Simone Maria Petek, BSc

PROTEIN ENGINEERING OF HYDROXYNITRILE LYASES

MASTER'S THESIS

to achieve the university degree of

Diplom-Ingenieurin

Master's degree programme: Biotechnology

submitted to

Graz University of Technology

Supervisor

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

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

Institute for Molecular Biotechnology - Graz University of Technology

Austrian Center for Industrial Biotechnology (acib) - Graz

Graz, 08/2015

AFFIDAVIT

I declare that I have authored this thesis independently. 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

Acknowledgment

An dieser Stelle möchte ich mich bei allen Personen bedanken, die zur Entstehung der vorliegenden Diplomarbeit beigetragen haben:

… Herrn Prof. Helmut Schwab für das Bereitstellen dieses interessanten Themas für die Diplomarbeit

…. meiner Betreuerin Frau Dr. Kerstin Steiner für die hervorragende Begleitung über den gesamten Arbeitsverlauf hinweg

… Frau Karin Reicher und Frau Myria Bekerle-Bogner für die Probenmessungen am Institut für organische Chemie

… allen acib Mitarbeitern für ihre freundliche Hilfsbereitschaft

Zahvalim se tudi pri moji družini, ki mi je vedno dragocena spremljevalka na moji poti.

Abstract

Hydroxynitrile lyases (HNLs) are enzymes, which catalyse the formation and cyanogenesis of cyanohydrins. Cyanohydrins have industrial relevance, since they are valuable intermediates in the pharmaceutical and agrochemical industry.

In this work HNLs were engineered by rational mutagenesis and directed evolution approaches. Recently, a bacterial cupin fold HNL from *Granulicella tundricola* (*Gt*HNL) was investigated for its activity on different substrates. Amino acids in the active site were investigated for their influence on accepted substrate species and enantioselectivity [1]. These investigations were continued in the first part of this work. Site specific mutations and site saturation mutagenesis in the active site of *Gt*HNL were performed. Single and double variants of *Gt*HNL and *Gt*HNL_A40HV42TQ110H (*Gt*HNL-Tr) were tested in their cyanogenesis and synthesis activity on (*R*)-mandelonitrile, (*R*) and (*S*)-2-chloromandelonitrile. The double mutant *Gt*HNL_A40HV42T showed a promising activity and enantioselectivity in both reaction directions.

The library screening of the site saturation mutagenesis approach identified amino acids, which were beneficial in the cyanogenesis towards (*R*) and (*S*)-2-chloromandelonitrile.

Several hydroxynitrile lyases not only catalyse the cleavage and formation of cyanohydrins but also show a promiscuous activity for the nitroaldol reaction. In the second part of this work a random library of a hydroxynitrile lyase from *Acidobacterium capsulatum* (*Ac*HNL) was screened for improved activity in the formation of nitroalcohols. The screening encompassed a library size of ~ 5700 enzyme variants, resulting in variants with slightly improved activities compared to the wild type. Furthermore, selected variants from *Gt*HNL and *Ac*HNL were tested for their nitroaldolase activity. The amino acid exchange A40H showed in both enzymes an improved nitroaldolase activity.

Zusammenfassung

Hydroxynitril-Lyasen (HNLs) sind Enzyme, welche die Spaltung und Synthese von Cyanohydrinen katalysieren. Cyanohydrine sind von industrieller Relevanz, da sie wertvolle Zwischenprodukte in der Pharma,- und Agrarindustrie darstellen.

In dieser Arbeit wurden rationale und gerichtete Mutagenese-Methoden auf Hydroxynitril-Lyasen angewendet.

Kürzlilch wurde eine bakterielle HNL mit einem Cupin-Faltungsmotiv von *Granulicella tundricola* charakterisiert. Das Enzyme sowie Varianten mit ausgetauschten Aminosäuren im katalytischen Zentrum wurden auf das Substratspektrum und auf die Enantioselektivität hin untersucht [1]. Diese Untersuchungen wurden im ersten Teil dieser Arbeit fortgeführt. Im aktiven Zentrum von *Gt*HNL wurden zum einem spezifische Mutationen eingeführt und zum anderen eine gesättigte Mutagenese, wiederum an spezifischen Stellen angewendet. Die Einzel,- und Doppelmutanten wurden auf (*R*)-Mandelonitril in beide Reaktionsrichtungen und in der Spaltung von (*R*) und (*S*)-2-Chloromandelonitril getestet. Die Doppelmutante *Gt*HNL_A40HV42T zeigte sowohl in der Spaltungs,- als auch in der Synthesereaktion eine hervorragende Aktivität und Enantioselektivität.

Das Screening der gesättigten Variantenbibliotheken auf (*R*) und (*S*)-2-Chloromandelonitril identifizierte Aminosäuren, welche sich günstig auf die Aktivität gegenüber dem (*R*) und (*S*) Enantiomer auswirkten.

Einige der Hydroxynitril-Lyasen aus der Cupin-Familie katalysieren nicht nur die Spaltung und Synthese von Cyanohydrinen, sondern zeigen auch eine promiskuitive Nitroaldol-Reaktion. Im zweiten Teil dieser Arbeit wurde eine zufallsgenerierte Variantenbibliothek der promiskuitiven Hydroxynitril-Lyase von *Acidobacterium capsulatum* hergestellt. Insgesamt wurden ~ 5700 Enzymvarianten auf ihre Nitroaldol-Aktivität getestet. Das Screening identifizierte Varianten, welche im Vergleich zum Wildtyp eine signifikant verbesserte Aktivität in der Synthese von Nitroalkoholen zeigten.

Weiters wurden in dieser Arbeit ausgewählte Enzymvarianten von *Gt*HNL und *Ac*HNL auf ihre Nitroaldol-Aktivität getestet, wobei der Aminosäureaustausch von Alanine in Position 40 zu Histidin (A40H) auf eine klare Verbesserung der Aktivität bei beiden Enzymen bewirkte.

Table of Contents

List of abbreviations

%	percent
°C	degree Celsius
3D	three dimensional
acib	Austrian Center for Industrial Biotechnology
ADH	alcohol dehydrogenase
bp	base pair
C-C	carbon-carbon
cfu	colony forming unit
Cl	chlorine
cm	centimeter
Cu	copper
Cup9	*Gt*HNL
dATP	deoxyadenosine triphosphate
dCTP	deoxycytidine triphosphate
dGTP	deoxyguanosine triphosphate
dH_2O	deionised water
DMSO	dimethylsulfoxid
DNA	deoxyribonucleic acid
dNTP	deoxynucleoside triphosphate
DTT	dithiotheitol
dTTP	deoxythymidine triphosphate
DWP	deep well plate
e.g.	for example
EDTA	ethylenediaminetetraacetic acid
ee	enantiomeric excess
epPCR	error prone polymerase chain reaction
et al.	et alii
EtBr	ethidium bromide
EtOH	ethanol
f	dilution factor
FAD	flavin adenine dinucleotide
FI	fluorescence intensity
Fig.	figure
g	g-force
g	gramme
GC	gas chromatography
GMC	glucose-methanol-choline
h	hour
H_2	hydrogen
HCN	hydrogenic cyanide
HEPES	2-[4-(2-hydroxyethyl)peperazin-1-yl]ethanesulfonic acid
HF	high fidelity
HNL	hydroxynitrile lyase
IPTG	isopropyl β-D-1-thiogalactopyranoside
ISO	isothermal
K_2HPO_4	dipotassium hydrogen phosphate
Kan	kanamycin
kb	kilobase
KCl	potassium chloride
kDA	kilodalton
KH_2PO_4	potassium dihydrogen phosphate
kPA	kilopascal
kV	kilovolt
L	litre
LB	Luria-Bertani
LDS	lithium dodecyl sulfate
LE	low electroendosmosis
M	molar
mA	milliampere
MES	2-(N-morpholino) ethanesulfonic acid
mg	milligramme
$MgCl_2$	magnesium dichloride
$MgSO_4$	magnesium sulfate
mmol	millimole
$MnCl_2$	manganese dichloride
mol	mole
MTBE	methyl tert-butyl ether
min	minute
mL	milliliter
mM	millimolar
Mn	manganese
(*R*)-2-ClMN	*(R)*-2-chloromandelonitrile
(*R*)-MN	*(R)*-mandelonitrile
®	registered trade mark
µg	mikrogramme
µL	mikrolitre
µmol	mikromole
Na_2HPO_4	disodium hydrogen phosphate
NaCl	sodium chloride
NaCN	sodium cyanide
NAD	nicotineamide adenine dinucleotide

NaH_2PO_4	sodium dihydrogen phosphate
NaOH	sodium hydroxyide
NBDH	4-hydrazino-7-nitrobenzofurazan
ng	nanogramme
Ni	nickel
nm	nanometer
nt	nucleotide
OD_{600}	optical density at λ=600nm
PAGE	polyacrylamide gel electrophoresis
PCR	polymerase chain reaction
PEG	polyethylenglycol
pmol	pikomole
PP	polypropylene
PS	polystyrene
qcPCR	QuikChange polymerase chain reaction
RSD	relative standard deviation
rpm	revolutions per minute
SDS	sodium dodecyl sulfate
(*S*)-2-ClMN	*(S)*-2-chloromandelonitrile
SOC	super optimal broth with catabolite repression
tab.	tabl
TAE	tris acetate EDTA
™	trade mark
Tris	2-amino-2-hydroxymethyl-propane-1,3-diol
TY	bacto tryptone yeast
U	µmol/min
UV	ultraviolet
V	volt
v/v	volume/volume
Vis	visible
V_{tot}	volume total
W	watt
w/v	weight/volume
WT	wild type
Zn	zinc
ΔA	change of absorption
ε	excitation coefficient
λ	wavelength

aa amino acids

G	glycine	**D**	aspartate
A	alanine	**E**	glutamate
V	valine	**N**	asparagine
L	leucine	**Q**	glutamine
I	isoleucine	**F**	phenylalanine
M	methionine	**Y**	tyrosine
P	proline	**W**	tryptophan
S	serine	**K**	lysine
T	threonine	**R**	arginine
C	cysteine	**H**	histidine

1 Introduction

1.1 Hydroxynitrile lyases

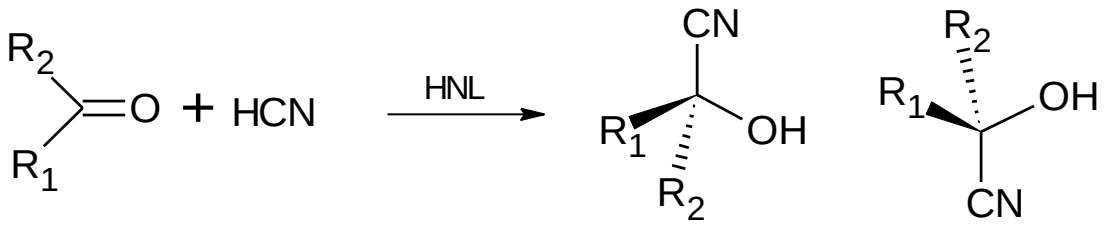
Biotransformations play a major role in industrial catalysis. Biocatalysis is a sustainable technology that can improve the economics of a given process and furthermore reduce environmental impact [2]. So far a great number of enzyme-based processes have been developed as an integrative step of process reactions and established in synthetic industry. In the last decade the replacement of conventional chemo-catalysts took place especially in the pharmaceutical and agrochemical industry [3]. Enzymes are widely recognised among the most active and selective catalysts for the preparation of enantiopure chemicals. Enantioselective enzymes allow an easy and waste-saving production of optically active intermediates which are used as building blocks, auxiliaries or advanced intermediates. They are applied in the chemical industry either in asymmetric synthesis reactions or for dynamic resolutions of racemic mixtures [4] [5].

1.1.1 Occurrence and function of hydroxynitrile lyases

Almost all hydroxynitrile lyases (HNLs) naturally occur in cyanogenic plants. A small amount is also found in other organisms e.g. bacteria. The natural role of HNLs is to cleave cyanohydrins and to release HCN when a plant is damaged. The role of cyanohydrins is also related to this mechanism of plant defense. Another biological function of cyanohydrins is the storage of nitrogen [6].

Hydroxynitrile lyases can be divided into different protein classes. Crystallographic 3D structure analysis and sequence alignment enable to classify the different HNLs. So far their protein structures are related to different fold families e.g. to α/β-hydrolases, carboxypeptidases, FAD-dependent oxidoreductases, zinc-dependent alcohol dehydrogenases and cupins. All fold families exhibit a different architecture of the active site, while catalysing the very same reaction. A lot of effort has been made to understand the catalytic mechanism of these enzymes. However, according to their enantioselectivity HNLs can be divided into (*R*) or (*S*) selective enzymes [7] [8] [9] [10].

In vitro, the synthesis of optically active cyanohydrins is catalysed by hydroxynitrile lyases. In this C-C bond-forming reaction hydrogenic cyanide (HCN) is condensed with an aldehyde or ketone to result a chiral cyanohydrin (Scheme 1). These products serve as starting material for chemical follow-up reactions. For example (*R*)-chloromandelic acid the product of (*R*)-chlorobenzaldehyde and HCN is an essential building block of the platelet aggregation inhibitor Clopidrogel [11].



Scheme 1: Synthesis reaction of cyanohydrins catalysed by HNL. Hydrogenic cyanide (HCN) is asymmetrically condensed with an aldehyde or ketone to form chiral cyanohydrins [12].

For the formation of enantiopure cyanohydrins on a preparative industrial scale, several HNLs are used. Table 1 lists the most common and well characterised HNLs [13].

Table 1: Most common hydroxynitrile lyases used on technical scale. All enzymes originate from different plant species belonging to four different protein folds. All of these HNLs were tailored for the industrial application by protein engineering.

	*Me*HNL	*Hb*HNL	*Pa*HNL	*Lu*HNL	*Sb*HNL	*At*HNL
Source organism	*Manihot esculenta*	*Hevea brasiliensis*	*Prunus amygdalus*	*Linum usitatissimum*	*Sorghum bicolor*	*Arabidopsis thaliana*
Stereoselectivity	*S*	*S*	*R*	*R* or *S*	*S*	*R*
Substrate range	aliphatic/aromatic-aldehydes & ketones	aliphatic/aromatic-aldehydes & methyl ketones	aliphatic/aromatic-aldehydes & methyl ketones	aliphatic aldehydes & methyl ketones	aromatic aldehydes & methyl ketones	aliphatic/aromatic-aldehydes & methyl ketones
Sequence- and structural similarity	α/β-hydrolase	α/β-hydrolase	GMC-oxidoreductase	Zn-dependent ADH	Serine carboxy-peptidase	α/β-hydrolase

1.1.2 Traits and demand on an HNL in industrial processes

To date many protocols based on chemical catalysts are applied for the synthesis of cyanohydrins [14]. Hydroxynitrile lyases have process relevant traits to compete with chemical methods. As already mentioned before the enzymatic reaction with HNLs result cyanohydrins with high enantiomeric excess (ee) and depending on the enzyme show opposite stereo-preferences. Due to this feature the need of protecting group chemistry can be minimised and the resolution techniques e.g. chromatography and crystallisation can be avoided. HNLs' broad substrate acceptance (aromatic-, heteroaromatic hydrocarbons and aliphatic compounds) enables the production of specific chiral cyanohydrins, which can be further converted in a large series of synthetic products. To give enough quantities of the enzyme used in the process, HNLs can be produced on a fermentative scale with recombinant expression systems like *Escherichia coli, Pichia pastoris* or *Saccharomyces cerevisiae*.

Demands on an HNL in the industrial process encompass their activity and stability under operating conditions. The production of cyanohydrins is performed at low pH (3.5 - 5.0) and low temperature (0° - 5°) to suppress the unspecific and spontaneous addition of HCN to the substrate resulting in non enantiopure products. The stability of the enzyme is strongly affected by the applied working mode. For example a biphasic system requires organic solvents, which can enormously decrease enzyme activity and stability. Furthermore, stirring the emulsion generates shear forces, which are disadvantageous for the enzyme stability [12] [15].

To fit the operating specifications and to improve catalytic characteristics of HNLs a lot of effort has been made. HNLs have been investigated to overcome shortcomings either by enzyme engineering or reaction engineering [13] [16].

1.1.3 Hydroxynitrile lyases with a cupin fold

Cupins are a protein superfamily named after the Latin term 'Cupa' meaning small barrel.

All members of the superfamily consist of a β-barrel shape. Two motifs G-$(X)_5$-H-X-H-$(X)_{3,4}$-E-$(X)_6$-G (motif 1) and G-$(X)_5$-P-X-G-$(X)_2$-H-$(X)_3$-N (motif 2) correspond to two conserved β-strands. Typically three histidines and a glutamate of this highly conserved sequence are involved in the metal ion binding in the active site. Most of the enzymatic cupins have Fe bound in their active site, but also Ni, Cu, Zn and Mn are found as cofactors. The various cofactors enable a different type of reaction in the conserved β-barrel structure. Therefore, cupins exhibit an extensive functional diversity with enzymatic and non-enzymatic functions [17] [18] [19] [20].

Recently, bacterial cupins with hydroxynitrile lyase activity were found. In this study gene libraries of different endophytic bacteria isolated from potato were screened for HNL activity. The data showed that proteins with HNL activity could be found among cupin superfamily proteins of so far unknown function [10]. A closer biochemical and structural characterization was done for the HNL from acidobacterium *Granulicella tundricola* (*Gt*HNL). The heterologous expression was successfully achieved with yields up to 50 % total protein in *E. coli*. Investigations revealed the first tertiary protein structure of a cupin hydroxynitrile lyase, which exhibits (*R*)-selectivity (Figure 2) [21]. Metal analysis showed that the enzyme activity is Mn-dependent [22].

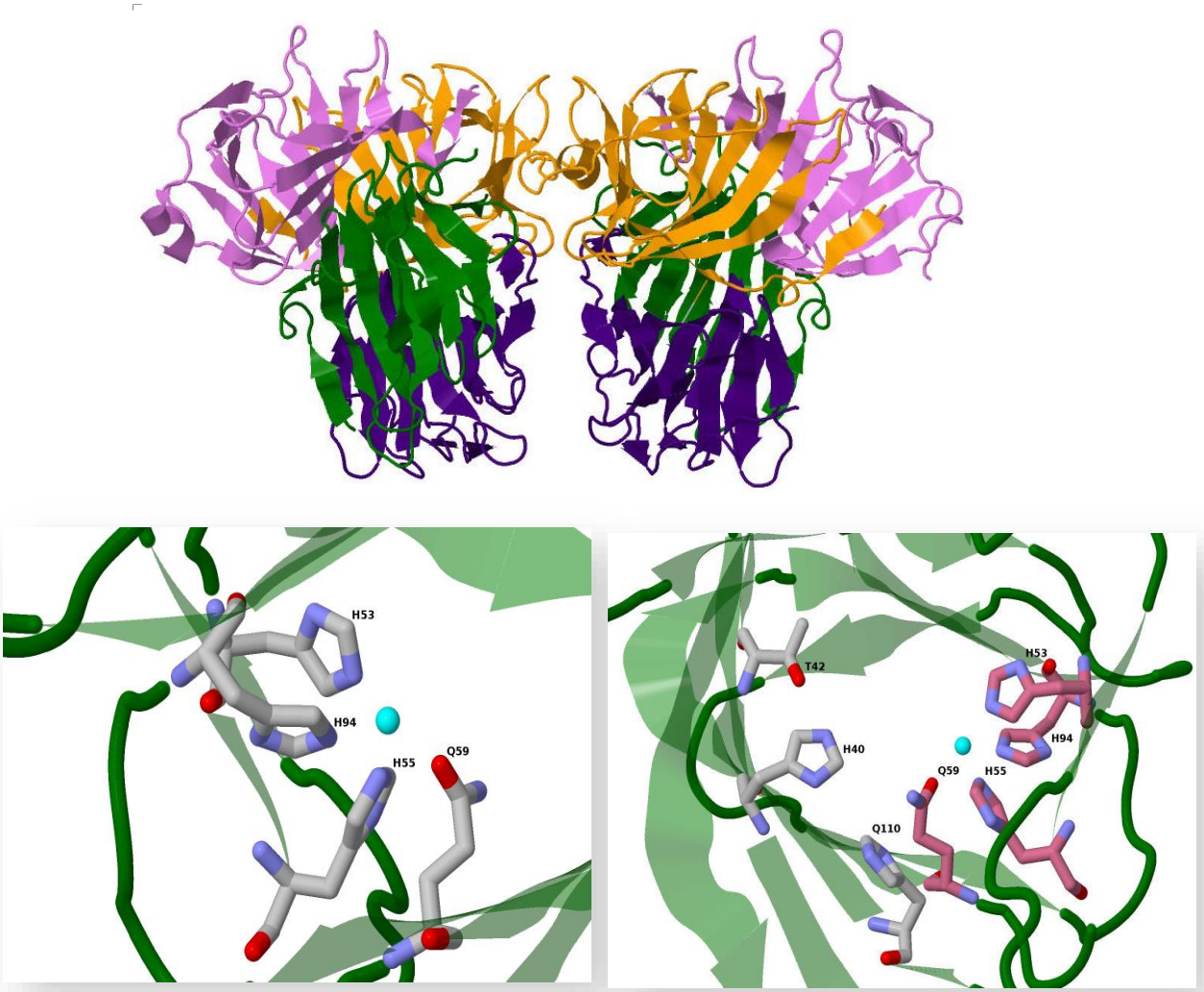


Figure 2: 3-D structure of GtHNL determined with crystallographic analysis. Top: Homotetramer of GtHNL (4bif). Bottom left: Amino acids involved in the Mn-binding at the active site of GtHNL. Bottom right: GtHNL-Tr. Exchanged amino acids compared with GtHNL (H40, T42, Q110) and metal binding amino acids (pale red). Data were obtained from PDB. The figures were prepared using Jmol: an open-source Java viewer for chemical structures in 3D. http://www.jmol.org/ [22].

*Gt*HNL was further the target for directed evolution and rational mutagenesis. The intention was to improve the enzymes activity, stability at low pH as well as the substrate scope. A combinatorial triple amino acid exchange (*Gt*HNL-A40HV42TQ110H, *Gt*HNL-Tr) in the active site (Figure 2) resulted in a variant with 490-fold higher specific activity against (*R*)-mandelonitrile. In the reversed synthesis reaction this enzyme achieved nearly full conversion with an excellent ee after 2 h. [1].

A highly similar cupin HNL from *Acidobacterium capsulatum* ATCC 51196 (*Ac*HNL) was recently discovered by database mining [23]. *Gt*HNL and *Ac*HNL share 77 % sequence identity and all active site amino acids are identical [1].

1.1.4 Catalytic promiscuity of hydroxynitrile lyases

To improve the applicability of HNLs it was considered to replace HCN in the synthesis reaction by nitroalkanes. This nitroaldolase activity, which is also known as Henry-reaction (Scheme 3) was the first example for an enzyme promiscuity of HNLs [24]. Biocatalytic promiscuity is well known for numerous enzymes and is part of the development of novel biocatalysts for organic synthesis purposes [25] [26] [27].

Scheme 3: Stereoselective addition of nitromethane to a carbonyl group in the presence of AcHNL.

The Henry-reaction is of huge importance in organic synthesis, since the products, the β-nitroalcohols, are key building blocks for the synthesis of complex and bioactive ingredients.

Griengl and coworkers examined the addition of nitromethane and nitroethane to various types of aldehydes in the presence of the HNL from *Hevea brasiliensis* (*Hb*HNL). The experiments were undertaken in a biphasic system due to reduced solubility of the nitroalkanes and aldehydes in water. The *Hb*HNL showed good conversion for the nitroalcohols with high enantiomeric excess. Further investigations of *Hb*HNL concerning reaction conditions, substrate scope and asymmetric retro-Henry-reaction were done by this group [28] [29]. Another HNL from the noncyanogenic plant *Arabidopsis thaliana* (*At*HNL) was also found to catalyse the Henry-reaction [30]. Some more examples from a biocatalysed Henry-reaction can be found in other enzyme classes e.g. hydrolases and transglutaminases. However, no enantioselectivity was reported so far [31] [32]. Although a number of biocatalytic protocols emerged in the past decade, this field of research is still in the fledgling stages [33].

Recently, a research team at the Austrian Center for Industrial Biotechnology in Graz (acib Graz) could prove the nitroaldolase activity for the cupin fold hydroxynitrile lyase from *A. capsulatum* (*Ac*HNL) and *G. tundricola* (Triple variant).

1.2 Protein engineering

Most of the enzymes used in industry are tailor-made enzymes for the applied process. Since enzymes rarely exhibit combinatorial properties required in organic synthesis, they are improved by protein engineering [34] [35].

In the late 70s recombinant DNA technology paved the way for protein engineering. In the last twenty years the methods and techniques in molecular biology, bioinformatics and related areas enormously improved. The progresses generated powerful tools for a cheap, easy and fast access to gene encoding sequences. The possibility to sequence heterologous genes, express and screen the proteins in just a few weeks pushed the field of protein engineering. In the past, the industrial process was engineered to match the requirements of the enzyme, today, the enzyme is designed to fit the process [36].

The main strategies in protein engineering are rational design and directed evolution or a combination thereof. In rational design knowledge about the structure and the relationship between structure and function is required. This information is obtained by three-dimensional structure analysis and bioinformatical tools. Directed evolution does not require profound knowledge about function and structure. Directed evolution can be defined as an iterative cycle of randomly created gene mutations and the screening of the variant library in a high-throughput fashion to identify mutants with improved properties. However, both methods need a suitable expression host and a sensitive detection system [37] [38].

Enzymes are engineered regarding their enantioselectivity, stability, process conditions, activity and specificity. Rational design focuses on interesting positions obtained from structure alignment studies or on mutations close to the active site, whereas directed evolution often results in mutations far from the active site. To increase enantioselectivity, substrate selectivity or alternate catalytic activity, mutations close to the active site are more efficient. Thus, the enzyme property to be improved always has to be considered, when the best mutagenesis method is chosen [39]. Furthermore, it is indispensable to create an intelligent diversity when applying protein engineering methods to potential industrial enzymes [40].

1.2.1 Selected examples of mutagenesis methods in protein engineering

To generate random libraries the most often applied method is error prone polymerase chain reaction (epPCR). With this technique a DNA-polymerase without proof reading function (e.g. Taq-DNA polymerase) introduces mutations into the gene of interest under mutagenic conditions. The mutation rate is controlled either by increasing the $MgCl_2$ concentration, addition of $MnCl_2$ or by altering dNTP

availability [41]. The mixed plasmid library is then transformed into a suitable expression host and the heterologous expressed enzyme variants are screened for improvement. Variants showing a promising result are subject to further rounds of evolution. Directed evolution is limited by the size of the library that may contain wild type or many variants that are either non-functional or that accumulate non-beneficial mutations [42]. To avoid non-functional variants and to keep the size of the variant library in a range able to screen the mutation rate should be controlled [43].

If there is no high-throughput screening available smaller and smarter libraries should be applied. On the basis of rational design, which requires structural and functional data, specific mutations can be introduced into enzymes to increase the probability to find a beneficial mutation [44]. One of these techniques is the creation of a site saturation library. It involves amino acid exchanges at a specific position in the enzyme. The method entails the introduction of all of the 20 canonical amino acids with a focused library of variants. The currently most widely used procedure is the creation of a mixed gene library by applying PCR with NNK (N: adenine/cytosine/guanine/thymine; K: guanine/thymine; MNN reversed complement) degenerated primers, which encodes all canonical amino acids (20 amino acids, 32 codons). In the next step a QuikChange protocol can be used to introduce the mutated gene into the plasmid. To ensure the coverage of a site saturation library the size may vary between a few hundred variants [45] [46].

1.3 Aim of this master thesis

In the first part of the thesis, variants of *Gt*HNL and *Gt*HNL-Tr originating from site directed and site saturation mutagenesis were investigated. The site directed amino acid exchanges at specific positions (A40, F19, W120) in the active site, should provide further insights into the reaction mechanism. The influence of the specific amino acids on stereo-preferences and substrate acceptance were tested for the cyanogenesis and synthesis reaction.

Site saturation libraries of *Gt*HNL and *Gt*HNL-Tr at the positions F19 and W120 were examined towards their activity on (*R*)-2-chloro-, and (*S*)-2-chloromandelonitrile. The goal was to identify amino acids, which show best activity on chlorated mandelonitrile for both the (*R*) as well as for the (*S*) enantiomer.

The second part of this thesis focuses on the creation of a random library of *Ac*HNL, which shows nitroaldolase activity. Variants should be expressed in *E. coli* and screened with a medium-through put assay for improved activity. Furthermore seven selected variants of *Gt*HNL and *Ac*HNL were examined towards their nitroaldolase activity.

2 Experimental section

2.1 Material

2.1.1 Chemicals, enzymes and primers

This chapter shows tabular lists of all growth media (Table 2), buffers (Table 3), enzymes (Table 4), reagents for SDS-PAGE and agarose gel electrophoresis (Table 5), reagents for the cyanohydrin synthesis (Table 6), solvents and reagents (Table 7) and primers (Table 8) used in this work.

- Growth media

Table 2: Growth media and supplements for the cultivation of E. coli.

LB-medium, 20 g/L, (Lennox, Roth)	10 g/L tryptone, 5 g/L yeast extract, 5 g/L NaCl
LB-agar plates, 35 g/L (Lennox, Roth)	10 g/L tryptone, 5 g/L yeast extract, 5 g/L NaCl, 15 g/L agar,
2xTY medium, 31 g/L (Roth)	16 g/L tryptone, 10 g/L yeas extract, 5 g/L NaCl
SOC medium	20 g/L bacto tryptone, 0.58 g/L NaCl, 5 g/L bacto yeast extract, 2 g/L $MgCl_2$, 0.18 g/L KCl, 2.46 g/L $MgSO_4$, 3.46 g/L glucose
Kanamycin stock (Roth)	40 mg/mL: 0.8 g kanamycin/20 mL dH_2O, filter sterilised
$MnCl_2$ stock (Sigma Aldrich)	100 mM: 198 mg $MnCl_2.4H_2O$/10 mL dH_2O, filter sterilised
1 M IPTG stock (Biosynth)	4.76 g IPTG/20 mL dH_2O, filter sterilised

- Buffers

Table 3: Buffers and their composition.

5x ISO reaction buffer	6 mL end volume: 25% (w/v) PEG-8000 (Sigma), 500 mM Tris/Cl pH 7.5, 50 mM $MgCl_2$, 50 mM DTT, 1 mM dATP, 1 mM dCTP, 1 mM dGTP, 1 mM dTTP, 5 mM NAD, dH_2O,
10 mM potassium dihydrogen phosphate buffer pH 7.0	Solution A: 10 mM KH_2PO_4 Solution B: 10 mM Na_2HPO_4
10 mM sodium dihydrogen phosphate buffer pH 7.0	Solution A: 10 mM NaH_2PO_4 Solution B: 10 mM Na_2HPO_4
50 mM potassium dihydrogen phosphate buffer pH 6.0	Solution A: 50 mM KH_2PO_4 Solution B: 50 mM Na_2HPO_4
1 M potassium dihydrogen phosphate buffer pH 5.1	Solution A: 1 M KH_2PO_4 Solution B: 1 M Na_2HPO_4

0.3 M citrate phosphate buffer pH 2.5	Solution A: 0.3 M citric acid Solution B: 0.3 M Na_2HPO_4
0.1 M citrate phosphate buffer pH 5.5	Solution A: 0.1 M citric acid Solution B: 0.1 M K_2HPO_4
0.1 M MES oxalate buffer pH 5.5	Solution A: 0.1 M oxalic acid dehydrate Solution B: 0.1 M MES; pH adjusted with 1 M NaOH
3 mM oxalate buffer pH 2.5	3 mM oxalic acid dehydrate; pH adjusted with 1 M NaOH
3 mM citrate phosphate buffer pH 3.0	Solution A: 3 mM citric acid Solution B: 3 mM K_2HPO_4
1 mM HEPES buffer	2 mL 1 M HEPES buffer diluted in 2 L dH_2O
1 M sodium acetate buffer pH 4.0	1 M sodium acetate

All chemicals were purchased from Carl Roth GmbH, if not stated otherwise.

- Enzymes

Table 4: List of enzymes and the corresponding buffers.

DNA polymerases	Pfu Polymerase, 2.5 U/µL (Thermo Scientific, USA)
	Dream Taq Polymerase, 5 U/µl (Thermo Scientific, USA)
	Phusion® High-Fidelity DNA Polymerase, 2 U/µl (Thermo Scientific, USA)
Restriction enzymes	NdeI, 20 U/µl, 5'-CA^TATG-3' (New England Biolabs, USA)
	HindIII, 20 U/µL, 5'-A^AGCTT-3' (New England Biolabs USA)
	DpnI, 10 U/µL, 5'-Gm6A^TC-3' (Thermo Scientific, USA)
Other enzymes	Taq DNA ligase, 40 U/µL (New England Biolabs, USA)
	T5 exonuclease, 10 U/µL (BioZym, USA)
	Lysozyme stock solution, 35 mg diluted in 1 mL dH_2O (Sigma Aldrich),
	Benzonase®, 250 U/µL (Merck, Germany)
Buffers	10x CutSmart®Buffer (New England Biolabs, USA
	1x Buffer Tango (Thermo Scientific, USA) Pfu Buffer with $MgSO_4$ (Thermo Scientific, USA)
	5x Phusion HF Buffer (Thermo Scientific, USA)
	10x DreamTaq Buffer (Thermo Scientific, USA)

- SDS-PAGE analysis and agarose gel electrophoresis

Table 5: Reagents used for SDS-PAGE analysis and for agarose gel electrophoresis.

20x NuPAGE® MES SDS Running buffer (Thermo Scientific, USA)	Dilution of 20x MES concentrate 1:20 with dH_2O
PageRuler Prestained Protein Ladder (Thermo Scientific)	
NuPAGE® 4-12% Bis-Tris Pre-Cast gels (Life Technologies, USA)	
4x NuPAGE LDS sample buffer (Life Technologies, USA)	
SDS gel staining solution	2.5 g Brilliant Blue, 500 mL EtOH (abs.), 75 mL acetic acid (100%)
SDS gel destaining solution	75 mL acetic acid (100%), 200 mL EtOH (abs.), 725 mL dH_2O
Biozym LE Agarose (Biozym Biotech Trading, Austria)	
GeneRuler™ 1 kb Plus DNA Ladder (Thermo Scientific)	
1x TAE buffer (Sigma Aldrich, USA)	40 mM Tris-Acetate, 1 mM EDTA
6x DNA loading dye (Thermo Scientific)	
ethidium bromide	

- HNL catalysed cyanohydrin synthesis

Table 6: List of chemicals used in the HNL synthesis reaction.

Methyl tert-butyl ether (MBTE)	
1,3,5 Triisopropyl benzene	
Sodium cyanide (NaCN)	
Dichlormethane	
Acetic acid (anhydride)	
Pyridine	
Benzaldehyde (fresh distilled)	

All chemicals were purchased by Sigma-Aldrich (USA).

- Solvents and reagents

Table 7: List of other chemicals and supplier.

4-hydrazino-7-nitrobenzofurazan (NBDH)/EtOH solution	0.16 mg/mL were dissolved in ethanol (abs.)
5x Bio-Rad Protein Assay Dye Reagent Concentrate (Bio-Rad, USA)	diluted 1:5 with dH_2O
Triton™ X-100 (Sigma Aldrich)	
(*R*)-mandelonitrile (DSM, Austria)	
(*R*)-2-chloromandelonitrile (DSM, Austria)	
(*S*)-2-chloromandelonitrile (DSM, Austria)	
Nitromethane (Sigma Aldrich)	
Dimethylsulfoxid – DMSO (Sigma Aldrich)	
dNTPs: dATP, dGTP, dCTP, dTTP,	
$MgCl_2$ stock (Carl Roth)	25 mM dissolved in dH_2O
$MnCl_2$ stock (Sigma Aldrich)	10 mM dissolved in dH_2O
Chloroform (Carl Roth)	
Copper (II) ethyloacetate (Sigma Aldrich)	1 % (w/v) solution in chloroform
4,4'-methylenebis(N,N-dimethylaniline (tetrabase) (Sigma Aldrich)	1 % (w/v) solution in chloroform
Albumin Faktor V (Carl Roth)	

- Primers

Table 8: Nucleotide sequences of all used primers with their size, GC content and melting temperature.

name of primer	Sequence [5'→ 3']	length [bp]	GC %	Tm °C
Cup9-F19V_for	CCGGCAGATTGGGTTACCGGCACCG	25	68	67.7
Cup9-F19V_rev	CGGTGCCGGTAACCCAATCTGCCGG	25	68	67.7
Cup9-W120R_for	CGGTAAAGCAGTTGATAGGATGGAACATGTTAC	33	42.4	60.2
Cup9-W120R_rev	GTAACATGTTCCATCCTATCAACTGCTTTACCG	33	42.4	60.2
Cu9A40HV42T_for	ACTGGTTGCCGGTCATAGCACTACCTTTGAACCGGG	36	55.6	68.5
Cu9A40HV42T_rev	CCCGGTTCAAAGGTAGTGCTATGACCGGCAACCAGT	36	55.6	68.5
Cu9A40RV42T_for	ACTGGTTGCCGGTAGGAGCACTACCTTTGAACCGG	36	58.3	69.5

Cu9A40RV42T_rev	CCCGGTTCAAAGGTAGTGCTCCTACCGGCAACCAGT	36	58.3	69.5
synCup9(pET26)_for	AATAATTTTGTTTAACTTTAAGAAGGAGATATACAT ATGGAAATTAAACGTGTTGGTAGC	60	26.7	62.5
synCupin9(pET26)_rev	GGTGGTGGTGGTGGTGCTCGAGTGCGGCCGCAAGC TTTTAGCGACGATACTGTTCATCG	59	59.3	73.6
Cup9_F19X_for NNK	CCGGCAGATTGGNNKACCGGCACCG	25		
Cup9_F19X_rev MNN	CGGTGCCGGTMNNCCAATCTGCCGG	25		
Cup9_W120X_for NNK	CGGTAAAGCAGTTGATNNKATGGAACATGTTAC	33		
Cup9_W120X_rev MNN	GTAACATGTTCCATMNNATCAACTGCTTTACCG	33		
Cup9_mut_fw1	GTTTAACTTTAAGAAGGAGATATACAT	27	25.9	49.3
T7 term_rev	CCGCTGAGCAATAACTAGC	19	52.6	53.4

2.1.2 Instruments and materials

Table 9 lists all instruments, devices and materials used in this work.

Table 9: List of all used instruments and materials in this project.

Cultivation of *E. coli*	100 mL shaking flasks
	1 L shaking flasks with baffles
	96 deep well plates
	96 PS-micro titer plates, clear, sterile, F-bottom (Greiner bio-one, Germany)
Incubators	BINDER Kühlbrutschränke (Binder GmbH, Germany)
	Certomat BS-1 (Sartorius, Germany)
	HT Multitron II (Infors AG, Switzerland)
	Thermomixer*(R)* comfort, 1.5 mL (Eppendorf, Germany)
Cell disruption	Sonifier 450 (Branson Ultrasonic, USA)
Centrifuges	Eppendorf table top centrifuge 5810 R (Eppendorf, Germany)
	Eppendorf Micro centrifuge 5415 R (Eppendorf, Germany)
	Avanti^{TM} J-20 XP Centrifuge (Beckman Coulter, USA)
	JA-10 and JA-25.50 rotor (Beckman Coulter, USA)
Photometers and plate readers	BioPhotometer Plus (Eppendorf, Germany)

	Cary 100 UV-Vis (Agilent Technologies, USA)
	FLUOstar Omega (BMG Labtech, Germany)
	EON microplate spectrophotometer (Bio-Tek, USA)
Thermocycler	GeneAmp PCR System 2700 (Applied Biosystems, USA)
Scale	Kern EW 1500-2M (Kern, Germany)
Electroporation	MicroPulser™ (BIO-RAD, USA)
	Electroporation cuvettes
Electrophoresis	PowerPac™ Basic + Sub-Cell GT (BIO-RAD, USA)
	XCell SureLock™ Mini-Cell Electrophoresis NuPAGE Novex System (Invitrogen)
Gel-imaging	UV transiluminator
	G-Box imaging system (Syngene, UK)
Reaction tubes	Microcentrifuge tubes, 1.5 mL volume with lid (Greiner bio-one, Germany)
	PP tubes, sterile, 15 mL volume (Greiner bio-one, Germany)
	PP tubes, sterile, 50 mL volume (Greiner bio-one, Germany)
	12-tube strip with separate flat cap strip, 0.2 mL volume (Greiner bio-one, Germany)
Microplates	96 PS-micro titer plates, clear, V-bottom (Greiner bio-one, Germany)
	96 PS-micro titer plates, clear, F-bottom (Greiner bio-one, Germany)
	96 UV-Star® micropate, µClear®, clear (Greiner bio-one, Germany)
Other devices	Nano Drop 200c Spectrometer (PEQLAB Biotechnologie, Germany)
	Stop watch
	Titramax 1000 (Heidolph, Germany)
Other material	Glass plates
	Whatman No.1 filter paper (supplier VWR, Austria)
	PP cuvettes 10 x 4 x 45 mm (Sarsted, Germany)
	Biodyne A nylon membranes (Pall, USA)
	Quartz cuvettes (Hellma GmbH, Germany)
	Mosquito net
	SilverSeal™ (Greiner bio-one, Germany)
	Rotilabo® SealPlate®, Polyester, sterile, 50 µm (Roth, Germany)
	Wizard® SV Gel and PCR Clean-Up System (Promega, USA)

	GeneJET Plasmid Miniprep Kit (Thermo Scientific, USA)
Cyanohydrin synthesis	6890N Gas chromatograph, Varian CP7503, CP-Chirasil-DEX CB column (25 m x 320 µm x 0.25 µm) (Agilent technologies, USA) PAL autosampler (CTC Analytics AG, Switzerland)

2.1.3 Strains and vectors

The *E. coli* strains used in this work were *E. coli* Top10F' and *E. coli* BL21-Gold(DE3).

E. coli Top 10 F' was used for standard cloning methods since this strain allows a stable replication of high copy number plasmids. The transformation efficiency was between 1 x 10^6-10^8 cfu/µg depending on the batch. Information about the genotype of this strain can be found on the web page of *Life Technologies* [47].

E. coli BL21-Gold(DE3) was used for protein expression. A detailed description of the strain can be found in the instruction manual from *Agilent Technologies* [48].

The vector pET26b(+) (Figure 4) was used for all cloning methods in this work. The vector contains a resistance gene for kanamycin, which is used for selective growth on kanamycin containing media. The replication of the plasmid in *E. coli* is enabled by the f1 origin of replication. The restriction sites of NdeI and HindIII, which were used in this work, are flanked by the T7 promoter and T7 terminator. The induction was carried out with IPTG under the control of the T7 lac promoter. The detailed description of pET vectors and their properties can be found in the user protocol from *Merck Millipore* [49].

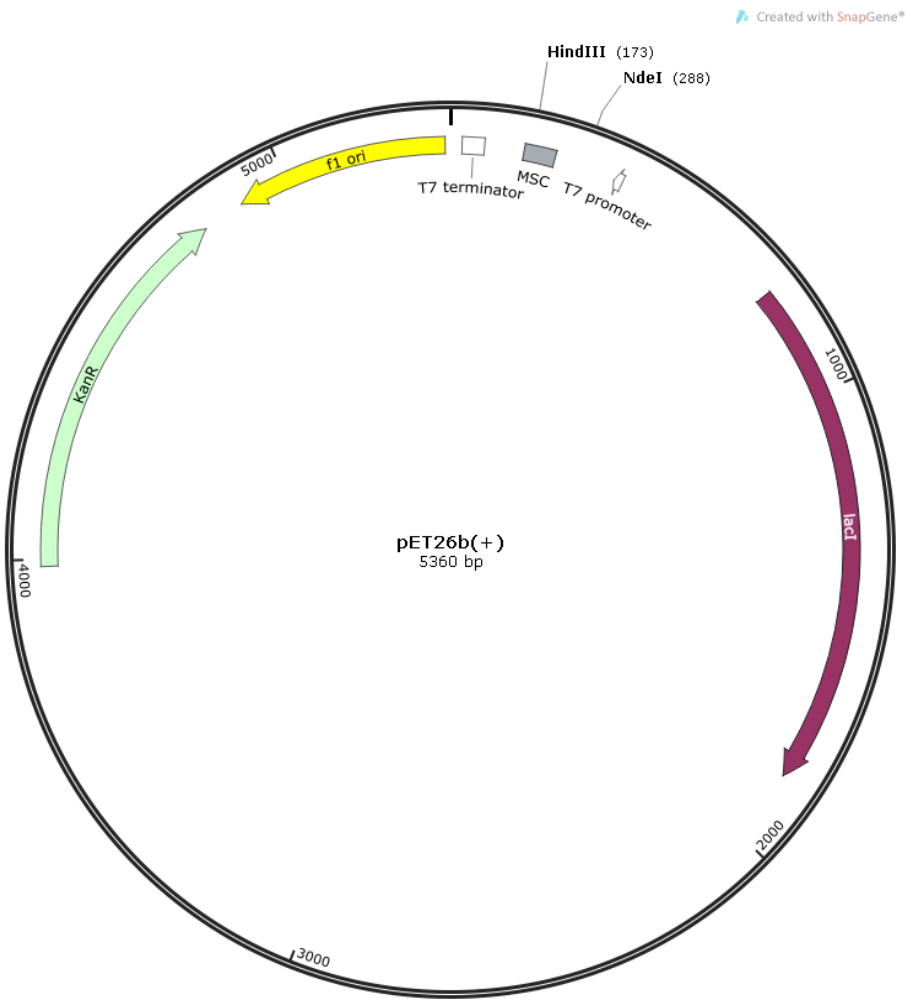


Figure 4: Expression vector pET26b(+). The vector provides a T7 promotor and terminator, a lacI coding sequence, as well as a kanamycin resistance marker and an origin of replication for E. coli. Restriction sites HindIII and NdeI are indicated in the vector. The graph was created with SnapGene software (from GSL Biotech; available at snapgene.com.)

2.2 General methods in molecular biology

2.2.1 Cell cultivation and protein expression

2.2.1.1 Cultivation and cell disruption by sonication

A pre-culture was obtained by using one single colony from the expression host *E. coli* BL21-Gold(DE3) harbouring pET26b(+) containing the gene of interest to inoculate 50 mL LB-medium with supplemented kanamycin (final concentration 40 µg/mL). This pre-culture was incubated at 37°C at 120 rpm overnight. The main culture was inoculated with 4 mL pre-culture and cells were grown in 1 L baffled flasks with 400 mL LB-medium supplemented with kanamycin (final concentration 40 µg/mL). Moreover, $MnCl_2$ (0.1 mM) was added to the main cultures at the inoculation time. Cells were incubated at 37°C, 120 rpm until they reached the optical density of 0.6. - 0.8 OD_{600} (λ = 600 nm). Expression of recombinant protein was initiated by addition of IPTG (0.5 mM) to the main culture and cultivation was continued at 25 °C, 120 rpm overnight.

Glycerol stocks were prepared by using 1 mL of the pre-cultures by adding 0.5 mL 50 % glycerol to the cells. The glycerol stocks were frozen at – 20 °C.

The main culture was harvested at 3600 g for 15 min at 4 °C in a JA-10 rotor. The pellet was weighed and used directly or stored at -20 °C. In any case the pellet was resuspended in 25 mL cold buffer and cells were disrupted by sonication for 6 min on ice (80 % duty cycle, 70 % output). Cell-free lysates were obtained by centrifugation for an hour at 48.250 g in a JA 25.50 rotor. For cyanohydrin synthesis, cleared lysates were concentrated to 50 mg/mL with Vivaspin 20 centrifugal concentrators. The lysates were stored at -20 °C until use.

2.2.1.2 Small-scale cultivation in Deep Well Plates (DWP)

96-well deep well plates were filled with 750 µL LB-medium/well supplemented with kanamycin (final concentration 40 µg/mL). The pre-culture was inoculated with the expression host *E. coli* BL21-Gold(DE3) harbouring pET26b(+) with the gene of interest. Inoculation was performed with a 96-pin stamp using cells from glycerol stocks stored in MTPs. The plates were incubated at 37 °C, 320 rpm and 80 % humidity overnight. For the main culture each well was filled with 750 µL LB-medium (kanamycin 40 µg/mL, 0.1 mM $MnCl_2$) and inoculated with 25 µL of the pre-cultures. Before inducing the protein expression, cells were grown further for 6 h at 37 °C, 320 rpm and 80 % humidity. Induction was carried out by adding IPTG (0.5 mM) to each well and cultivation was continued at 25 °C, 320 rpm and 80 % humidity for 15-16 h. Cells were harvested at 3220 g and plates containing the pellets were frozen at -20°C.

The frozen cells were lysed by adding 300 µL lysis buffer stock solution (for 4 DWP: 125 mL potassium phosphate buffer pH 7.0, 1 mL lysozyme stock (35 mg/mL), 2 µL benzonase, 250 µL Triton™ X-100) per well and lysis was continued for 1 h at 1200 rpm. The lysate was cleared from cell debris by centrifugation at 3220 g for 20 min. The supernatant was used for further investigations.

2.2.1.3 Small-scale cultivation in Micro Titer Plates (MTP) for glycerol stocks

Micro-scale cultivation of *E. coli* BL21-Gold(DE3) was performed in sterile 96 – micro titer plates. Each well was filled with 100 µL LB-medium supplemented with kanamycin (40 µg/mL) and inoculated with single colonies from LB-agar plates with a sterile toothpick. The plates were incubated at 37 °C overnight. To prevent evaporation of the medium, the plates were placed in a closed box filled with some water. Glycerol stocks were obtained by adding 50 µL of 50 % glycerol to each well. The plates were sealed with SilverSeal™ and stored at -20°C.

2.2.2 Preparation of electro competent cells

Overnight culture with a volume of 50 mL 2xTY medium was inoculated with one single colony of *E. coli* Top10F'/BL21-Gold(DE3) and incubated at 37 °C at 200 rpm overnight. For the main culture 2 L baffled flasks with 400 mL 2xTY medium were used. The main culture was inoculated with 4 mL of the overnight culture and incubated at 37 °C until it reached an optical density (OD_{600}) of 0.7-0.8. Then the cells were cooled on ice for 1 h and harvested at 3000 g for 10 min at 4 °C in a JA-10 rotor. The remaining medium was removed and the pellet resuspended in 1 mM HEPES buffer. The cell suspension was centrifuged at 4000 g at 4 °C for 10 min. Again the pellet was resuspended in 1 mM HEPES buffer and before a further centrifugation step 10 % glycerol was added to the cell suspension. The cell suspension was centrifuged for 20 min at 4500 g and 4°C. In the last step the cell pellet was resuspended in 10 % glycerol and centrifuged one more time for 15 min at 5000 g at 4 °C. The pellet was resuspended in 4 mL 10 % glycerol and aliquots of 80 µL electro competent cells were stored at - 80 °C.

2.2.3 Protein analysis with SDS-PAGE

For the sample preparation 15 µL of cleared lysate with a protein concentration of 1 mg/mL were mixed with 5 µl NuPAGE 4x loading dye. The samples were incubated at 93 °C for 10 min. The polyacrylamide gel (NuPAGE® Bis-Tris Pre-Cast gel) was positioned in the gel chamber, which was filled with 1x MES buffer. 10 µL of the denatured samples were centrifuged for one minute and then used directly to load the gel. 3 µL of NuPage Prestained Protein ladder were used as standard (Figure 5). To start the electrophoresis the program was set to a running time of 35 min, 200 V, 120 mA, 25.0 W.

To visualise the separated proteins, the polyacrylamide gel was covered with the staining solution Coomassie Brilliant Blue and after a short heating up in the microwave the gel was incubated for 20 min with shaking. The polyacrylamide gel was then decoloured at least for 2 h with the destaining solution containing ethanol and acetic acid at 500 rpm. The visualised protein bands were imaged with the G-box imaging system.

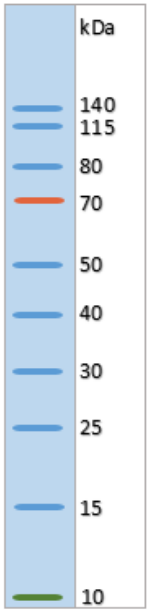


Figure 5: Migration pattern of Page Ruler Prestained Protein Ladder in MES buffer. Protein sizes are given in kDa.

2.2.4 Determination of the protein concentration with the Bio-Rad Protein Assay

The Bradford method for the quantification of proteins is based on the ability of Coomassie-Brilliant-Blue-250 (Bio-Rad concentrate) to build a complex with proteins, which absorbs at 595 nm. The absorption of the colour complex is increasing with the concentration of the protein. To quantify the total protein concentration of cleared lysate a standard curve (0.24 mg/mL, 0.21 mg/mL, 0.18 mg/mL, 0.15 mg/mL, 0.12 mg/ml, 0.12 mg/mL, 0.09 mg/mL, 0.06 mg/mL, 0.03 mg/mL) with Albumin Faktor V was obtained and

samples with unknown protein concentration were calculated with the linear equation. All lysates were prepared by diluting them 1:50 (concentrated samples 1:500) in the sample buffer at least in duplicate. 50 µL of these dilutions were mixed with 950 µL Bio-Rad Protein Assay Dye Reagent, which was diluted before 1:5 with dH_2O. After vortexing and an incubation time of 10 min the absorbance was measured at 595 nm with the photometer.

2.2.5 Agarose gel electrophorese

To prepare the agarose gel, 2 g of agarose were melted in 200 mL 1 x TAE buffer after a short incubation in the microwave. The hot agarose was then cooled down and one drop of EtBr was added. The liquid agarose was then poured in the gel chamber and cured for 20 min. The chamber was then filled with 1 x TAE buffer. The DNA samples were mixed with 6 x loading dye and then loaded onto the gel. To estimate the DNA size of the samples 5 µl of the 1 kb Plus DNA Ladder (Figure 6) were loaded on the gel. The electrophoresis was started by applying a voltage of 90 V and the running time was set to 2 h for a preparative agarose gel. The separated DNA bands were visualised with the UV transiluminator system and in case of a preparative gel cut off the gel and cleaned up.

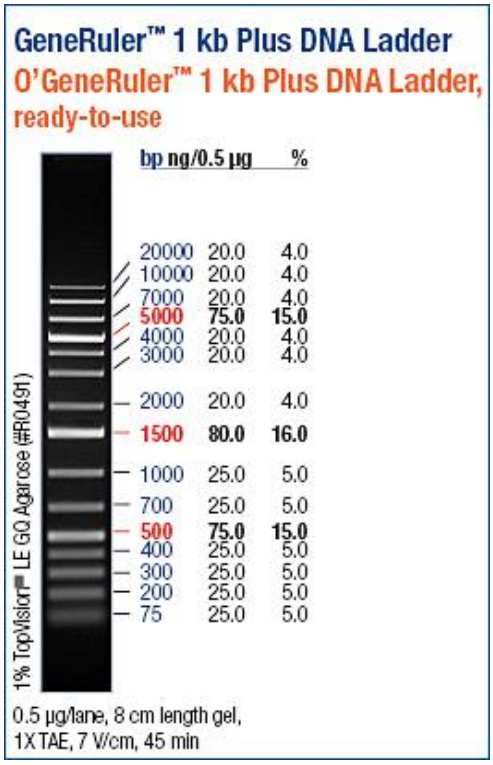


Figure 6: DNA size standard GeneRuler™ 1 kb Plus DNA Ladder.

2.2.6 Plasmid isolation protocol and cleanup system for DNA from agarose gel

Plasmid isolation

Plasmids were isolated from cells by using the GeneJET Plasmid Miniprep Kit. The procedure was done as described in the supplier's manual. Two steps of the protocol were altered. Step 10: To remove residual Wash Solution it was centrifuged for 2 min. Step 11: Samples were diluted in 50 µL H_2O (Fresenius). The plasmid concentration was determined with the Nano Drop 200 c Spectrometer.

DNA cleaning up from agarose gel

DNA was cleaned from the agarose gel as described in the manual from Wizard® SV Gel and PCR Clean-Up System. Other than described in the manual the DNA samples were not eluted with 50 µL nuclease-free water but with 30 µL H_2O (Fresenius). The DNA concentration was determined with the Nano Drop 200c Spectrometer.

2.2.7 Vector linearisation and template preparation of *Gt*HNL, *Gt*HNL-Tr and *Ac*HNL

The plasmid pET26b(+) was linearised with the restriction enzymes NdeI and HindIII. The restriction enzyme reactions were carried out by applying the standard protocol from New England Biolabs [50].

Table 10: Reaction conditions for the restriction of pET26b(+) with NdeI and HindIII.

restriction enzyme	7.5 U
plasmid DNA	1 µg
10x CutSmart buffer	5 µL
total reaction volume	50 µL
incubation temperature	37 °C
incubation time	overnight

The DNA of the restriction reaction was separated with a preparative agarose gel electrophorese. Bands at the desired size were cut out of the gel and cleaned up from the gel. The restriction was checked with a retransformation into *E. coli* Top10F' cells. Therefore the cells were plated onto selective LB-agar plates (kanamycin 40 µg/mL) and incubated at 37 °C overnight. The restriction was successful, if there were no colonies observed, since the plasmid with the selection marker was restricted.

2.2.8 Gibson cloning

The technique allows a rapid cloning without restriction enzymes and ligase by in vitro recombination [51]. The 5 x isothermal reaction buffer was prepared with a final volume of 6 mL. All components listed in Table 11 were mixed well by vortexing and then 100 µL were aliquoted and stored at -20 °C until use.

Table 11: Components of the 5x ISO buffer for the Gibson assembly.

component	stock solution	volume, amount
25 % (w/v) PEG-8000	PEG-8000	1.5 g
500 mM Tris/Cl pH 7.5	1 M Tris/Cl pH 7.5	3000 µL
50 mM $MgCl_2$	2 M $MgCl_2$	150 µL
50 mM DTT	1 M DTT	300 µL
1 mM dATP	100 mM dATP	60 µL
1 mM dCTP	100 mM dCTP	60 µL
1 mM dGTP	100 mM dGTP	60 µL
1 mM dTTP	100 mM dTTP	60 µL
5 mM NAD	100 mM NAD	300 µL
	Sterile H_2O	up to 6000 µL
V_{tot}		6000 µL

The assembly master mix was prepared by combining the components listed in Table 12.

Table 12: Components of the assembly master mix for the Gibson cloning.

component	volume to add
5 x ISO reaction buffer	320 µL
T5 exonuclease, 10 U/µL	0.64 µL
Phusion® High-Fidelity DNA Polymerase, 2 U/µL	20 µL
Taq DNA ligase, 40 U/µL	160 µL
sterile dH_2O	699.36 µL
V_{tot}	1200 µL

After well mixing, aliquots of 15 µL were stored at -20 °C until use.

For the Gibson assembly were used at least 100 ng DNA of each fragment in equimolar amounts. To calculate the amount of the insert, the following formula was used:

$$\frac{ng\ vector\ \times bp\ size\ of\ interest}{bp\ size\ vector} = ng\ insert$$

5 µL of the DNA to be assembled were added to 15 µL of the assembly master mix and incubated at 50 °C for 60 min.

Before transformation the Gibson assembly mix was desalted. For this procedure a membrane was put in a plate with water and the assembly mix was pipetted on top. After 15 – 30 min the assembly mix was ready for transformation.

2.2.9 Transformation of electro competent *E. coli* cells

For the transformation of *E. coli* Top10F', electro competent cells were carefully thawed on ice. 80 µL of these competent cells were mixed with 5 µL of the Gibson assembly mix and transferred in sterile, pre-cooled 0.2 cm electroporation cuvettes. For the transformation of electro competent *E. coli* BL21-Gold(DE3) cells 1 ng/µL plasmid DNA was used. The cells were pulsed with the electroporator by applying the program Ec_2 at 2.5 kV. For regeneration of the transformation mix 800 µL LB-medium were added and cells were incubated for 1 h at 350 rpm. After the regeneration, 1 x 100 µL and 1 x 50 µL of the transformation reaction were plated on LB-agar plates supplemented with kanamycin 40 µg/mL and incubated at 37 °C overnight.

2.3 Generation and screening of site directed and site saturation libraries of *Gt*HNL

2.3.1.1 Overlap extension PCR

Based on rational design, amino acids in the active site of *Gt*HNL, and *Gt*HNL-Tr were mutated. The targeted amino acid exchanges in the site specific mutagenesis experiment were A40RV42T and A40HV42T in *Gt*HNL and F19V and W120R in *Gt*HNL-Tr. For the site saturation experiment the positions

5'–ATGGAAATTAAACGTGTTGGTAGCCAGGCAAGCGGTAAAGGTCCGGCAGATTGGTTTACCGGCACCGT
TCGTATTGATCCGCTGTTTCAGGCACCGGATCCGGCACTGGTTGCCGGTGCAAGCGTTACCTTTGAACCGG
GTGCACGTACCGCATGGCATACCCATCCGCTGGGTCAGACCCTGATTGTTACCGCAGGTTGTGGTTGGGCA
CAGCGTGAAGGTGGTGCAGTTGAAGAAATTCATCCGGGTGATGTTGTTTGGTTTAGTCCGGGTGAAAAAC
ATTGGCATGGTGCAGCACCGACCACCGCAATGACCCATCTGGCAATTCAGGAACGTCTGGACGGTAAAGC
AGTTGATTGGATGGAACATGTTACCGATGAACAGTATCGTCGCTAA–3'

Figure 7: Gene sequence of the hydroxynitrile lyase from G. tundricola. Mutated nucleotides of GtHNL-Tr in blue letters. Targets of site directed mutagenesis in grey and for site saturation libraries underlined.

F19 and W120 were chosen in both *Gt*HNL and *Gt*HNL-Tr. To insert the mutations into the coding sequence of *Gt*HNL (Figure 7) and *Gt*HNL-Tr an overlap extension PCR was performed.

The overlap extension PCR was carried out in two PCR steps (PCR 1 and PCR 2). The exact primer combinations for the site specific mutagenesis are listed in Table 13 and for the site saturation mutagenesis in Table 14. In the first PCR (Table 15) the mutagenic forward and reversed primers, which were complementary to each other were used in two separated reactions (reaction 1 and reaction 2). The template was either pET26b(+)_*Gt*HNL or pET26b(+)_*Gt*HNL-Tr. In the second PCR the products from the first PCR were combined with outer primers to give the full length of the product.

Table 13: Primer combination for the insertion of site specific mutations with overlap extension PCR.

aa-exchange	PCR 1	template	mutagenic primer	outer primer
A40HV42T	reaction 1	pET26b(+)_*Gt*HNL	Cu9A40HV42T_for	synCup9(pET26)_rev
	reaction 2	pET26b(+)_*Gt*HNL	Cu9A40HV42T_rev	synCup9(pET26)_for
A40RV42T	reaction 1	pET26b(+)_*Gt*HNL	Cu9A40RV42T_for	synCup9(pET26)_rev
	reaction 2	pET26b(+)_*Gt*HNL	Cu9A40RV42T_rev	synCup9(pET26)_for
F19V	reaction 1	pET26b(+)_*Gt*HNL-Tr	Cup9-F19V_for	synCup9(pET26)_rev
	reaction 2	pET26b(+)_*Gt*HNL-Tr	Cup9-F19V_rev	synCup9(pET26)_for
W120R	reaction 1	pET26b(+)_*Gt*HNL-Tr	Cup9-W120R_for	synCup9(pET26)_rev
	reaction 2	pET26b(+)_*Gt*HNL-Tr	Cup9-W120R_rev	synCup9(pET26)_for

Table 14: Primer combination to create site saturation libraries with overlap extension PCR.

aa-exchange position	PCR 1	template	mutagenic primer	outer primer
F19X	reaction 1	pET26b(+)_*Gt*HNL	Cup9_F19X_for NNK	synCup9(pET26)_rev
	reaction 2	pET26b(+)_*Gt*HNL	Cup9_F19X_rev MNN	synCup9(pET26)_for
	reaction 1	pET26b(+)_*Gt*HNL-Tr	Cup9_F19X_for NNK	synCup9(pET26)_rev
	reaction 2	pET26b(+)_*Gt*HNL-Tr	Cup9_F19X_rev MNN	synCup9(pET26)_for
W120X	reaction 1	pET26b(+)_*Gt*HNL	Cup9_W120X_for NNK	synCup9(pET26)_rev
	reaction 2	pET26b(+)_*Gt*HNL	Cup9_W120X_rev MNN	synCup9(pET26)_for
	reaction 1	pET26b(+)_*Gt*HNL-Tr	Cup9_W120X_for NNK	synCup9(pET26)_rev
	reaction 2	pET26b(+)_*Gt*HNL-Tr	Cup9_W120X_rev MNN	synCup9(pET26)_for

Table 15: Components of the overlap extension PCR for the first PCR reaction.

H_2O		35.5 µL
template pET26b(+)_*Gt*HNL or *Gt*HNL_Tr	1 ng/µL	1 µL
forward primer or reverse primer	10 pmol/µL	1 µL
dNTPs	10 mM	1 µL
Phusion HF Buffer	5 x	10 µL
Phusion DNA Polymerase		0.5 µL
V_{tot}		50 µL

To use the PCR 1 product for PCR 2 the samples were loaded onto an agarose gel. After electrophoresis the PCR-fragments were cut off the gel at the correct size and purified. The samples were then ready to use in PCR 2. The reaction set up of PCR 2 is listed in Table 16.

Table 16: Reaction set up of the overlap extension PCR 2.

H_2O		31.5 µL
PCR 1 reaction product 1 for	10 pmol/µL	2.5 µL
PCR 1 reaction product 2 rev	10 pmol/µL	2.5 µL
synCup9(pET26)_for	10 pmol/µL	1 µL
synCup9(pET26)_rev	10 pmol/µL	1 µL

dNTPs	10 mM	1 µL
Phusion HF Buffer	5 x	10 µL
Phusion DNA Polymerase		0.5 µL
V_{tot}		50 µL

PCR cycling programme for PCR 1 and PCR 2:

98°C 1 min

98°C 15 s
55°C 30 s } 25 cycles
72°C 15 s

72°C 7 min

4°C ∞

After the PCR 2 reaction, the products were loaded on a preparative agarose gel for electrophoresis. After purification of the DNA from the agarose gel the concentration was determined with the Nano Drop 200c Spectrometer. The mutated fragments were cloned into the linearised pET26b(+) vector by applying the Gibson cloning protocol.

The plasmids originating from the site specific mutagenesis were transformed into *E. coli* Top10F' cells, purified and retransformed into the expression host *E. coli* BL21-Gold(DE3). The desired mutations in the gene sequence were confirmed by sequencing. The variants (*Gt*HNL_A40HV42T, *Gt*HNL_A40RV42T, *Gt*HNL-Tr_F19V, *Gt*HNL-Tr_W120R) were expressed by fermentation in 400 mL LB-medium supplemented with kanamycin (40 µg/mL). The cells were harvested and disrupted by sonication. The obtained cleared lysates were used for the analysis in SDS-PAGE, cyanogenesis and cyanohydrin synthesis.

In the case of site saturation mutagenesis, 2.5 µL of the obtained Gibson assembly were transformed into *E. coli* Top10F' cells. The transformants were used to inoculate an overnight culture of 5 mL LB-medium containing kanamycin (40 µg/mL). The cells of the overnight culture were harvested and the mixed plasmids were isolated and further transformed into *E. coli* BL21-Gold(DE3). To obtain single colonies aliquots of 50 µL and 100 µL were plated onto LB-Kan-agar plates and incubated at 37 °C overnight. Four colonies of each library were sequenced and evaluated for a randomised mutation at the desired position. 168 single colonies of each library were picked from agar plates and deposited in 96 well MTP containing LB broth. As references four colonies of pET26b(+), *Gt*HNL and *Gt*HNL-Tr were picked into each MTP. After incubation at 37° overnight and generation of cryo cultures, the MTPs served as master plates for the colony based screening for hydroxynitrile lyase activity on (*R*)-2-chloromandelonitrile and (*S*)-2-

chloromandelonitrile. Variants found in the colony-based screening were re-screened and promising hits were cultivated in 400 mL LB-medium to obtain enough lysate for SDS-PAGE analysis and for the HNL-Activity Assay.

2.3.1.2 Colony based filter assay for hydroxynitrile lyase activity

To screen the site saturation libraries of *Gt*HNL and *Gt*HNL-Tr, colonies from MTP glycerol stocks were stamped onto nylon membranes. The membranes were laid on LB-Kan-agar plates and incubated at 37 °C overnight. To express the enzyme variants the membranes were transferred to LB-Kan-agar plates containing IPTG (0.1 mM) for induction and $MnCl_2$ (0.1 mM). The plates were further incubated at room temperature overnight. The grown colonies with the expressed protein were then ready to screen.

The screening procedure was carried out as described in the acib Protocol: "Colorimetric screening for hydroxynitrile lyase activity" (Author: Romana Wiedner). The procedure was described before by Krammer et al. [52]. The colony based screening assay allows the detection of hydroxynitrile lyase activity with the aid of an HCN-sensitive detection paper. The principle of the test relies on the formation of a blue coloured complex.

Preparation of HCN-sensitive detection paper:

A 1 % (w/v) solution of copper (II) ethylacetoacetate and 1 % (w/v) solution of 4,4'-methylenbis (N,N-dimethyl-aniline) (tetra base) were prepared in chloroform and combined. Whatman No. 1 filter sheets were soaked with the prepared solution and dried under the hood. The pale green, almost colourless detection filters were stored until use in the fridge at 4°C.

Preparation of substrate solution:

(*R*)-2-chloromandelonitrile (10 mM) and (*S*)-2-chloromandelonitrile (20 mM) were dissolved in 40 % DMSO and 0.1 M citrate-phosphate buffer pH 3.5.

Prior to screening the membranes with the colonies were equilibrated for 15 min on a filter sheet soaked in 0.1 M citrate phosphate buffer at pH 3.5. The screening itself was performed as a sandwich-like assembly. First a filter sheet was soaked with substrate solution. The membranes with the colonies were transferred with colony side down onto the substrate filter. The build-up was then covered with a thin permeable mosquito net. Subsequently, the HCN sensitive detection paper was placed on the net and finally covered with a glass plate. The time until blue spots appeared on the HCN-detection paper was measured and photos were taken.

2.3.1.3 Cyanohydrin Synthesis

The cyanohydrin synthesis was carried out at the Institute for Organic Chemistry at the Technical University of Graz by Karin Reicher. The procedure is described in detail in the acib protocol from 29.02.2012: "Cyanohydrin synthesis with HCN in organic solvent". (Authors: Ivan Hajnal, Kerstin Steiner, Helmut Schwab).

The synthesis reaction was performed in a two-phase system. The organic phase with a volume of 1 mL contained 0.467 M benzaldehyde, 2 v/v % of 1,2,5 Triisopropylbenzene (internal standard), 2 M HCN and as organic solvent MTBE. The aqueous phase contained 450 µL cleared lysate (~ 35 mg/mL total protein), which was pre-equilibrated by adding 50 µL 1 M sodium acetate buffer pH 4.0. The reaction was carried out in 2 mL reaction tubes on a thermomixer shaking at 1000 rpm and 4°C. After two hours 50 µL of the organic phase were withdrawn and used for derivatization at room temperature for 1 h. The derivatization solution contained 850 µL dichloromethane, 100 µL acetic acid anhydride, 50 µL pyridine. The acetylated cyanohydrins were analyzed with gas chromatography (GC). The parameters for the GC are listed in Table 17.

Table 17: Parameters of the GC analysis of acetylated cyanohydrins after HNL catalysed synthesis reaction.

	GC settings
Carrier gas:	Constant pressure mode at 100 kPA H_2
Injector:	250 °C
Temperature programme:	60 °C, 10 °C/min to 140°C

The retention time for the internal standard was 8.5 min, for benzaldehyde 4. 5 min, for (*R*)-cyanohydrin acetate 9.4 min and for (*S*)-cyanohydrin acetate 10.2 min.

Data obtained from the gas chromatography were analysed with the Agilent ChemStation software. Conversion and enantiomeric excess (ee) values were calculated with the formulas:

$$100 * \left(1 - \frac{\frac{benzaldehyde}{internal\ standard}(t_{2h})}{\frac{benzaldehyde}{internal\ standard}(t_{0h})}\right) = conversion\ [\%]$$

$$100 * \frac{A - B}{A + B} = ee\ [\%]$$

A= (*R*) enantiomer; B= (*S*) enantiomer

2.3.1.4 HNL-Activity Assay

The cyanogenesis activity of the HNLs was determined spectrophotometrically in 96 well UV-Star® MTPs by measuring the absorption of released benzaldehyde (280 nm) or 2-chlorobenzaldehyde (300 nm). The substrates (*R*)-mandelonitrile, (*R*)-2-chloromandelonitrile and (*S*)-2-chloromandelonitrile were dissolved depending on the respective buffer system either in 3 mM citrate phosphate buffer pH 3.0 or 3 mM oxalate buffer pH 2.5. To increase the solubility of the chlorinated substrates 40 % DMSO were added. Each reaction was tested separately with two buffer systems, either with 0.1 M citrate phosphate buffer pH 5.5 or 0.1 MES oxalate buffer at pH 5.5.

The reaction set up contained 20 µL cleared lysate between a total protein quantity of 2-500 µg per well, 50 µL substrate stock solution with a final concentration of 18 mM and 130 µL respective reaction buffer. For the blank, the same amount of reaction buffer was used instead of lysate. Before the reaction start the temperature of the plate reader was set to 25 °C. The reaction was started by adding substrate to the other components. The absorption of the released aldehyde was measured over 10 min. The obtained slopes were used to compare the enzyme activity of the variants.

2.4 Random mutagenesis and screening of *Ac*HNL

2.4.1.1 Randomisation of AcHNL with error prone PCR

To introduce mutations randomly into the gene sequence (396 nt) of *Ac*HNL (Figure 8) the error prone PCR (epPCR) protocol was applied. In the epPCR a mutagenic mega primer of the whole gene was created, which was then introduced into the plasmid in a second step with the QuikChange PCR (qcPCR). The 5'-end primer used in the first PCR was complementary to the bases in the pET26b(+)-vector before the start codon. The 3'-end primer was complementary to the T7 terminator region of the vector, which is located some bases downstream of the stop codon.

5'–GTTTAACTTTAAGAAGGAGATATACATATGCAGATTACCCGTAATGGTAGCCAGCCGAGCGGTCGTGGTC
CGGCAGAATATTTCACCGGCACCGTTCGTGTTGATCCGCTGTTTGCAGCACCGGAACCGGCACGTGTTGCCGG
TGCAAGCGTTACCTTTGAACCGGGTGCACGTACCGCATGGCATACCCATCCGCTGGGTCAGACCCTGATTGTT
ACCAGCGGTTGTGGTCGTGTTCAGCGTGAAGGTGGTCCGGTTGAAGAAATTCGTCCGGGTGATGTTGTTTGG
TTTACACCGGGTGAACGTCATTGGCATGGTGCAAGCCCGAGCACCGCAATGACCCATATTGCAATTCAAGAAA
AACTGGATGGCAAAGTTGTTGAATGGCTGGAACATGTTACCGATGCAGAATATGCAGGTTAAAAGCTTGCGG
CCGCACTCGAGCACCACCACCACCACCACTGAGATCCGGCTGCTAACAAAGCCCGAAAGGAAGCTGAGTTGG
CTGCTGCCACCGCTGAGCAATAACTAGGC–3'

Figure 8: Nucleotide sequence of AcHNL flanked by fragments of pET26b(+). The gene encoding sequence is highlighted in dark grey. The primer binding sites for the error prone PCR are highlighted in light grey.

The epPCR was carried out with the enzyme Dream Taq Polymerase, which lacks proof reading activity. To cause misincorporated nucleotides during the polymerase reaction, mutagenic conditions were applied by using unusual metal salt concentrations. The final concentration in the reaction mixture of $MgCl_2$ was 5 mM and 0.5 mM for $MnCl_2$. All components of the epPCR are listed in Table 18.

Table 18: Error prone PCR reaction set up to generate a mutated mega primer of AcHNL.

H_2O		26,5 µL
template pET26b+_*Ac*HNL	1 ng/µL	1 µL
forward primer Cup9_mut_fw1	10 pmol/µL	2 µL
reverse primer T7 term	10 pmol/µL	2 µL
dNTPs	10 mM	1 µL
Dream Taq buffer	10 x	5 µL
$MgCl_2$	25 mM	10 µL
$MnCl_2$	2.5 mM	2 µL
Dream Taq Polymerase		0.5 µL

V_{tot}		50 µL

Temperature programme for the epPCR:

95°C 2 min

95°C 30 s
55°C 30 s } 35 cycles
72°C 20 s

72°C 10 min

4°C ∞

To purify the mega primer the PCR sample was run on a preparative agarose gel, cut off the gel and cleaned up. The DNA concentration was determined and the mutated mega primer was further used in the qcPCR. This PCR enabled the integration of the randomly mutated fragment into the expression vector pET26b(+). All components used for the qcPCR are listed in Table 19.

Table 19: Reaction mixture of the QuikChange PCR to integrate the mutated mega primer in the vector.

H_2O		37 µL
template pET26b+_*Ac*HNL	50 ng/µL	1 µL
megaprimer	300 ng	4-5 µL
dNTPs	10 mM	1 µL
Pfu buffer $MgSO_4$ Thermo Scientific	10 x	5 µL
Pfu Polymerase Thermo Scientific		1 µL
V_{tot}		50 µL

The cycling programme for the qcPCR:

95°C 2 min

95°C 1 min
65°C 30 s } 30 cycles
72°C 10 min

72°C 10 min

4°C ∞

After qcPCR the sample was again purified with electrophoresis and cleaned up from the agarose gel. In the next step the template strand of the double stranded plasmid was digested with DpnI by using 3 µL of Tango buffer and 0.5 µL restriction enzyme DpnI for 1 h at 37°C. Digestion was stopped by incubating the mixture for 20 min at 80°C. 4 µL of the sample were desalted and transformed into *E. coli* Top10F' cells. The plasmids harbouring the mutated gene of *Ac*HNL were retransformed in *E. coli* BL21-Gold(DE3) and aliquots of 50 µL and 100 µL were plated on LB-Kan-agar plates. To obtain glycerol stocks of the random library, single colonies from the LB-Kan-agar plates were picked into each well of 96 well micro titer plates and cultivated in small scale. In total the random library represented 5760 clones of *Ac*HNL. To determine the mutation rate, six randomly selected variants were sequenced.

2.4.1.2 Screening of AcHNL random library for nitroaldolase activity

To obtain a sufficient enzyme concentration for the screening, the fermentation of the variants was performed in deep well plates. To inoculate DWPs, *E. coli* BL21-Gold(DE3) cells hosting the random library *of Ac*HNL were stamped from MTP glycerol stocks. After expression of the enzyme, harvesting of the cells and cell lysis as described above, the obtained crude lysate was used in the enzymatic nitroaldol reaction. The enzymatic reaction was carried out in a 96 well microtiter plate (F-form, PS). To 90 µl of crude lysate, 10 µL 1 M potassium phosphate buffer pH 5.1 were added. The reaction was started by adding 100 µL substrate stock solution consisting of benzaldehyde and nitromethane (final concentration benzaldehyde: 10 mM and nitromethane: 0.6 M) diluted in 50 mM potassium phosphate buffer pH 6.0. The reaction plate was sealed with a sealing tape. One hour after reaction start the remaining benzaldehyde concentration was determined. Therefore 10 µL of the enzymatic reaction mix were diluted in 140 µL 0.3 M citrate phosphate buffer pH 2.5 using a 96 well micro titer plate (V-form, PS). The detection solution was prepared by dissolving 0.16 mg/mL 4-hydrazino-7-nitrobenzofurazan (NBDH) in > 99.9 % ethanol by sonication in a water bath for 15 min. Remaining solids were removed by centrifugation at 3220 g for 5 min. To start the detection of benzaldehyde, 50 µL of the fluorogenic NBDH/Ethanol solution were added to the diluted reaction mixture in each well. The fluorescence intensity was measured in the plate reader (FLUOstar Omega) by applying an excitation filter at 485 nm and an emission filter at 520 nm. Further settings for the plate reader included the cycle number of 1, cycle time of 8 s, number of flashes/well 3 and a gain of 1414.

Variants, which showed better nitroaldolase activity than the wild type *Ac*HNL, were re-screened. Subsequently, the gene sequence of 39 *Ac*HNL variants was determined by sequencing.

3 Results

3.1 Site directed mutagenesis

3.1.1.1 SDS-PAGE

Single and double mutants of *Gt*HNL (A40HV42T, A40RV42T) and *Gt*HNL-Tr (F19V, W120R) generated with PCR were expressed in shake flasks for activity analysis in the cleavage and synthesis reaction. The soluble total protein fraction of the variants was analysed with SDS-PAGE (Figure 9). The successfully overexpressed variants of *Gt*HNL and *Gt*HNL-Tr showed protein bands at the expected size of 15 kDA, whereas the negative control pET26b(+) did not show a band. A dimer and traces of a tetramer can be seen for the double mutant *Gt*HNL_A40RV42T (lane 4).

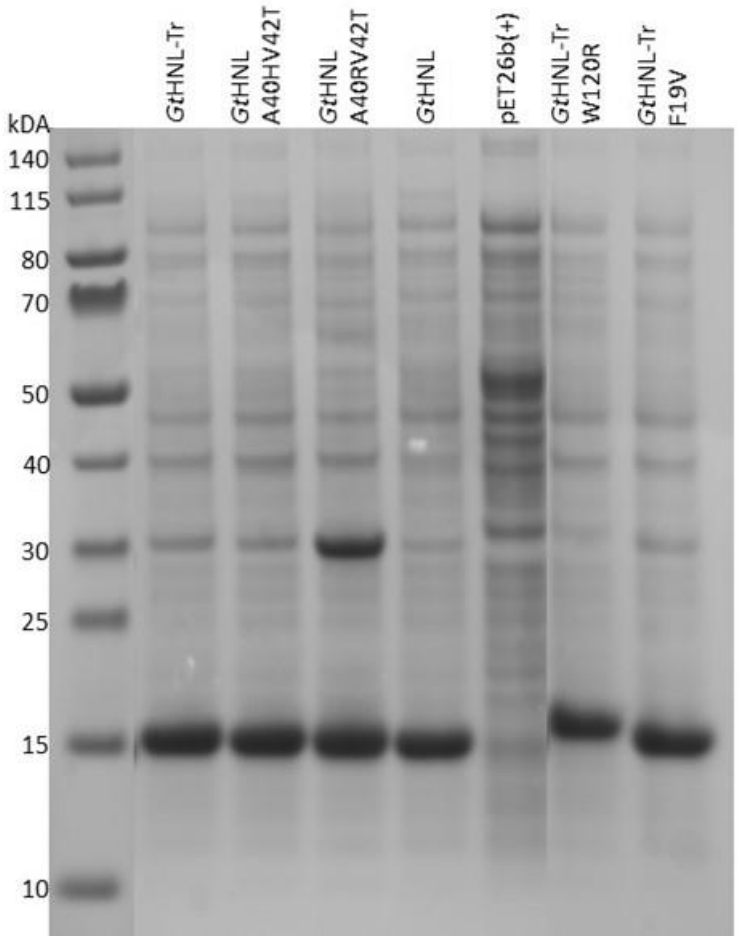


Figure 9: SDS gel of double and single mutants of GtHNL and GtHNL-Tr. Lane 1: Page Ruler Prestained Protein Ladder. Lane 2 and 5: wild type controls. Lane 6: negative control pET26b(+). Lane 3-4, 7-8: lysates from GtHNL and GtHNL-Tr variants.

3.1.1.2 Cyanohydrin synthesis

The *Gt*HNL (A40HV42T, A40RV42T) and *Gt*HNL-Tr (F19V, W120R) variants originating from site directed mutagenesis were examined in the synthesis reaction to the respective cyanohydrin using benzaldehyde and HCN as substrates. The reaction was carried out in a biphasic system for two hours and the resulting products were analysed with GC. The experiments for the synthesis of mandelonitrile showed improved conversion and enantioselectivity for all variants compared to the wild type *Gt*HNL (Table 20). *Gt*HNL-Tr with an excellent enantioselectivity (ee 100 %) and very good conversion (98 %) was compared to the newly generated variants. The variant A40HV42T synthesised enantiopure (*R*)-mandelonitrile (ee 100 %), thus showed as good enantioselectivity as *Gt*HNL-Tr, while A40RV42T showed an enantiomeric excess of 95.7 %. The enantioselectivity of the *Gt*HNL-Tr variants, F19V (ee 98.1 %) and W120R (ee % 99.1) was also slightly decreased. The best conversion achieved the variant *Gt*HNL_A40HV42T. Compared to *Gt*HNL-Tr the double mutant showed only a very little difference achieving a conversion of 97.7 %. By contrast the double mutant A40RV42T and W120R showed markedly lower conversions compared to *Gt*HNL-Tr.

Table 20: Conversion and enantioselectivity of GtHNL and GtHNL-Tr variants in the synthesis reaction. Reaction was carried out in a bisphasic system for 2h. The resulting mandelonitrile was analysed with GC.

enzyme	conversion [%]	enantiomeric excess (ee) [%]
*Gt*HNL	38.0	82.0
*Gt*HNL-Tr	98.0	> 99.9
*Gt*HNL_A40HV42T	97.7	> 99.9
*Gt*HNL_A40RV42T	50.5	95.7
*Gt*HNL-Tr_F19V	61.3	98.1
*Gt*HNL-Tr_W120R	94.5	99.6

3.1.1.3 HNL-Activity Assay

Variants of *Gt*HNL (A40HV42T and A40RV42T) and *Gt*HNL-Tr (F19V and W120R) were examined in the cleavage reaction of (*R*)-mandelonitrile and/or (*R*)-2-chloromandelonitrile and (*S*)-2-chloromandelonitrile. The HNL activity assay was performed in a plate reader using cleared lysate in triplicate and a substrate concentration of 18 mM. Two buffer systems were tested either citrate phosphate buffer or oxalate both at pH 5.5. The reaction was tracked by following the released benzaldehyde at 280 nm or in the case of 2-chlorobenzaldheyde at 300 nm for 10 min. The enzyme activity of the different variants was calculated with the following formula:

$$U = \frac{\frac{\Delta A}{t}}{\varepsilon * d} * \frac{V_t}{V_s} * f$$

U = *enzyme activity in µmol/min*
ΔA/t = *change of absorption over time ΔA/min*
ε = *excitation coefficient in l * mmol^{-1} * cm^{-1} (for benzaldehyde 1.376, for 2-chlorobenzaldehyde 1.521)*
d = *thickness in cm*
V_t = *total volume in µL*
V_s = *sample volume in µL*
f = *dilution factor*

The slope ΔA/t was calculated at the beginning of the reaction in the linear range (~ 2 min). The specific activity U/mg was calculated by including the total protein concentration of the lysate and the specific activity U/mg protein was calculated by assuming that ~ 50 % of total protein is expressed protein (Table 21). Calculation parameters are attached in the Appendix.

The strongest specific activity on (*R*)-mandelonitrile (*R*-MN) in oxalate buffer exhibited the double mutant *Gt*HNL_A40HV42T with 11.15 ± 3.13 U/mg, while the other examined enzymes showed < 0.24 U/mg lysate. Interestingly *Gt*HNL-Tr showed weak activity in oxalate buffer, whereas the specific activity in citrate phosphate buffer was increased to 22.75 ± 2.61 U/mg lysate. The double mutant *Gt*HNL_A40HV42T showed a similar specific activity at 20.19 ± 6.11 U/mg lysate. The variant *Gt*HNL-Tr F19V was also more active (4.69 U/mg lysate) in citrate phosphate buffer, while the other variants showed very low activity independently of the applied buffer system.

Table 21: Photometrical analysis of the specific activity for GtHNL and GtHNL-Tr variants. Two buffers were tested with (R)-mandelonitrile ((R)-MN) using a final concentration of 18 mM. The increase of benzaldehyde was followed at 280 nm in a plate reader.

buffer	oxalate buffer pH 5.5		citrate phosphate buffer pH 5.5	
substrate	(*R*)-MN		(*R*)-MN	
specific activity	U/mg lysate	U/mg protein	U/mg lysate	U/mg protein
*Gt*HNL-Tr	0.24 ± 0.04	0.48	22.75 ± 2.61	45.50
*Gt*HNL_A40HV42T	11.15 ± 3.13	22.30	20.19 ± 6.11	40.38
*Gt*HNL_A40RV42T	0.20 ± 0.01	0.40	0.17 ± 0.00	0.34
*Gt*HNL-Tr_F19V	0.09 ± 0.02	0.18	4.69 ± 0.08	9.38
*Gt*HNL-Tr_W120R	0.02 ± 0.00	0.04	0.05 ± 0.01	0.10

The specific activity of the variants from *Gt*HNL-Tr (F19V and W120) on (*R*)-2-chloromandelonitrile ((*R*)-2-ClMN) and (*S*)-2-chloromandelonitrile ((*S*)-2-ClMN) was determined (Table 22). The variant *Gt*HNL-Tr_W120R showed in both applied buffers oxalate and citrate phosphate a weak activity on both

enantiomers. *Gt*HNL-Tr was moderate active on (*R*)-2-ClMN in oxalate buffer (0.31 U/mg lysate) and achieved a specific activity of 5.26 U/mg lysate in citrate phosphate buffer. However, there was no activity detected against the (*S*) enantiomer. The best activity towards (*S*)-2-ClMN showed the variant *Gt*HNL-Tr_F19V, which was not active in oxalate buffer, but showed a specific activity of 0.27 U/mg lysate in citrate phosphate buffer.

The activity on the (*R*) enantiomer was increased to 1.51 U/mg lysate in oxalate and 9.38 U/mg lysate in citrate phosphate buffer.

Table 22: Specific activity of GtHNL-Tr single mutants in two different buffers tested on 18 mM (R)-2-chloromandelonitrile and (S)-2-chloromandelonitrile. Release of 2-chlorobenzaldehyde was detected at 300 nm in a plate reader.

buffer	oxalate buffer pH 5.5				citrate phosphate buffer pH 5.5			
substrate	(*R*)-2-ClMN		(*S*)-2-ClMN		(*R*)-2-ClMN		(*S*)-2-ClMN	
specific activity	U/mg lysate	U/mg protein	U/mg lysate	U/mg protein	U/mg lysate	U/mg protein	U/mg lysate	U/mg protein
*Gt*HNL-Tr	0.31 ± 0.06	0.62	-	-	5.26 ± 0.04	10.52	-	-
*Gt*HNL-Tr_F19V	1.51 ± 0.29	3.02	-	-	9.38 ± 0.29	18.76	0.27 ± 0.13	0.54
*Gt*HNL-Tr_W120R	0.05 ± 0.00	0.10	0.05 ± 0.01	0.10	0.20 ± 0.00	0.40	-	-

3.2 Site saturation mutagenesis

3.2.1.1 Evaluation of site saturation libraries

To evaluate the site saturation libraries, four randomly picked colonies were sequenced. The sequencing results showed, that all targeted positions were mutated (Table 23).

Table 23: Evaluation of the site saturation libraries. The mutated nucleotides are highlighted in grey.

	nucleotide exchange	amino acid exchange	nucleotide exchange	amino acid exchange
*Gt*HNL	TTT	F19X	TGG	W120X
1	GGT	F19G	GCG	W120A
2	GCG	F19A	ATG	W120M
3	GCT	F19A	GCG	W120A
4	GGG	F19G	TCT	W120S

*Gt*HNL-Tr				
1	CAT	F19H	TGT	W120C
2	ACT	F19T	TGT	W120C
3	CGT	F19R	ATG	W120M
4	GAG	F19E	GTG	W120V

3.2.1.2 Colony based HNL-Assay

The active site amino acids phenylalanine (F19) and tryptophan (W120) were investigated by site saturation mutagenesis studies for *Gt*HNL and *Gt*HNL-Tr. The site saturation libraries (*Gt*HNL_F19X, *Gt*HNL_W120X, *Gt*HNL-Tr_F19X and *Gt*HNL-Tr_W120X) with ~ 170 single colonies (in the case of *Gt*HNL-Tr_W120X with ~ 250 colonies) were screened with the colony based screening assay for improved activity towards both (*S*)-2-chloromandelonitrile ((*S*)-2-ClMN) and (*R*)-2-chloromandelonitrile ((*R*)-2-ClMN). The assay uses a HCN sensitive filter paper. Activity of the variants could be determined by the release of HCN and the formation of blue colour on the filter paper. The elapsed time until the signal appeared was recorded. Colonies, producing a faster signal than the wild type (*Gt*HNL or *Gt*HNL-Tr respectively) were collected for re-screening.

The site saturation library screening on (*R*)-2-ClMN of *Gt*HNL_F19X revealed in the first round 37 variants, which produced a faster signal than the wild type. In the re-screening the signal of *Gt*HNL arose after 3:20 min, whereas the best signal was achieved with the colonies F31, F61 and B81 (2:00 min) as well as C62 (2:50 min) (Figure 10). All four variants were subjected to sequencing. The primary screening of the same library against the other enantiomer (*S*)-2-ClMN revealed 33 variants with a better activity compared to the wild type. In the second screening the fastest signal was produced by the colonies B101, G112 and B122 (1:20 min), which were then also sequenced. The colony of the wild type gave a signal after 3:50 min (Figure 11).

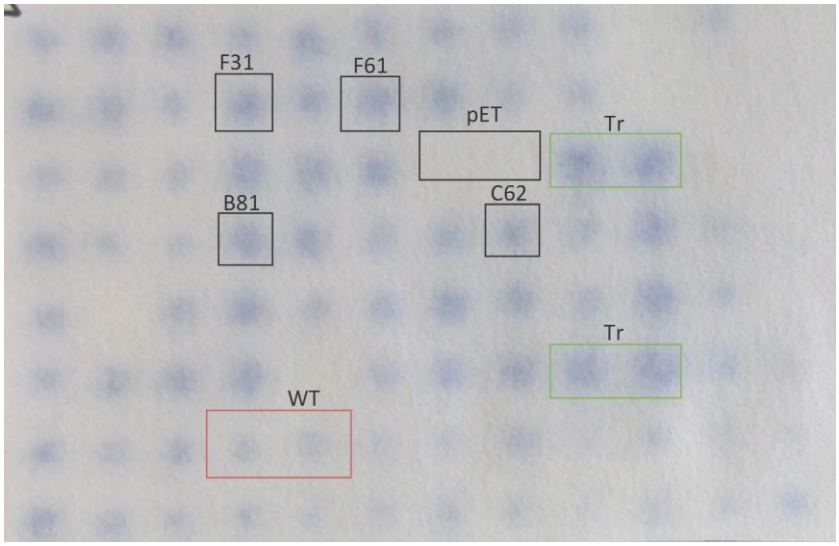


Figure 10: Re-screen filter of GtHNL_F19X library, screened against 10 mM (R)-2-chloromandelonitrile. Photo was taken 10 min after reaction start. pET indicates the negative control pET26b(+), Tr indicates GtHNL-Tr. The signal of GtHNL (WT) arose after 3:20 min. Variants F31, F61 and B81 signal arose after 2:00 min and C62 after 2:50 min.

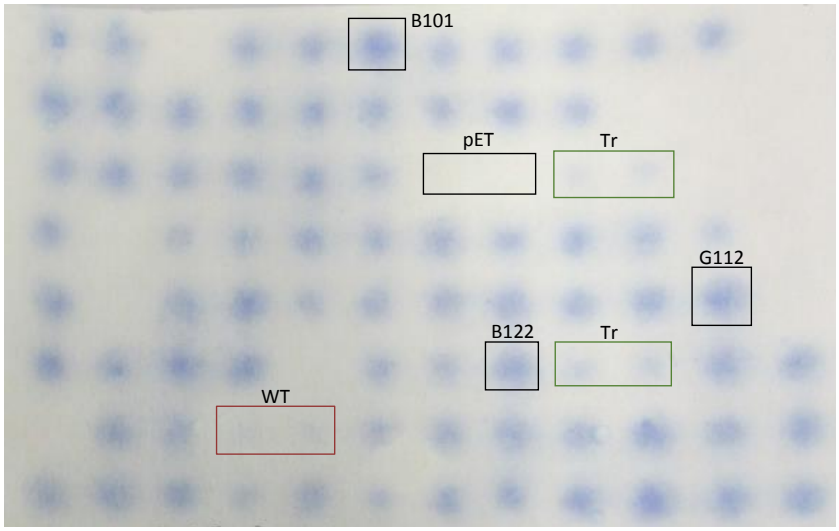


Figure 11: Re-screen filter of the GtHNL_F19X site saturation library screened towards (S)-2-chloromandelonitrile (20 mM). The photo was taken 10 min after reaction start. WT indicates GtHNL, pET the vector pET26b(+), Tr the triple variant GtHNL-Tr. Blue spots of B101, B122 and G112 appeared after 1:20 min. WT signal arose after 3:50 min.

In the first round of screening the *Gt*HNL_W120X library, in total 39 colonies were identified to be more active on (*R*)-2-ClMN than the wild type *Gt*HNL. After re-screening three colonies were chosen to be sequenced. These colonies (B42, F112, C51) produced the signal in the re-screening after 2:10 min, that was faster than the other examined colonies. The wild type *Gt*HNL needed 3:30 min to produce a visible signal (Figure 12).

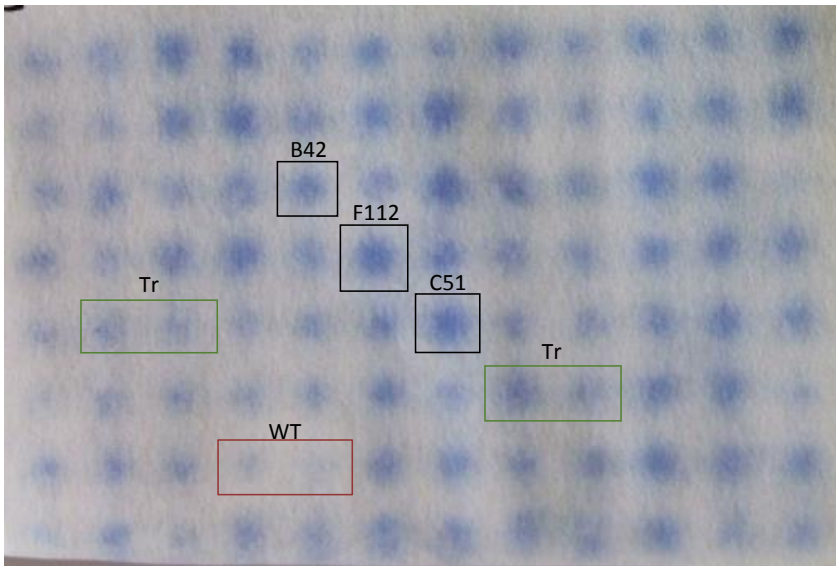


Figure 12: Re-screen filter of GtHNL_W120 10 min after reaction start. Colonies were screened with 10 mM (R)-2-chloromandelonitrile. WT indicates GtHNL, Tr indicates GtHNL-Tr. Variant C51, B42 and F112 appeared after 2:10 min. Wild type signal arose after 3:30 min.

A distinct colour signal was achieved by screening the *Gt*HNL_W120X library towards (*S*)-2-ClMN. In the first screening for example the colonies C61 and G51 produced a blue signal after 1:40 min, while the signal of the wild type *Gt*HNL arose after 3 min (Figure 13). In total 15 variants, which gave a faster signal than the wild type, were re-screened. After the second screening four variants were chosen to be sequenced.

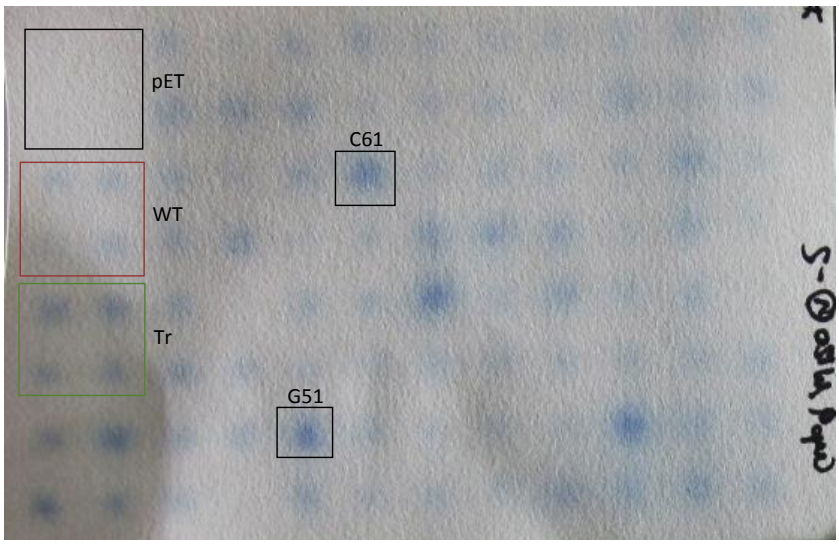


Figure 13: Screening filter of GtHNL_W120X library after 10 min of reaction. Colonies were screened with 20 mM (S)-2-chloromandelonitrile. WT indicates GtHNL, pET indicates the negative control pET26b(+), Tr is GtHNL-Tr. Variant C61 and G51 signal arose after 1:42 min, wild type signal after 3 min.

The *Gt*HNL-Tr variant already shows a very good activity on (*R*)-2-ClMN. By investigating the site saturation library of amino acid position 19 in the first round of screening on (*R*)-2-ClMN, 45 colonies were found which gave a similar fast signal than *Gt*HNL-Tr. These variants were re-screened (Figure 14). Three variants (E41, G21 and C112), which produced the signal in the same time as *Gt*HNL-Tr (2:30 min) were chosen to be sequenced.

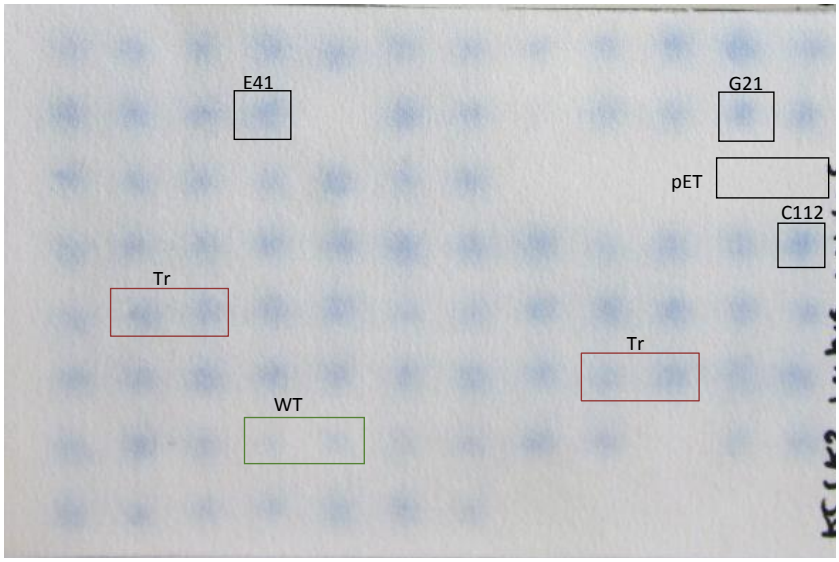


Figure 14: Example of a re-screen filter of GtHNL-Tr_F19X screened towards (R)-2-chloromandelonitrile (10mM). Photo was taken 10 min after reaction start. Tr indicates GtHNL-Tr, pET is the vector pET26b(+) and WT is GtHNL. The signal of GtHNL-Tr arose after 2:30 min. The same time was measured for the variants (E41, G21 and C112).

In the screening against (*S*)-2-ClMN of the site saturation library *Gt*HNL-Tr_F19X 32 colonies were found, which gave a faster signal than *Gt*HNL-Tr. Three of these variants (B81V, D111, F91), which were later also sequenced, produced the blue signal in the first screening round after 1:22 min, whereas the signal of *Gt*HNL-Tr appeared after 2:50 min (Figure 15).

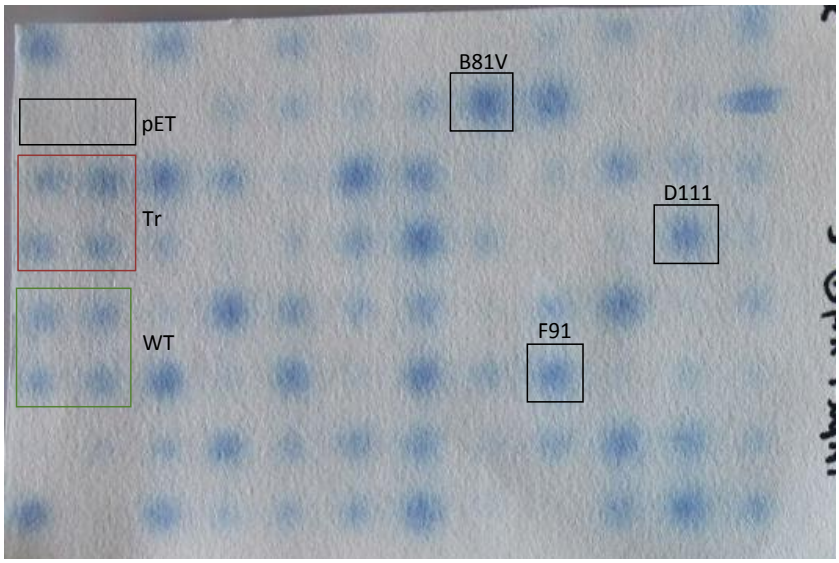


Figure 15: Screening filter of GtHNL-Tr_F19X library after 10 min of reaction. Colonies were screened with 20 mM (S)-2-chloromandelonitrile (20 mM). Tr indicates GtHNL-Tr, pET indicates the negative control pET26b(+) and WT is GtHNL. B81V appeared after 1:22 min at the same time as D111 and F91. Signal from GtHNL-Tr arose after 2:50 min.

The colony based screening of the site saturation library *Gt*HNL-Tr_W120X against (*R*)-2-ClMN did not give any colony that produced the signal faster than *Gt*HNL-Tr. However, 40 variants were used in the re-screening. Three variants (B10a, B11a and D6a) of the rescreening are depicted in Figure 16. The signal of these colonies appeared after 2:50 min, whereas the signal of *Gt*HNL-Tr arose earlier at 2:20 min. In total seven colonies, which gave a signal faster than the other colonies were selected for sequencing.

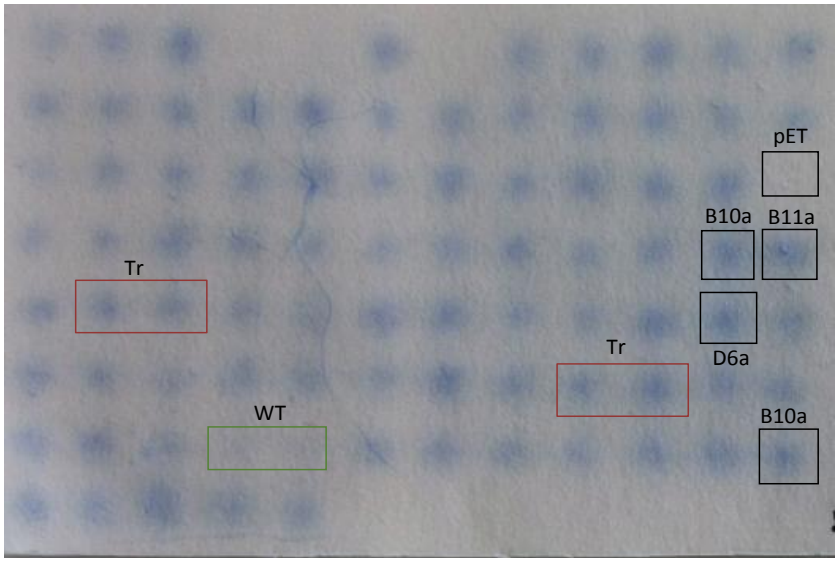


Figure 16: Example of a re-screen filter of GtHNL-Tr_W120X site saturation library. Photo was taken 10 min after reaction start with (R)-2-chloromandelonitrile (10 mM). Tr indicates GtHNL-Tr and WT indicates GtHNL. Blue spots of Tr appeared after 2:20 min. Variants B10a, B11a and D6a arose after 2:50 min.

In the site saturation library screening of *Gt*HNL-Tr_W120X against (*S*)-2-ClMN in total 39 variants were found, which were chosen for re-screening. The signal of the fastest variant (G52) appeared in the re-screening after 2:00 min, while the signal of *Gt*HNL-Tr appeared after 3:00 min (Figure 17). After re-screening three variants were sequenced.

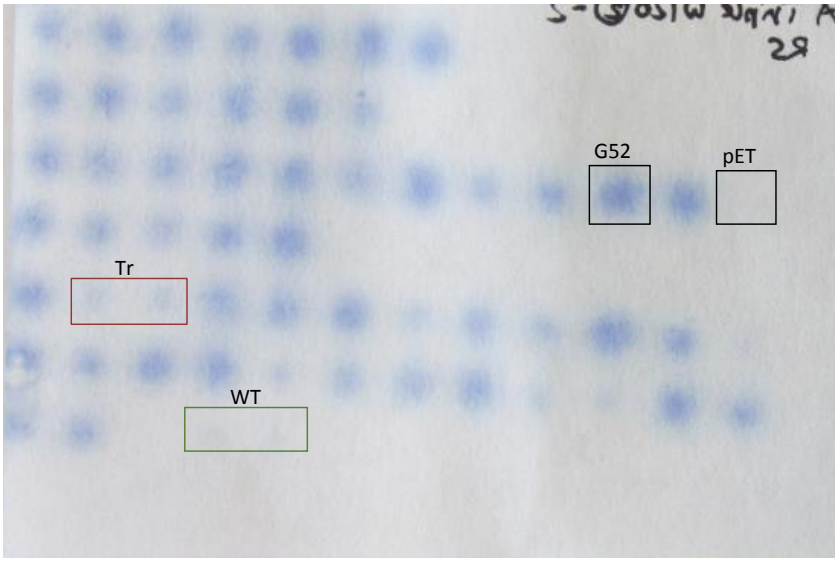


Figure 17: Re-screen filter of GtHNL-Tr_W120X 10 min after reaction start. Colonies were screened with 20 mM (S)-2-chloromandelonitrile. Tr indicates GtHNL-Tr, WT is GtHNL and pET the vector pET26b(+). Blue spot of variant G52 appeared after 2:00 min. Tr arose after 3:00 min.

The sequencing results (Table 24) revealed, that some mutations were identified in several variants. For example, in the library of *Gt*HNL_W120X the variant W120L was found three times with two different triplets (TTG and CTT) when screened against (*R*)-2-ClMN. The same amino acid exchange was also found in the library of *Gt*HNL-Tr when screened against both substrates. Another example for a multiple hit was *Gt*HNL-Tr_W120A, which was also found three times. Interestingly three variants with valine at position 19 appeared in both libraries *Gt*HNL and *Gt*HNL-Tr, when screened towards (*S*)-2-ClMN. The same amino acid exchange was previously examined with site directed mutagenesis (Chapter 3.1). However, the specific activity on (*R*)-2-ClMN and (*S*)-2-ClMN for 13 of these hits was further determined.

Table 24: Results from sequenced hits of the site saturation libraries.

Site saturation library	(*R*)-2-chloromandelonitrile			(*S*)-2-chloromandelonitrile		
	sample name	mutation	aa exchange	sample name	mutation	aa exchange
*Gt*HNL_F19X	C62	ACG	F19T	B101	GTT	F19V
	F31	ATT	F19I	G112	TGT	F19C
	F61	CTT	F19L	B122	ATT	F19I
	B81	CGT	F19R			
*Gt*HNL_W120X	C51	TTG	W120L	G51	GAG	W120E
	B42	CTT	W120L	C61	GAG	W120E
	F112	CTT	W120L	H22	CAT	W120H
	D71	CGT	W120R	H82	CAT	W120H
*Gt*HNL-Tr_F19X	E41	TGT	F19C	D111	ACG	F19T
	G21	CGG	F19R	F91	ACG	F19T
	C112	GCT	F19A	B81V	GTG	F19V
				G82	GTG	F19V
				F102	CTT	F19L
*Gt*HNL-Tr_W120X	B10a	CTG	W120L	F1120	TTG	W120L
	B11a	GCT	W120A	A122	TGT	W120C
	H1a	AAG	W120K	G52	CTG	W120L
	D6a	GCT	W120A			
	C32	GCT	W120A			
	A102	CCT	W120P			
	C42	CCT	W120P			

3.2.1.3 SDS-PAGE

Selected *Gt*HNL and *Gt*HNL-Tr hits from the colony based screening of the site saturation libraries were expressed in shake flasks in the presence of $MnCl_2$. The lysates of the total protein of *E. coli* BL21-Gold(DE3) were analysed with SDS-PAGE (Figure 18). The overexpression could be determined for all

analysed variants at an expected size of ~ 15 kDa. The variants *Gt*HNL_W120E (lane 6), W120H (lane 7) and W120K (lane 13) with a newly introduced charged amino acid showed a shorter migration distance. Also the variant *Gt*HNL-Tr_W120P (lane 11) carrying proline, showed due to putatively introduced kinks in the protein structure a shorter migration distance.

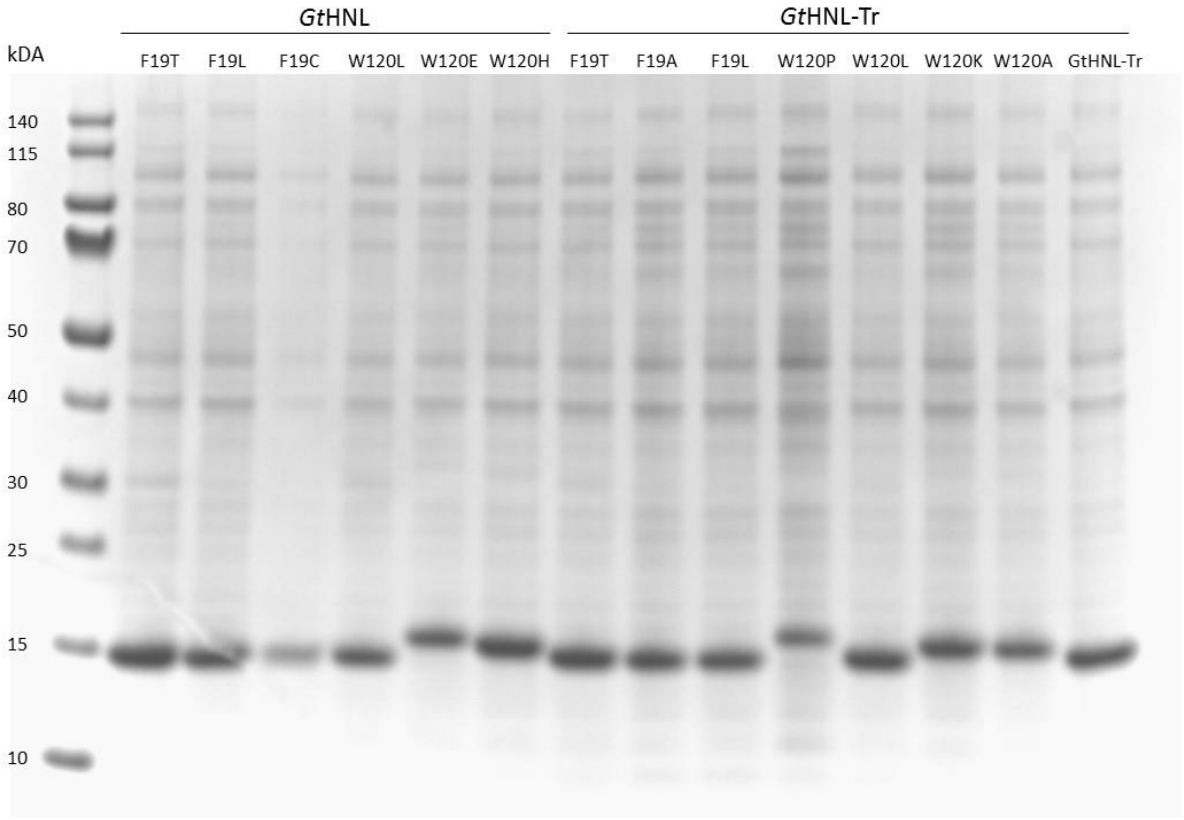


Figure 18: SDS gel of variants from site saturation library. Lane 1: Page Ruler Prestained Protein Ladder. Lane 2-7: Lysate from single mutants of GtHNL_F19 and GtHNL_W120. Lane 8-14: Lysates from GtHNL-Tr with single mutation at the position F19 and W120. Lane 15: positive control.

3.2.1.4 HNL-Activity Assay

Selected variants (Table 25) originating from the colony based screening of the site saturation libraries were expressed in shake flasks and the obtained lysates were used for the HNL activity assay. To investigate the specific activity against chloromandelonitrile the variants were tested on both the (*R*) and (*S*) enantiomer. The specific activity was calculated by including the total protein concentration of the lysate U/mg lysate and U/mg protein assuming that ~ 50 % of total protein is expressed enzyme. The wild type of *Gt*HNL showed a very weak specific activity on both (*R*)-2-chloromandelonitrile (0.04 ± 0.01 U/mg lysate) and (*S*)-2-chloromandelonitrile (0.01 ± 0.00 U/mg lysate). The best variant of the *Gt*HNL_F19X library on (*S*)-2-chloromandelonitrile was F19L (0.43 ± 0.03 U/mg lysate), but also the other tested variants

showed activities up to ≤ 0.35 U/mg lysate. In the site saturation library of GtHNL_W120X the amino acid exchange to a glutamate (W120E) showed the best result (0.24 ± 0.04 U/mg lysate). Compared to the *Gt*HNL_F19X library, none of the variants from the W120X library was found to be more active on (*S*)-2-chloromandelonitrile. The best variants from both libraries *Gt*HNL_F19X and *Gt*HNL_W120X on (*R*)-2-chloromandelonitrile were F19L and F19T. Both showed compared to the wild type a slightly increased activity of 0.10 U/mg lysate.

*Gt*HNL-Tr did not show any activity on (*S*)-2-chloromandelonitrile, while it exhibited a specific activity of 1.67 ± 0.01 U/mg lysate on the (*R*) enantiomer. The variant with the best activity on (*S*)-2-chloromandelonitrile was found in the site saturation library of *Gt*HNL-Tr_W120X. A specific activity of 0.30 ± 0.02 U/mg lysate was achieved with the variant W120L and 0.18 ± 0.00 U/mg lysate with variant W120P, while the other variants from both libraries showed ≤ 0.10 U/mg lysate. The activity on (*R*)-2-chloromandelonitrile was increased in comparison to *Gt*HNL-Tr. Four variants showed improved activity, whereby the amino acid exchange F19T resulted in the best variant (4.01 ± 0.08 U/mg lysate). The best variant from the *Gt*HNL-Tr_W120X library was W120L, which achieved a specific activity of 3.21 ± 0.12 U/mg lysate.

Interestingly, the amino acid exchanges F19T and F19L in *Gt*HNL were beneficial on the activity against (*S*)-2-chloromandelonitrile (≥ 0.35 U/mg lysate), while the same amino acid exchange in *Gt*HNL-Tr showed a very weak or no specific activity (≤ 0.05 U/mg lysate). Calculation parameters of the photometrical HNL activity assay are attached in the Appendix.

Table 25: Specific activity of site saturation mutants on (R)-2-chloromandelonitrile and (S)-2-chloromandelonitrile.

substrate		(*R*)-2-chloromandelonitrile		(*S*)-2-chloromandelonitrile	
specific activity		U/mg lysate	U/mg protein	U/mg lysate	U/mg protein
enzyme	aa exchange				
*Gt*HNL	none	0.04 ± 0.01	0.08	0.01 ± 0.00	0.02
	F19T	0.10 ± 0.01	0.20	0.35 ± 0.00	0.70
	F19L	0.10 ± 0.00	0.20	0.43 ± 0.03	0.86
	F19C	0.08 ± 0.00	0.16	0.30 ± 0.02	0.60
	W120L	0.05 ± 0.02	0.10	0.05	0.10
	W120E	0.03 ± 0.01	0.06	0.24 ± 0.04	0.48
	W120H	0.05 ± 0.02	0.10	0.03 ± 0.01	0.06
*Gt*HNL-Tr	none	1.67 ± 0.01	3.34	0	0

	F19T	4.01 ± 0.08	8.02	0.05 ± 0.01	0.10
	F19A	1.42 ± 0.10	2.84	0.04 ± 0.01	0.08
	F19L	2.03 ± 0.09	4.06	0	0
	W120P	1.64 ± 0.21	3.28	0.18 ± 0.00	0.36
	W120L	3.21 ± 0.12	6.42	0.30 ± 0.02	0.60
	W120K	0	0	0.02 ± 0.01	0.04
	W120A	1.93 ± 0.01	3.86	0.10 ± 0.00	0.20

3.3 Screening random library

3.3.1 Evaluation of the random library of *Ac*HNL

In order to determine the mutation rate in the random library of *Ac*HNL, six randomly picked clones were used in the plasmid isolation and were sequenced afterwards. Sequencing results revealed one wild type, while the remaining clones showed 2-5 mutations/mutated region (533 bp) corresponding to 0.4 - 0.9 % mutations/kbp. All mutations of the five clones at the nucleotide level as well as the amino acid exchange are listed in Table 26.

Table 26: Sequenced variants from random library for determination of mutation rate.

sample number	sum of mutations	nucleotide exchange	amino acid exchange	mutations/kbp %
121	5	A160G	T54A	0.9
		G196A	G66S	
		T263A	F132Y	
		T311C	M104T	
		A223G	K112E	
122	2	T56A	F19Y	0.4
		T436A	stop146R	
123	2	G19A	G7S	0.4
		G91A	A31T	
124	5	T15C	-	0.9
		C145T	R49C	
		C163T	H55Y	

		T291A	-	
		A330G	-	
125	5	G109A	V37I	0.9
		A295C	S99R	
		G355A	E119K	
		A357G		
		A463G	K155E	
126	0	wild type		0

To analyse the mutational bias often associated with the error prone PCR the gene mutations of the five randomly selected clones were typified. The analysis revealed, based on 19 nucleotide mutations a transition/transversion ratio of 2.8, indicating a mutational bias for transitions (Table 27).

Table 27: Analysis of mutations in nucleotide sequence inserted by error prone PCR. The bias of the enzyme is expressed as the ratio of transitions to transversions. The mutations per gene were calculated in average for the gene of AcHNL (393 bp).

Type of mutations	number
transition	
A↔G	10
T↔C	4
transversion	
A↔T	4
A↔C	1
G↔C	0
G↔T	0
transition/transversion	2.8
Ø mutations/gene	3.2

3.3.2 Screening of the *Ac*HNL random library for improved activity on nitroaldol reaction

In total a number of 5760 randomised variants of *Ac*HNL were screened for an improved nitroaldol activity. In the reaction (in duplicate) at room temperature and pH 6.0, benzaldehyde (10 mM) and nitromethane (0.6 mM) were consumed to form a nitroalcohol. After the nitroaldol reaction (1 h) the decrease of benzaldehyde was determined by measuring the fluorescence intensity (FI), which was proportional to benzaldehyde.

The fluorescence intensity was set relative to the slightly active wild type *Ac*HNL (100 %). Variants, which were better than the wild type on both plates, were used in the re-screening. In total ~ 500 variants were re-screened in quadruplicate.

Examples of the screening results for the random library of *Ac*HNL

An example of the first round of screening is shown in Figure 19. On MTP 6521 at position H10 was found a variant (34-FE5-65H10) with a relative FI of 42 %. In total 8 variants of this plate were chosen for re-screening. Thereof three variants (32-FF4-65A7, 34-FE5-65H10, 33-FB5-65H4) could be reproduced and were sequenced. Sequencing results revealed one wild type and two variants with an amino acid exchange (Table 28).

	1	2	3	4	5	6	7	8	9	10	11	12
A	95	92	97	88	89	87	85	100	104	86	88	
B	105	96	117	104	101	110	94	103	108	101	97	
C	99	106	102	109	92	111	103	100	99	106	110	
D	97	98	103	93	96	107	106	84	92	102	99	
E	102	90	111	96	92	93	120	111	115	90	101	
F	95	96	93	101	95	92	117	89	102	97	90	
G	89	108	108	104	108	113	101	113	147	131	123	
H	110	99	96	75	90	84	99	76	81	42	62	

Figure 19: Screening plate 6521 after nitroaldol reaction. Values show FI % relative to AcHNL (in red). In green: negative control pET26b(+). Intensity of blue colour increases with ≤ 100 %. Variant 32-FF4-65A7 at position A7, 34-FE5-65H10 at H10 and 33-FB5-65H4 at H4.

Another promising variant (15-CB11-37H7) was found in the first round of screening on MTP 3711 (Figure 20) at position H7 with a relative FI of 68 %. In total 13 variants of the MTP 3711 were re-screened, but only variant 15-CB11-37H7 could be reproduced.

	1	2	3	4	5	6	7	8	9	10	11	12
A	110	103	112	92	95	98	87	89	91	80	83	
B	105	91	83	88	93	91	114	105	97	107	102	
C	92	93	90	92	105	109	99	97	124	116	132	
D	119	96	95	81	105	106	107	93	133	106	107	
E	101	108	117	118	106	114	109	114	103	101	97	
F	97	98	96	114	112	118	115	115	113	105	113	
G	103	112	100	108	97	100	99	105	114	117	122	
H	118	87	89	106	98	86	68	106	122	88	87	

Figure 20: Relative FI % to wild type AcHNL (in red) of screening plate 3711. In green: negative control pET26b(+). Intensity of blue colour increases with distance from the mean value of AcHNL. Variant 15-CB11-37H7 at position H7.

Another screening plate (1011) from the first screening round is shown in Figure 21. Six variants from this plate were screened again (MTP positions A3, A8, A10, B3, H6, H10). Thereof only the variant (3-AE8-10B3) with a relative FI of 76 % at the MTP position B3 could be reproduced in the re-screening, where it showed a relative FI of 83 % (MTP position E8, re-screening plate A121) (Figure 22). On this re-screening plate other variants were also found (2-AA6-7E6, 6-AH11-17C6) at MTP position A6 and H11, which showed a relative FI of 77 % and 86 %. Both were chosen to be sequenced. Other promising variants were identified as wild types or could not be reproduced in parallel determinations e.g. variant at MTP position A12 with a relative FI of 68 %.

	1	2	3	4	5	6	7	8	9	10	11	12
A	108	98	91	116	112	117	109	74	102	74	117	
B	104	95	76	89	99	97	101	99	118	104	100	
C	97	106	95	106	97	86	88	97	90	99	109	
D	109	87	104	97	92	88	98	100	95	91	96	
E	92	89	92	101	89	99	98	101	102	110	113	
F	111	108	100	102	100	103	106	111	111	89	87	
G	101	109	100	92	99	93	105	97	114	108	103	
H	107	101	94	101	95	78	92	101	101	78	87	

Figure 21: Relative FI % after nitroaldol reaction of variants from AcHNL (plate 1011). In green: negative control pET26b(+). In red: wild type AcHNL. Intensity of blue colour increases with ≤ 100 %. Variants at position A3, A8, A10, B3, H6, and H10 were re-screened. Variant 3-AE8-10B3 at position B3.

	1	2	3	4	5	6	7	8	9	10	11	12
A	103	89	94	89	96	77	105	114	89	92	89	68
B	110	99	103	102	93	103	124	102	100	96	84	
C	107	99	99	98	105	96	103	90	118	107	96	88
D	107	100	120	96	102	105	108	88	110	120	107	
E	111	106	94	104	117	102	100	83	80	89	89	81
F	116	100	102	92	96	104	103	143	112	105	137	
G	93	101	97	97	98	102	103	85	98	103	88	80
H	112	90	91	100	99	103	102	116	98	81	86	

Figure 22: Re-screening plate A121 of promising variants from the first library screening. Values correspond to FI % relative to the wild type (in red). In green: negative control pET26b(+). Intensity of blue colour increases with distance from the mean value of AcHNL. Variant 3-AE8-10B3 at position E8, 2-AA6-7E6 at A6 and 6-AH11-17C6 at H11.

Another example for an analysed screening plate (4221) is depicted in Figure 23. The variant (18-DE2-42C11) was found at position C11 with a relative FI of 84 %. In the re-screening (plate D221) the better activity was confirmed, where the same variant at position E2 showed a relative FI of 77 % (Figure 24). In this re-screening another variant (21-DH2-43A10) at MTP position H2 with a relative FI of 73 % could also be reproduced. In total six variants were reproduced and sequenced. Three of these sequenced variants

were wild types. For example the variant 22-DD11-51B3 at MTP position D11 showed a relative FI of 81 %, but after sequencing it was identified as wild type.

	1	2	3	4	5	6	7	8	9	10	11	12
A	106	95	100	126	114	114	108	115	109	83	114	
B	129	114	140	122	121	108	126	113	129	113	133	
C	103	88	107	97	100	107	86	99	103	96	84	
D	115	94	112	110	113	94	129	111	111	120	121	
E	111	102	102	90	110	104	105	100	110	142	118	
F	137	112	118	119	106	115	98	109	115	107	121	
G	111	100	96	98	84	101	109	104	114	105	117	
H	126	88	95	115	117	99	104	98	103	96	88	

Figure 23: Example (plate 4221) for a screening on improved nitroaldolase activity. FI % is set relative to AcHNL (in red). In green: negative control pET26b(+).Intensity of blue colour increases with ≤ 100 %. A promising variant 18-DE2-42C11 is at position C11.

	1	2	3	4	5	6	7	8	9	10	11	12
A	109	104	94	114	114	107	105	102	107	92	106	111
B	113	115	120	97	125	107	119	111	123	109	108	
C	112	111	87	114	117	110	108	98	124	113	119	127
D	120	85	89	97	98	90	103	98	106	105	81	
E	109	77	96	92	98	97	99	112	118	126	109	120
F	104	90	90	91	86	89	96	97	91	110	121	
G	101	93	94	102	103	98	89	100	100	115	108	113
H	102	73	86	104	112	92	90	124	136	106	91	

Figure 24: Re-screening plate D221 of promising variants from the first library screening. Values show FI % relative to AcHNL (in red). In green: negative control pET26b(+). Intensity of blue colour increases with distance from the mean value of AcHNL. Variant 18-DE2-42C11 is at position E2 and 21-DH2-43A10 at H2.

After the re-screening of the random library all 39 reproduced variants were sequenced (Table 28). The sequencing results revealed 38.5 % wild type clones without any amino acid exchange. Thereof 28.2 % of the variants carried silent mutations and 10.3 % were found with the parental gene sequence of *Ac*HNL. However, some amino acid exchanges were found two or even three times in different variants. The amino acid position 99 was found to be mutated in three variants R93HS99C, S99N and S99GA103SE111G. Mutations that appeared two times in different variants were E119G, Q9R, M104V, W87R, E17G and E17V.

Table 28: Sequenced AcHNL variants originating from random library screening. Amino acid mutations found in more than one variant are highlighted in grey. Variants without amino acid mutations in red letters.

sampel		mutation		silent mutation
number	name	nucleotide ex	aa exchange	nucleotide ex
1	1-AA3-4E8	-	-	A48G
2	2-AA6-7E6	T199C	C67R	T240C
3	3-AE8-10B3	A26G G256A A356G	Q9R V86I E119G	T207G, T240A, T378C
4	4-AA10-11D5	-	-	-
5	5-AB11-17A6	-	-	-
6	6-AH11-17C6	G278A A295T	R93H S99C	C150T
7	7-BE3-19F10	A310G	M104V	
8	8-BG3-19H9	A383T	E128V	
9	9-BH4-21H3	-	-	G168A
10	10-BH11-19A7	G88A	A30T	A144G
11	11-CH2-31A10	-	-	T111A, C195T,A336G
12	12-CH3-31E4	A329T	Q110L	A120G, T258C, A308G
13	13-CD4-31H5	-	-	-
14	14-CA7-35E2	C24A	S8R	
15	15-CB11-37H7	T259C	W87R	C114A, T207C
16	16-CC3-31F4	T387A	$Y129^{stop}$	-
17	17-CG5-34A10	A377C	D126A	T327C
18	18-DE2-42C11	T259A A335G	W87R K112R	A99G, T141C, T351C
19	19-DE3-43F3	C28T	P10S	T255C
20	20-DG2-42H11	-	-	T75C
21	21-DH2-43A10	A310G	M104V	
22	22-DD11-51B3	-	-	T87C, T55C

23	23-DH11-51F5	-	-	T126C,A348G
24	24-EF8-59A7	-	-	G168A, T186A
25	25-EF9-60A10	-	-	A390G
26	26-EH10-61E11	G296A G307T A332G	S99N A103S E111G	T39C, T126, C315A
27	27-EB11-61F11	C95A	P32Q	T231C, C375T
28	28-EB4-55E7	A26G A50T	Q9R E17V	
29	29-EE8-59A6	A50G	E17G	
30	30-FA2-64E3			G6A, C33T, C82T, T174A
31	31-FA3-69D11	T326C	I109T	T132C, T189C
32	32-FF4-65A7	G76A	D26N	T75C
33	33-FB5-65H4	-	-	A396G
34	34-FE5-65H10	A365G	E119G	T216C
35	35-FF5-66A7	-	-	T21C, T75C, T117C, T147C
36	36-FA6-66G9	-	-	-
37	37-FA7-67C11	A295G	S99G	
38	38-FE2-64G10	G103A	A35T	A90T, A120G, G177A, T342C
39	39-FH10-16A4	C226T A368G	P76S H123R	T240A

For a statistical point of view the screening assay was evaluated by calculating the relative standard deviation (RSD) of *Ac*HNL. For the first round of screening 576 datasets of *Ac*HNL (72 DWPs with four controls screened in duplicate) resulted in a mean RSD % between 3.9 - 15.1 and in the re-screening (6 DWPs with four controls screened in quadruplicate resulting in 96 datasets) the mean RSD % was between 3.1 – 10.7.

To compare the activity of the sequenced variants to each other, they were screened again on the nitroaldol reaction. The screening was performed in octuple. The measured fluorescence intensity was set relative to the wild type *Ac*HNL. Considering the relative standard deviation, significant better activity

(~ 7 - 30%) on the nitroaldol reaction was obtained with 6 variants (Figure 25). Two of these variants carried the mutation E119G. The mutation M104V and W87R only showed significant better activity in one of the both tested variants. The other two variants with significant improved activity were C67R and A35T, both were found only once in the screened random library.

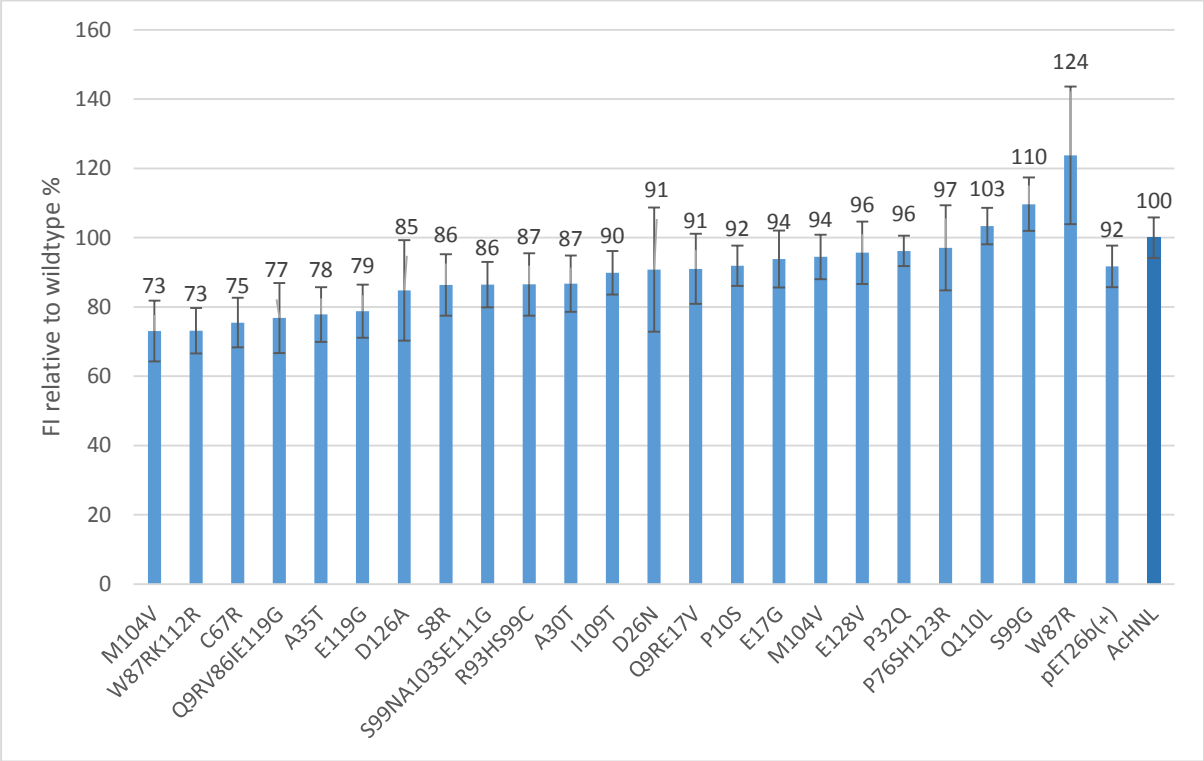


Figure 25: Relative fluorescence intensity of variants from AcHNL compared to the wild type. Nitroaldol reaction was carried out at pH 6.0 with 10 mM benzaldehyde and 0.6 M nitromethane at room temperature. Benzaldehyde decrease was determined after 1h. Error bars: RSD %.

To visualise the amino acid exchange position of these six variants a 3D-model of *Ac*HNL was used. All variants (A35T, C67R, W87RK112R, M104V, E119G and Q9RV86IE119G) found to be significantly better than the wild type *Ac*HNL occurred distant from the active site (Figure 26).

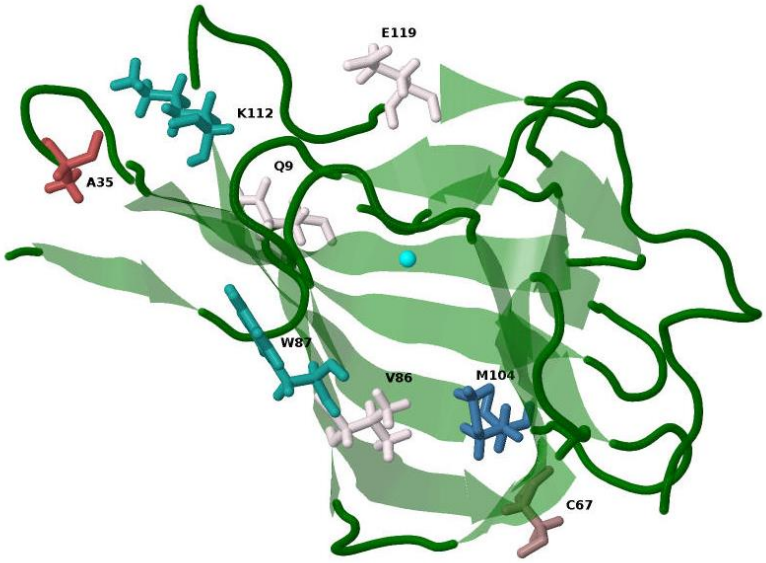


Figure 26: 3D-model of AcHNL [53] indicating amino acids, which were altered by random mutagenesis.

3.4 Screening of selected *Ac*HNL and *Gt*HNL variants

Selected variants of *Ac*HNL and *Gt*HNL, which were obtained in previous experiments by members of acib Graz and showed increased HNL activity compared to the WT (but less than the triple variants) [1], were examined for their nitroaldol activity with the fluorescence based determination of benzaldehyde decrease. The variants were cultivated and expressed in multiple wells in DWPs (at least quadruplicates). The nitroaldol reaction was performed in MTP with potassium phosphate buffer at pH 6.0 for one hour.

All examined variants carried a mutation at active site amino acids. In comparison to the wild type *Ac*HNL, the mutation A40H showed distinct lower fluorescence intensities, thus an increased benzaldehyde consumption (Figure 27). The best variant was found to be *Ac*HNL_A40H, which consumed up to ~ 80 % more benzaldehyde than the wild type *Ac*HNL.

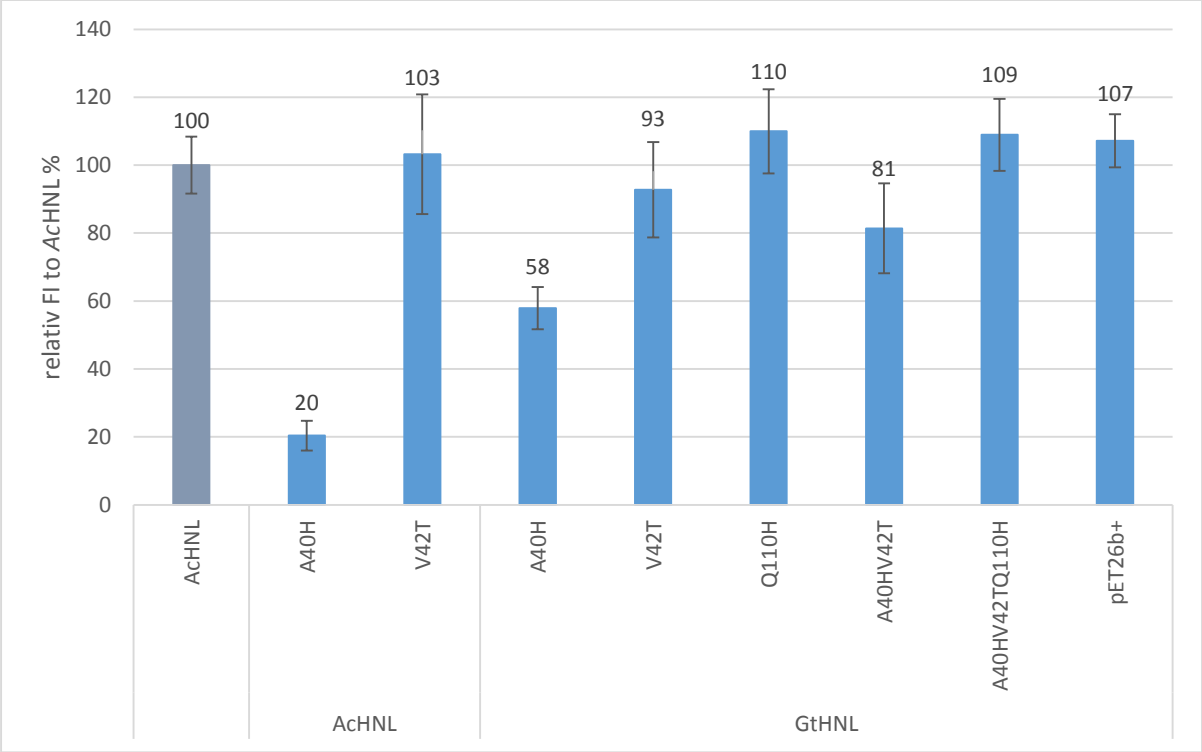


Figure 27: Comparison of selected variants of GtHNL and AcHNL on their nitroaldol activity. FI % is set relative to AcHNL. Reaction conditions: Substrates: 10 mM benzaldehyde and 0.6 M nitromethane. Reaction time: 1 h at room temperature in potassium phosphate buffer pH 6.0.

4 Discussion

The HNL-activity assay is a simple photometrical method to determine the enzyme activity of hydroxynitrile lyases in the cleavage of cyanohydrins. When discussing the accuracy of the assay there are some aspects to consider e.g. the applied method to determine the protein concentrations, sample quality and the used buffer system.

To determine the protein concentrations in this work the Bradford protocol was used. The dye Coomassie brilliant blue interacts with basic and nonpolar amino acid residues. Van der Waals forces and hydrophobic interactions are involved in the binding mechanism of the reagent. The dye forms a blue coloured complex with proteins, which can be determined photometrical at 595 nm. The Bradford assay is known to show variations in the determined protein concentrations, since the colour intensity is dependent on the protein composition respectively the amino acid residues. This imprecise method to determine protein quantities is well known for the Bradford protocol [54]. In this work a standard protein (Albumin Faktor V) was used for calibration, which may vary in the binding positions for Coomassie brilliant blue compared to the protein to be quantified. Since all samples were quantified with the Bradford reagent, the data should be consistent, although the protein concentrations may be slightly different amongst the variants, since newly amino acids with potential binding positions for the Bradford reagent were introduced.

Another point to consider at the HNL-activity assay is the sample quality. In this experiments it is known that approximately 50 % of the total protein is heterologous expressed cupin [22]. Therefore the specific activity was calculated for the amount of expressed enzyme. In general it is preferable to work with purified protein, since lysate components may interact or interfere with the assay components and subsequently decrease or increase the enzyme activity.

The next aspect, which influences the HNL-activity assay is the applied buffer system. In this work differences in enzyme activities were found when using oxalate buffer and citrate phosphate buffer, although the pH was constant at 5.5. For example *Gt*HNL_A40HV42T showed a specific activity of 11.15 ± 3.13 U/mg in oxalate buffer and 20.19 ± 6.11 U/mg in citrate phosphate buffer when testing towards (*R*)-mandelonitrile. These results showed that the nature of the buffer components is influencing the activity assay. Buffers may change the electron status in the active site of the enzyme and cause different activities. The electron status depends on the properties of the amino acid side chains, therefore the activity of variants can be higher or even lower [55]. An explanation for lower enzyme activities of the variants in oxalate buffer may also be an unidentified promiscuous activity towards the oxalate, exhibiting a potential oxalate decarboxylase activity, thus a competitive enzyme inhibition may occur. Oxalate

decarboxylase activities are well known for cupin fold manganese dependent enzymes, although this activity has not been observed in *Gt*HNL [20] [19] [56] [57].

Promiscuous reactions are very often undiscovered, for example due to their low activity or a missing screening they proceed in the background. One promiscuous reaction that was found for hydroxynitrile lyases is the nitroaldol reaction. For further investigations, a directed evolution concept was applied by generating a random library of *Ac*HNL. When performing a random library the quality of the library has to be considered.

In this work the evaluation of the random library showed a mutation rate of 0.4 – 0.9 %/kbp. Furthermore one wild type was found out of six sequenced variants. A low mutation rate is preferable, since many mutations can yield in a library with a large amount of inactive variants, e.g. caused by incorrect folded protein.

In average 3.2 mutations were calculated per gene for the random library of *Ac*HNL. With 5760 screened variants only a very small fraction of the theoretically possible variants was covered. To consider all possible sequences the library size would be at ~ 9.24×10^9 variants [58].

The transition/transversion ratio indicated the bias of the enzyme used in the error prone PCR. No bias towards transitional or transversional substitutions would yield the ratio 0.5, indicating a high quality of the random library with an optimal diversity [59]. The random library of *Ac*HNL showed a bias to transitional mutations which occurred 2.8-fold more often than transversions.

The screening assay was arranged to analyse the substrate decrease of benzaldehyde. The screening range was subjected to some limitations. On the one hand the substrate benzaldehyde undergoes auto oxidation in air. This process is accelerated by light and yields in benzoic acid [60]. On the other hand the screening range is reduced due to the natural activity of *Ac*HNL on the nitroaldol reaction. To find variants with improved activity on the nitroaldol reaction, screening results were compared to the wild type.

The variability of the assay was evaluated after the screening. A mean relative standard deviation (RSD) for *Ac*HNL (576 datasets) of the first screening was calculated to be between 3.9 - 15.1 %. An explanation of this high variability could be irregularities in growth of *E. coli* in deep well plates. Variants at the edge of the DWP might have a better oxygen supply. This effect can lead to differences in growth and the amount of expressed protein. Furthermore fluorescence is a very precise method, detecting even very small variations, which could be caused by the used material e.g. multichannel pipettes and the related

multiple pipetting steps. The high variability of the screening assay could also be an explanation why sequencing results revealed 38.5 % wild type.

The sequenced variants, which carried an amino acid exchange, were screened again. Considering the RSD, six variants were found, which were significantly better than the wild type. This variants showed improved activity on the nitroaldol reaction in the range of 7-30 % compared to the wild type *Ac*HNL.

Amino acid residues, which are interconnected by different physical forces e.g. hydrogen bonds, van der Waals forces and ionic bonds, are involved in the structure organisation and the related function of the protein. Even small changes in solubility, conformational status, substrate recognition and reaction mechanism can alter the enzyme activity. An amino acid exchange at the surface of the protein is very often responsible for the solubility of the protein. Proteins, which are soluble in aqueous environment, carry polar residues at the surface, whereas the hydrophobic residues are packed into the interior of the protein. Hydrophobic and aliphatic side chains are very flexible and non-reactive, thus play a greater role in structure and substrate recognition [61]. The randomly introduced amino acids in *Ac*HNL brought up some mutations, which were found more than once in the screened library e.g. E119G, M104V and W87R. These mutations changed one of the mentioned protein characteristics and increased the enzymes activity. To understand the mechanism in more detail a 3D structure analysis of *Ac*HNL could provide more insights. However, a comparison of the found amino acid exchanges to the 3D-model of *Ac*HNL, which is based on the structure of *Gt*HNL and *Gt*HNL-Tr, showed that the mutations occurred distant from the active site. Therefore it is more likely that these mutations changed solubility of the enzyme or a conformational change increased the enzyme activity.

The amino acid exchange A40HV42T in *Gt*HNL provided enantioselectivity of the enzyme, hindering the cleavage and formation of the (*S*) enantiomer. Alanine the smallest proteinogen amino acid at position 40 was exchanged by the larger histidine, which carries an imidazole ring with a slightly basic nature. And valine a small amino acid with aliphatic properties was exchanged by a polar threonine. The effect of this double amino acid exchange was a stereospecific cleavage of (*R*)-mandelonitrile. But also the reversed reaction showed a higher specific activity compared to the wild type, especially in the citrate phosphate buffer. The most promising mutation found in this work was A40H, which is a crucial amino acid exchange of the examined *Gt*HNL and *Ac*HNL, since catalysis was improved on both the hydroxynitrile reaction and the nitroaldol reaction. This enhancement in enzyme activity may be explained by an altered electron status and a narrowed catalytic cavity by the larger histidine with the respective basic imidazole ring.

On the other hand an enlargement of the catalytic cavity can show a contrary result. For example the exchange of F19V from phenylalanine to the smaller amino acid at the entrance to the active site, was disadvantageous, since enantioselectivity was lost. The same effect was shown for W120R.

Results of the site saturation screening showed leucine at position 19 and 120 to be the preferred amino acid, enhancing the enzymes activity towards chlorinated mandelonitrile. Again the size of the catalytic cavity was enlarged, thus providing more space for the substrate, but the enantioselectivity for example of GtHNL-Tr was lost. Interestingly, the site saturation library of wild type *Gt*HNL brought up the amino acid exchange W120E to be preferable in catalyzing (*S*)-2-chloromandelonitrile. The properties of the active site of *Gt*HNL were changed by the glutamate with its polar reactive side chain and in consequence enhanced only the activity towards the (*S*) enantiomer.

5 Outlook

An interesting approach in future would be to switch the (*R*) enantioselectivity of *Gt*HNL, since cyanohydrins with (*S*) oriented groups also play a great role as intermediates in the synthesis of fine chemicals. Furthermore, examinations concerning the substrate scope could be continued.

The randomly generated variants of *Ac*HNL, which were significantly better than the wild type, when screening on the nitroaldol reaction, can further be tested with a quantitative analysing method e.g. HPLC. As mentioned before directed evolution experiments are always iterative approaches. Improved variants found in the first round of evolution could be a starting pool for a second randomization step with subsequent screening. In consequence of a defined substrate scope of *Ac*HNL, further examinations and evolutionary based approaches could be applied to broaden the enzymes application.

Furthermore a resolved structure of *Ac*HNL would bring further insights into the relationship between structure and function, subsequently directing the experiments to a more rational approach. An improvement of the enzymes properties e.g. substrate scope and catalysing activity could fit in the enzyme at an industrial scale.

6 References

[1] Wiedner R, Kothbauer B, Schwab H, Steiner K, et.al, „Improving the Properties of Bacterial R - Selective Hydroxynitrile Lyases for Industrial Applications," vol. 7, no. 2, pp. 325-332, 2015.

[2] Schoemaker H, „Dispelling the Myths - Biocatalysis in Industrial Synthesis," *Science,* vol. 299, no. 5613, pp. 1694-1697, 2003.

[3] Choi J, Han S, Kim H, „Industrial applications of enzyme biocatalysis: Current status and future aspects," *Biotechnology Advances,* p. in press, 2015.

[4] Breuer M, Ditrich K, Habicher T, Zelinski T, et.al, „Industrial methods for the production of optically active intermediates," *Angewandte Chemie - International Edition,* vol. 43, no. 7, pp. 788-824, 2004.

[5] García-Urdiales E, Alfonso I, Gotor V, „Enantioselective enzymatic desymmetrizations in organic synthesis," *Chemical Reviews,* vol. 105, no. 1, pp. 313-354, 2005.

[6] Griengl H, Schwab H, Fechter M, „The synthesis of chiral cyanohydrins by oxynitrilases," *Trends in Biotechnology,* vol. 18, no. 6, pp. 252-256, 2000.

[7] Dadashipour M, Asano Y, „Hydroxynitrile lyases: Insights into biochemistry, discovery, and engineering," *ACS Catalysis,* vol. 1, no. 9, pp. 1121-1149, 2011.

[8] Gruber K, Gartler G, Krammer B, Schwab H, Kratky C, „Reaction mechanism of hydroxynitrile lyases of the α/ β-hydrolase superfamily: The three-dimensional structure of the transient enzyme-substrate complex certifies the crucial role of Lys236," *Journal of Biological Chemistry,* vol. 279, no. 19, pp. 20501-20510, 2004.

[9] Gruber K, Kratky C, „Biopolymers for Biocatalysis: Structure and Catalytic Mechanism of Hydroxynitrile Lyases," *Journal of Polymer Science Part A: Polymer Chemistry,* vol. 42, pp. 479-486, 2004.

[10] Hussain Z, Wiedner R, Steiner K, Schwab H, et.al, "Characterization of two bacterial hydroxynitrile lyases with high similarity to cupin superfamily proteins," *Applied and Environmental Microbiology,* vol. 78, pp. 2053-2055, 2012.

[11] Breuer M, Hauer B, „Carbon-carbon coupling in biotransformation," *Current Opinion in Biotechnology,* vol. 14, no. 6, pp. 570-576, 2013.

[12] Purkarthofer T, Skranc W, Schuster C, Griengl H, „Potential and capabilities of hydroxynitrile lyases as biocatalysts in the chemical industry," *Applied Microbiology and Biotechnology,* vol. 76, no. 2, pp. 309-320, 2007.

[13] Andexer J, Langermann J, Kragl U, Phol M, „How to overcome limitations in biotechnological processes - examples from hydroxynitrile lyase applications," *Trends in Biotechnology,* vol. 27, pp. 599-607, 2009.

[14] Wang W, Liu X, Lin L, Feng X, „Recent progress in the chemically catalyzed enantioselective synthesis of cyanohydrins," *European Journal of Organic Chemistry,* vol. 25, pp. 4751-4769, 2010.

[15] Johnson D, Zabelinskaja-Mackova A, Griengl H, „Oxynitrilases for asymmetric C – C bond formation," *Curr Opin Chem Biol,* vol. 1, pp. 103-109, 2000.

[16] Glieder A, Weis R, Skranc W, Gruber K, et.al, „Comprehensive Step-by-Step Engineering of an (R)-Hydroxynitrile Lyase for Large-Scale Asymmetric Synthesis," *Angewandte Chemie - International Edition,* vol. 42, no. 39, pp. 4815-4818, 2003.

[17] Agarwal G, Rajavel M, Gopal B, Srinivasan N, „Structure-based phylogeny as a diagnostic for functional characterization of proteins with a cupin fold," *PLoS ONE,* vol. 4, no. 5, 2009.

[18] Dunwell J, Culham A, Carter C, Goodenough P, et.al, „Evolution of functional diversity in the cupin superfamily," *Trends in Biochemical Sciences,* vol. 26, no. 12, pp. 740-746, 2001.

[19] Dunwell J, Purvis A, Khuri S, „Cupins: The most functionally diverse protein superfamily?," *Phytochemistry,* vol. 65, pp. 7-17, 2004.

[20] Khuri S, Bakker F, Dunwell J, „Phylogeny, function, and evolution of the cupins, a structurally conserved, functionally diverse superfamily of proteins.," *Molecular biology and evolution,* vol. 18, no. 4, pp. 593-605, 2001.

[21] Yskowski A, Steiner K, Hajnal, Gruber K et.al, „Crystallization of a novel metal-containing cupin from Acidobacterium sp. and preliminary diffraction data analysis," *Acta Crystallographica Section F: Structural Biology and Crystallization Communications,* vol. 68, pp. 451-454, 2012.

[22] Hajnal I, Lyskowski A, Schwab H, Steiner K, et.al, „Biochemical and structural characterization of a novel bacterial manganese-dependent hydroxynitrile lyase.," *The FEBS journal,* vol. 280, no. 22, pp. 5815-28, 2013.

[23] Wiedner R, Gruber-Khadjawi M, Schwab H, Steiner K, „Discovery of a novel (R)-selective bacterial hydroxynitrile lyase from Acidobacterium capsulatum," *Computational and Structural Biotechnology Journal,* vol. 10, no. 16, pp. 58-62, 2014.

[24] Purkarthofer T, Gruber K, Gruber-Khadjawi M, Griengl H et. al, „A biocatalytic Henry reaction-the hydroxynitrile lyase from Hevea brasiliensis also catalyzes nitroaldol reactions," *Angewandte Chemie (International ed. in English),* vol. 45, no. 21, pp. 3454-6, 2006.

[25] Humble M, Berglund P, „Biocatalytic promiscuity," *European Journal of Organic Chemistry,* vol. 19, pp. 3391-3401, 2011.

[26] Bornscheuer U, Kazlauskas R, „Catalytic promiscuity in biocatalysis: Using old enzymes to form new bonds and follow new pathways," *Angewandte Chemie - International Edition,* vol. 43, no. 45, pp. 6032-6040, 2004.

[27] Miao Y, Rahimi M, Geertsema E, Poelarends G, „Recent developments in enzyme promiscuity for carbon–carbon bond-forming reactions," *Current Opinion in Chemical Biology,* vol. 25, pp. 115-123, 2015.

[28] Gruber-Khadjawi M, Purkarthofer T, Skranc W, Griengl H, „Hydroxynitrile Lyase-Catalyzed Enzymatic Nitroaldol (Henry) Reaction," *Advanced Synthesis & Catalysis,* vol. 349, no. 8-9, pp. 1445-1450, 2007.

[29] Yuryev R, Briechle S, Gruber-Khadjaw M, Liese A, et.al, „Asymmetric Retro-Henry Reaction Catalyzed by Hydroxynitrile Lyase from Hevea brasiliensis," *ChemCatChem,* vol. 2, no. 8, pp. 981-986, 2010.

[30] Fuhshuku K, Asano Y, „Synthesis of (R)-β-nitro alcohols catalyzed by R-selective hydroxynitrile lyase from Arabidopsis thaliana in the aqueous-organic biphasic system.," *Journal of biotechnology,* vol. 153, no. (3-4), pp. 153-9, 2011.

[31] Tang R, Guan Z, He Y, Zhu W, „Enzyme-catalyzed Henry (nitroaldol) reaction," *Journal of Molecular Catalysis B: Enzymatic,* vol. 63, no. 1-2, pp. 62-67, 2010.

[32] Wang J, Li X, Xie H, Liu B, Lin X, „Hydrolase-catalyzed fast Henry reaction of nitroalkanes and aldehydes in organic media," *Journal of Biotechnology,* vol. 145, no. 3, pp. 240-243, 2010.

[33] Milner S, Moody T, Maguire A, „Biocatalytic approaches to the Henry (nitroaldol) reaction," *European Journal of Organic Chemistry,* vol. 2012, no. 16, pp. 3059-3067, 2012.

[34] Strohmeier G, Pichler H, May O, Gruber-Khadjawi M, „Application of designed enzymes in organic synthesis," *Chemical Reviews,* vol. 111, no. 7, pp. 4141-4164, 2011.

[35] Bornscheuer U, Pohl M, „Improved biocatalysys by directed evolution and rational protein design," *Curr Opin Chem Biol,* vol. 5, no. 2, pp. 137-43, 2001.

[36] Behrens G, Hummel A, Padhi S, SchätzleS, Bornscheuer U, „Discovery and protein engineering of biocatalysts for organic synthesis," *Advanced Synthesis and Catalysis,* vol. 353, no. 13, pp. 2191-2215, 2011.

[37] Ryan E. Cobb, Ran Chao H, „Directed Evolution: Past, Present, and Future," *AlChE,* vol. 59, no. 5, pp. 1432-1440, 2013.

[38] Bornscheuer U, Huisman G, Kazlauskas R, Robins K, et.al, „Engineering the third wave of biocatalysis," *Nature,* vol. 485, no. 7397, pp. 185-194, 2012.

[39] Morley K, Kazlauskas R, „Improving enzyme properties: When are closer mutations better?," *Trends in Biotechnology,* vol. 23, no. 5, pp. 231-237, 2005.

[40] Luetz S, Giber L, Lalonde J, „Engineered enzymes for chemical production," *Biotechnology and Bioengineering,* vol. 101, no. 4, pp. 647-653, 2008.

[41] Fromant M, Blanquet S, Plateau P, „Direct random mutagenesis of gene-sized DNA fragments using polymerase chain reaction," *Analytical biochemistry,* vol. 224, no. 1, pp. 347-353, 1995.

[42] Denard C, Ren H, Zhao H, „Improving and repurposing biocatalysts via directed evolution," *Current Opinion in Chemical Biology,* vol. 25, no. 55-64, pp. 55-64, 2015.

[43] Reetz M, „Laboratory evolution of stereoselective enzymes: A prolific source of catalysts for asymmetric reactions," *Angewandte Chemie - International Edition,* vol. 50, no. 1, pp. 138-174, 2011.

[44] Steiner K, Schwab H, „Recent advances in rational approaches for enzyme engineering," *Computational and Structural Biotechnology Journal,* vol. 2, no. 3, 2012.

[45] Bommarius A, Blum J, Abrahamson M, „Status of protein engineering for biocatalysts: How to design an industrially useful biocatalyst," *Current Opinion in Chemical Biology,* vol. 15, no. 2, pp. 194-200, 2011.

[46] Reetz M, Kahakeaw D, Lohmer R, „Addressing the numbers problem in directed evolution," *ChemBioChem,* vol. 9, no. 11, pp. 1797-1804, 2008.

[47] „http://www.lifetechnologies.com"

[48] „http://www.chem.agilent.com/library/usermanuals/public/230130.pdf"

[49] „http://www.merckmillipore.com/INTL/en/product/pET-26b%28%2B%29-DNA---Novagen,EMD_BIO-69862#anchor_USP“

[50] „https://www.neb.com/protocols/2012/12/07/optimizing-restriction-endonuclease-reactions“

[51] Gibson DG, et. al, „Enzymatic assembly of DNA molecules up to several hundred kilobases,“ *Nat. Methods,* vol. 6, no. 5, pp. 343-345, 2009.

[52] Krammer B, Rumbold K, Tschemmernegg M, Schwab H, et.al, „A novel screening assay for hydroxynitrile lyases suitable for high-throughput screening,“ *Journal of Biotechnology,* vol. 129, pp. 151-161, 207.

[53] Dr. Georg Steinkellner, AG Gruber.

[54] Georgiou CD, Grintzalis K, Zervoudakis G, Papapostolou I, „Mechanisem of Coomassie brilliant blue G-250 binding to proteins: a hydrophobic assay for nanogram quantities of proteins,“ vol. 391, Anal Bioanal Chem, 2008, pp. 391-403.

[55] Acker MG, Auld DS, „Considerations for the design and reporting of enzyme assays in high-throughput screening applications,“ vol. 1, no. (1-6), Perspectives in Sciences, 2014, pp. 56-73.

[56] Tanner A, Bornemann S, „Bacillus subtilis YvrK Is an Acid-Induced Oxalate Decarboxylase, “ vol. 182, no. 18, J Bacteriol, 2000, pp. 5271-5273.

[57] Dr. Kerstin Steiner, personal communication.

[58] Kuchner O, Arnold FH, „Directed Evolution of enzyme catalysts,“ vol. 15, TIBTECH, 1997, pp. 523-530.

[59] Wong TS, Roccatano D, Zaccharias M, Schwaneberg U, „A Statistical Analysis of Random Mutagenesis Methods Used for Directed Protein Evolution,“ vol. 355, no. 4, Journal of Molecular Biology, 2006, pp. 858-871.

[60] Brühne F, Weight E, „Benzaldehyde,“ vol. 5, Ullmann's Encyclopedia of Industrial Chemistry; Wiley-VCH Verlag, 2012, pp. 223-235.

[61] Betts MJ, Russel RB, „Amino Acid Properties and Consequences of Substitutions,“ Chapter 14, Bioinformatics for Geneticists, 2003, pp. 289-316.

7 List of figures

8 List of tables

9 Appendix

A: Protein concentrations and HNL-Activity-Assay raw data for site directed variants

Appendix Table 1: Protein concentrations of GtHNL variants.

Enyzme	total protein concentrations
	mg/mL
*Gt*HNL-Tr	20.74
*Gt*HNL_A40HV42T	12.15
*Gt*HNL_A40RV42T	10.80
*Gt*HNL-Tr_F19V	16.33
*Gt*HNL-Tr_W120R	9.65

Appendix Table 2: Parameters to calculate the specific enzyme activity on (R)-mandelonitrile.

Enzyme	oxalate buffer			citrate phosphate buffer		
	ΔA/t	thickness [cm]	dilution factor	ΔA/t	thickness [cm]	dilution factor
*Gt*HNL-Tr	0.0015	0.5340	210.00	0.1753	0.5180	210.00
	0.0038	0.5123	105.00	0.3015	0.5240	105.00
*Gt*HNL_A40HV42T	0.0905	0.4923	121.50	0.1769	0.5247	121.50
	0.1237	0.5030	60.75	0.2254	0.5163	60.75
*Gt*HNL_A40RV42T	0.1510	0.5465	1.08	0.1297	0.5430	1.08
	0.1444	0.5465	1.08	0.1315	0.5430	1.08
*Gt*HNL-Tr_F19V	0.0018	0.4993	65.32	0.0839	0.5140	65.32
	0.0026	0.5140	32.66	0.1628	0.5107	32.66
*Gt*HNL-Tr_W120R	0.0146	0.5267	0.97	0.0335	0.5472	0.97
	0.0115	0.5267	0.97	0.0468	0.5472	0.97

Appendix Table 3: Parameters to calculate the specific enzyme activity on (R)-2-chloromandelonitrile.

Enzyme	oxalate buffer			citrate phosphate buffer		
	ΔA/t	thickness [cm]	dilution factor	ΔA/t	thickness [cm]	dilution factor
*Gt*HNL-Tr	0.0017	0.4263	210.00	0.0351	0.4407	210.00
	0.0047	0.4343	105.00	0.0712	0.4430	105.00
*Gt*HNL-Tr_F19V	0.0286	0.4397	65.32	0.1651	0.4530	65.32
	0.0431	0.4353	32.66	0.3100	0.4443	32.66
*Gt*HNL-Tr_W120R	0.0354	0.4190	0.97	0.1317	0.4422	0.97
	0.0335	0.4190	0.97	0.1307	0.4422	0.97

Appendix Table 4: Parameters to calculate the specific enzyme activity on (S)-2-chloromandelonitrile.

Enzyme	oxalate buffer			citrate phosphate buffer		
	ΔA/t	thickness [cm]	dilution factor	ΔA/t	thickness [cm]	dilution factor
*Gt*HNL-Tr	-0.0003	0.4487	210.00	-0.0034	0.4400	210.00
	-0.0007	0.4373	105.00	-0.0030	0.4327	105.00
*Gt*HNL-Tr_F19V	-0.0035	0.4460	65.32	0.0060	0.4413	65.32
	-0.0045	0.4838	33.66	0.0060	0.4410	33.66
*Gt*HNL-Tr_W120R	0.0390	0.4542	0.97	0.0005	0.4452	0.97
	0.0274	0.4542	0.97	0.0043	0.4452	0.97

B: Protein concentrations and HNL-Activity-Assay raw data for site saturation variants

Appendix Table 5: Protein concentration of variants originating from site saturation libraries.

Enzyme	total protein concentration
	mg/mL
*Gt*HNL	
*Gt*HNL_F19T	7.92
*Gt*HNL_F19L	8.38
*Gt*HNL_F19C	9.24
*Gt*HNL_W120L	3.60
*Gt*HNL_W120E	6.17
*Gt*HNL_W120H	9.30
*Gt*HNL-Tr	9.26
*Gt*HNL-Tr_F19T	11.64
*Gt*HNL-Tr_F19A	8.10
*Gt*HNL-Tr_F19L	7.38
*Gt*HNL-Tr_W120P	9.50
*Gt*HNL-Tr_W120L	9.94
*Gt*HNL-Tr_W120K	7.23
*Gt*HNL-Tr_W120A	7.20

Appendix Table 6: Parameters to calculate the specific enzyme activity on (R)-2-chloromandelonitrile.

Enzyme	ΔA/t	thickness [cm]	dilution factor
*Gt*HNL	0.0181	0.4228	1.12
*Gt*HNL_F19T	0.0468	0.4233	1.06
*Gt*HNL_F19L	0.0478	0.4227	1.12
*Gt*HNL_F19C	0.0395	0.4298	1.23
*Gt*HNL_W120L	0.0122	0.4287	1.00
*Gt*HNL_W120E	0.0122	0.4337	1.03
*Gt*HNL_W120H	0.0224	0.4318	1.24
*Gt*HNL-Tr	0.0108	0.4253	92.60
*Gt*HNL-Tr_F19T	0.0255	0.4170	116.40

*Gt*HNL-Tr_F19A	0.0088	0.4193	81.00
*Gt*HNL-Tr_F19L	0.0133	0.4350	73.80
*Gt*HNL-Tr_W120P	0.0102	0.4313	95.00
*Gt*HNL-Tr_W120L	0.0202	0.4128	99.40
*Gt*HNL-Tr_W120K	-0.0048	0.4327	72.30
*Gt*HNL-Tr_W120A	0.0127	0.4320	72.00

Appendix Table 7: Parameters to calculate the specific enzyme activity on (S)-2-chloromandelonitrile.

Enzyme	ΔA/t	thickness [cm]	dilution factor
*Gt*HNL	0.0063	0.4275	1.12
*Gt*HNL_F19T	0.1639	0.4165	1.06
*Gt*HNL_F19L	0.2067	0.4200	1.12
*Gt*HNL_F19C	0.1498	0.4352	1.23
*Gt*HNL_W120L	0.0343	0.4167	1.00
*Gt*HNL_W120E	0.1051	0.4327	1.03
*Gt*HNL_W120H	0.0160	0.4247	1.24
*Gt*HNL-Tr	-0.0033	0.4262	5.79
*Gt*HNL-Tr_F19T	0.0047	0.4182	7.28
*Gt*HNL-Tr_F19A	0.0038	0.4320	5.06
*Gt*HNL-Tr_F19L	0.0002	0.4218	4.61
*Gt*HNL-Tr_W120P	0.0182	0.4227	5.94
*Gt*HNL-Tr_W120L	0.0316	0.4323	6.21
*Gt*HNL-Tr_W120K	0.0026	0.4507	4.52
*Gt*HNL-Tr_W120A	0.0104	0.4242	4.50