilac circles/lines).....	57
Figure 28. Screening result of GtHNL-A40R rescreening plate 4. The screening was performed according to the standard screening assay with NE. Each circle indicates a single variant, but the potential hits were positioned in pairs of two and the measurement was carried out in duplicate (purple and lilac circles/lines).....	57
Figure 29. Screening result of GtHNL-A40R rescreening plate 5. The screening was performed according to the standard screening assay with NE. Each circle indicates a single variant, but the potential hits were positioned in pairs of two and the measurement was carried out in duplicate (purple and lilac circles/lines).....	58
Figure 30. Screening result of GtHNL-A40R rescreening plate 6. The screening was performed according to the standard screening assay with NE. Each circle indicates a single variant, but the potential hits were positioned in pairs of two and the measurement was carried out in duplicate (purple and lilac circles/lines).....	58
Figure 31. Screening result of GtHNL-A40R rescreening plate 7. The screening was performed according to the standard screening assay with NE. Each circle indicates a single variant, but the potential hits were positioned in pairs of two and the measurement was carried out in duplicate (purple and lilac circles/lines).....	58
Figure 32. Structure of one monomer of the GtHNL (purple) with the metal-binding site and the manganese ion highlighted in lilac and the amino acid positions that were exchanged in at least one of the hit-variants highlighted in turquoise. The Figure was prepared using the PyMOL Molecular Graphics System and is based on the PDB entry 4bif.	60

Figure 33. SDS-PAGE gel of cell free lysates (~5 µg total protein per sample) and corresponding pellet fractions of the GtHNL-A40R library's hits. MW GtHNL monomer: ~15 kDa. Lanes: GtHNL-A40R/Q110L lysate (1), GtHNL-A40R/Q110L pellet (2), GtHNL-A40R/A99T/Q110L lysate (3), GtHNL-A40R/A99T/Q110L pellet (4), GtHNL-A31T/A40R/C67S lysate (5), GtHNL-A31T/A40R/C67S pellet (6), GtHNL-V23A/A40R lysate (7), GtHNL-V23A/A40R pellet (8), GtHNL-A40R/T50S lysate (9), GtHNL-A40R/T50S pellet (10), GtHNL-A40R/K93R lysate (11), GtHNL-A40R/K93R pellet (12); STD: PageRuler Prestained Protein Ladder.....	61
Figure 34. Relative decrease of BA, applying cell free lysates of the hits identified during the screening of the AchNL-A40H library. The percentage of BA converted was related to the remaining amount of BA in the pET-26b control sample. Measurements were performed six times for each sample according to the screening assay with cell free lysate and NE.	62
Figure 35. Relative decrease of BA, applying cell free lysates of the hits identified during the screening of the GtHNL-A40R library. The percentage of BA converted was related to the remaining amount of BA in the pET-26b control sample. Measurements were performed six times for each sample according to the screening assay with NE and cell free lysate.	63
Figure 36. BA conversion achieved by the hits measured with HPLC after 2 and 24 hours of reaction time. Blue bars indicate the controls. Shown are the average conversions of the duplicate measurements.....	65
Figure 37. Representation of the formation of the four different products upon enzymatic conversion with AchNL-A40H variants after 2 or 24 hours of reaction in the biphasic system. The peak areas correspond to the respective product concentration. (1S,2R) corresponds to product (1S,2R) 2-nitro-1-phenyl-1-propanol, (1R,2S) to product (1R,2S) 2-nitro-1-phenyl-1-propanol, (1S,2S) to product (1S,2S) 2-nitro-1-phenyl-1-propanol and (1R,2R) to product (1R,2R) 2-nitro-1-phenyl-1-propanol. ...	67
Figure 38. Representation of the formation of the four different products upon enzymatic conversion with GtHNL-A40R variants after 2 or 24 hours of reaction in the biphasic system. The peak areas correspond to the respective product concentration. (1S,2R) corresponds to product (1S,2R) 2-nitro-1-phenyl-1-propanol, (1R,2S) to product (1R,2S) 2-nitro-1-phenyl-1-propanol, (1S,2S) to product (1S,2S) 2-nitro-1-phenyl-1-propanol and (1R,2R) to product (1R,2R) 2-nitro-1-phenyl-1-propanol. ...	69
Figure 39. Plasmid map of the expression vector including the cupin-fold HNL gene, either AchNL-A40H or GtHNL-A40R.	76
Figure 40. Time dependent relative decrease of BA, applying cell free lysates of AchNL-A40H as well as the empty vector lysate pET-26b as background control. The decrease of BA was related to the BA concentration in the 0 minutes sample of pET-26b. Measurements were performed in duplicate for the AchNL-A40H cell free lysate: red and blue triangles; grey circles correspond to the pET-26b controls. Measurements were performed applying the time dependent screening assay with NM... ..	77
Figure 41. HPLC chromatogram of the control sample pET-26b, 24 h. Peaks were detected spectrophotometrically at 210 nm. Each major peak is assigned to the corresponding substance.....	82
Figure 42. HPLC chromatogram of the control sample AchNL-A40H, 2 h. Peaks were detected spectrophotometrically at 210 nm.....	82
Figure 43. HPLC chromatogram of the control sample GtHNL-A40R, 2 h. Peaks were detected spectrophotometrically at 210 nm.....	83

10. List of Tables

Table 1. Summary of the components of the epPCR mixture.	11
Table 2. Temperature program of the epPCR.	11
Table 3. Summary of the components of the QuikChange PCR mixture.	13
Table 4. Temperature program of the QuikChange PCR.	13
Table 5. Summary of the components of the DpnI digest reaction mixture.	13
Table 6. Positions of the controls in the 384 well MTPs of the two random mutagenesis libraries. ...	16
Table 7. List of chemicals and related data required for the screening assay.	18
Table 8. Summary of the components of the QuikChange PCR mixture for site specific and site saturation mutagenesis.	24
Table 9. Summary of the components of the colony PCR mixture.	25
Table 10. Temperature program of the colony PCR.	25
Table 11. List of the variants and controls of both sets measured with HPLC.	30
Table 12. List of the components of organic substrate solution for a single HPLC sample. However, for each of the two sets of samples an extra, pre-mixed batch of substrate solution had been prepared.	30
Table 13. Summary of all the potential hits included in the rescreening of the AchNL-A40H library. Indicated are the positions on the 384 well MTP and on the corresponding DWP. PS functions as an abbreviation for “pre-screening plate”	41
Table 14. Summary of the nucleotide mutations and amino acid exchanges found in the AchNL-A40H library’s best hits’ sequences. Variants chosen for further characterization are underlined.	43
Table 15. Total protein concentrations of the cell free lysates of the controls and the variant AchNL-T50A.	47
Table 16. Summary of the identified nucleotide mutations and amino acid exchanges of the most promising variants of the AchNL-A40H/T50X library.	50
Table 17. Total protein concentrations of the cell free lysates of the hits identified during the screening of the AchNL-A40H library.	51
Table 18. Summary of all the potential hits included in the rescreening of the GtHNL-A40R library. Indicated are the positions on the 384 well MTP and on the corresponding DWP.	55
Table 19. Summary of the nucleotide mutations and amino acid exchanges found in the GtHNL-A40R library’s best hits’ sequences. Variants chosen for further characterization are underlined.	59
Table 20. Total protein concentrations of the cell free lysates of the hits identified during the screening of the GtHNL-A40R library.	60
Table 21. Summary of the BA conversions and stereochemical properties of the reaction products yielded upon enzymatic conversion with the chosen hits of both libraries.	64
Table 22. Summary of the relative production of the desired product, (1R,2S) 2-nitro-1-phenyl-1-propanol. The percentage of (1R,2S) was calculated as the ratio of the (1R,2S) product’s peak area to the sum of all the four products’ peak areas.	68
Table 23. Summary of the relative production of the desired product, (1R,2S) 2-nitro-1-phenyl-1-propanol. The percentage of (1R,2S) was calculated as the ratio of the (1R,2S) product’s peak area to the sum of all the four products’ peak areas.	70
Table 24. Amino acid- and nucleotide sequences of the GtHNL-A40R and AchNL-A40H enzymes, nucleotide sequences are both codon-optimized for E. coli.	76
Table 25. Mutation rates of the various random mutagenesis libraries.	77
Table 26. Scheme according to which the potential hits of the AchNL-A40H library were distributed on the rescreening plate 1 (RS 1). At positions E7 and F7 the clone E-O18 was accidentally used instead of E-O19 as had actually been planned. So the correct variant was only present in duplicate	

on RS 2. X indicates sterile wells, pET-26b indicates the empty vector control; AchNL-A40H and AchNL-WT indicate the parent variant and the wild type, respectively..... 78

Table 27. Scheme according to which the potential hits of the AchNL-A40H library were distributed on the rescreening plate 2 (RS 2). X indicates sterile wells, pET-26b indicates the empty vector control; AchNL-A40H and AchNL-WT indicate the parent variant and the wild type, respectively..... 78

Table 28. Scheme according to which the potential hits of the AchNL-A40H library were distributed on the rescreening plate 3 (RS 3). X indicates sterile wells, pET-26b indicates the empty vector control; AchNL-A40H and AchNL-WT indicate the parent variant and the wild type, respectively..... 78

Table 29. Scheme according to which the potential hits of the AchNL-A40H library were distributed on the rescreening plate 4 (RS 4). X indicates sterile wells, pET-26b indicates the empty vector control; AchNL-A40H and AchNL-WT indicate the parent variant and the wild type, respectively..... 79

Table 30. Scheme according to which the potential hits of the GtHNL-A40R library were distributed on the rescreening plate 1 (RS 1). X indicates sterile wells, pET-26b indicates the empty vector control; AchNL-A40H and GtHNL-A40R indicate the positive controls. 79

Table 31. Scheme according to which the potential hits of the GtHNL-A40R library were distributed on the rescreening plate 2 (RS 2). X indicates sterile wells, pET-26b indicates the empty vector control; AchNL-A40H and GtHNL-A40R indicate the positive controls. 79

Table 32. Scheme according to which the potential hits of the GtHNL-A40R library were distributed on the rescreening plate 3 (RS 3). X indicates sterile wells, pET-26b indicates the empty vector control; AchNL-A40H and GtHNL-A40R indicate the positive controls. 80

Table 33. Scheme according to which the potential hits of the GtHNL-A40R library were distributed on the rescreening plate 4 (RS 4). X indicates sterile wells, pET-26b indicates the empty vector control; AchNL-A40H and GtHNL-A40R indicate the positive controls. 80

Table 34. Scheme according to which the potential hits of the GtHNL-A40R library were distributed on the rescreening plate 5 (RS 5). X indicates sterile wells, pET-26b indicates the empty vector control; AchNL-A40H and GtHNL-A40R indicate the positive controls. 80

Table 35. Scheme according to which the potential hits of the GtHNL-A40R library were distributed on the rescreening plate 6 (RS 6). X indicates sterile wells, pET-26b indicates the empty vector control; AchNL-A40H and GtHNL-A40R indicate the positive controls. 81

Table 36. Scheme according to which the potential hits of the GtHNL-A40R library were distributed on the rescreening plate 7 (RS 7). X indicates sterile wells, pET-26b indicates the empty vector control; AchNL-A40H and GtHNL-A40R indicate the positive controls. 81