Angewandte Chemie
Eine Zeitschrift der Gesellschaft Deutscher Chemiker

Supporting Information

for

Angew. Chem. Int. Ed. Z52141

69451 Weinheim, Germany

Comprehensive step by step engineering of a (*R*)-hydroxynitrile lyase for large scale asymmetric synthesis**

Anton Glieder*, Roland Weis, Wolfgang Skranc, Peter Poechlauer, Ingrid Dreveny, Sandra Majer, Marcel Wubbolts, Helmut Schwab and Karl Gruber

Experimental

DNA isolation from *Prunus amygdalus*

Dried almond seeds (Farmgold, Lot L4532, 1999) were grinded with liquid nitrogen. 5 mL breaking buffer (100 mM NaAc; 50 mM EDTA; 500 mM NaCI, pH 5,5; 1,4% SDS and 20 µg/mL RNAse A) at 65 °C were added to 0.1 g frozen almond powder. After incubation and shaking at 65 °C for 15 min and following centrifugation (10 min at 7000 g) the supernatant was mixed with the same volume of 10 M ammonium acetate and then incubated on ice for 10 min. After centrifugation at 10.000 g for 15 min the supernatant was 2x extracted with phenol/chloroform (1/1, phenol equilibrated with 50 mM Tris, pH 8.0) After a further extraction with 2 volumes chloroform/isoamylalcohol (24/1) the DNA was precipitated with one volume of isopropanol, dried and dissolved in 50 µL water.

Cloning of the *Pahnl*5 gene into *Pichia pastoris*

Based on the sequences of the *mdl*5 gene from *Prunus serotina* (Hu und Poulton, 1999) and the *mdl*1 gene of *Prunus amygdalus* (Suelves et al., 1998) the *hnl*5 gene from *Prunus amygdalus* was cloned with the following PCR primers:
Primer 1: 5'- CGGAATTCACAATATGGAGAAATCAACAATGTCAG-3'
Primer 2: 5'- CGGAATTCTTCACATGGACTCTTGAATATTATG-3'

The PCR was performed in 50 µL with 1.2 U "Hotstar" Taq DNA Polymerase (Qiagen, Hilden, Germany), 50 ng genomic almond DNA, 200 ng each of primers 1 and 2, five µL of a dNTP mix (2 mM each) and 1x PCR buffer of the Hotstar kit (Qiagen, Hilden, D) and the following program: 15 min 95 °C , 30 cycles (1 min 95 °C, 30 sec 64 °C, 2 min 72 °C) and a final incubation step at 72 °C for 5 min. The 2.16 kb PCR product was directly sequenced, cloned into a pBSSK(-) vector, resulting in a plasmid, named pBSPamHNL5g and used for splicing by PCR. The DNA sequence was deposited in the genbank database under the accession number AY321296.

Three introns with the consensus sequence "GT.....AG" have been identified by sequence comparison with the cDNA sequences of published *mdl* genes. The introns of the *Prunus amygdalus hnl*5 gene have been spliced out by an overlap extension PCR strategy. Exons II and III were amplified with the primer pairs PamHNL5b/PamHNL5c (PCR1-1) and PamHNL5d/PamHNL5e (PCR1-2), respectively. The 50 µL PCRs contained 1x PCR buffer (Qiagen), 100 pmol of each primer, 2.5 U "Hotstar" Taq DNA polymerase (Qiagen), 5 µL dNTP mix (2 mM each) and 10 ng of plasmid pBSPamHNL5g. The following program was used: 15 min 95° C, 30 cycles for 1 min at 95 °C, 30 sec 68 °C, 1 min 72 °C and a final extension step for 5 min

at 72 °C. The products from PCR1-1 und PCR1-2 were purified by agarose gel electrophoresis and eluted from the gel using a Qiaexll Kit from Qiagen. For extension the product of PCR1-1 was amplified (PCR2) with the primers PamHNL5a2 and PamHNL5c with the following program: 15 min 95 °C, 30 cycles with 1 min 95 °C, 30 sec 68 °C, 1 min 72 °C and 1 time 5 min 72 °C. Also this PCR product was purified by isolation from an agarose gel. The product from PCR1-2 and PCR2 were combined by primerless PCR (PCR3: 5 cycles with 1 min at 94 °C, 30 sec at 68 °C and 1.5 min 72 °C, with ~100 ng of the products from PCR1-1 und PCR2 in a 50 µL reaction in 1x PCR buffer (Qiagen), 5 µL dNTP mix (2 mM each) and 2.5 U "Hotstar" Taq DNA polymerase (Qiagen). Extension to a full length product and amplification was made in a fourth PCR with 100 pmol of each of the primers PamHNL5a1 and PamHNL5f. They were directly added to PCR 3 and the following program was applied: 20 cycles with 1 min 95 °C, 30 sec 63 °C, 1.5 min 72 °C and finally 5 min 72 °C. The product was purified by isolation from an agarose gel, cloned into a pBSSK(-) vector and reamplified with the primers PCRHNL5-a und PCRHNL5-e. After purification this product was cut with *Eco*RI and cloned into the vector pHILD2 (Invitrogen, CA), resulting in plasmid pHILDPaHNL5a. *Pichia pastoris* GS115 was transformed following the standard protocol from Invitrogen.

PamHnl5a15'-
GAAGATCTGAATTCCATGGAGAAATCAACAATGTCAGTTATACTATTTGTGTTGCATCTTCTTG-3'

PamHnl5a2
5'-CTATTTGTGTTGCATCTTCTTGTTCTTCATCTTCAGTATTCAGAGGTTCACTCGCTTGCCAATACTTC-3'

PamHnl5b
5'-GTTCACTCGCTTGCCAATACTTCTGCTCATGATTTTAGCTACTTGAAGTTTGTGTACAACGCCACTG-3'

PamHnl5c
5'-GATGTATTGGAAGAGAAGAGGATCTTCTCTACT-3'

PamHnl5d
5' GATCCTCTTCTCTTCCAATACATCAAATTTGTCAGCTATTGGAGTCATATATACGG 3'

PamHnl5e
5'-CAACCGGATTGACCTTTCTTGCAGGATTTGAAGGCCCACATACCTTCCTAACATCAGATAGAAGCC-3'

PamHnl5f
5´-
AAGATCTGGAATTCTTCACATGGACTCTTGAATATTATGAATAGCCTCCAACCGGATTGACCTTTCTTGC
AG-3'

PCRHNL5-a
5'-
TCGAATTCGAGCTCGGTACCCGGGGATCCTCTAGAAATAATTTTGTTTAACTTTAAGAAGGAGATATACA
TATGGAGAAATCAACAATGTCAGTTATACTATTTGTGTTGCATC-3'

PCRHNL5-e
5'-CGAATTCGCCCTTTCGCATGCTCACATGGACTCTTGAATATTATGAATAGCCTC-3'

Change of secretory leader sequence and introduction of mutation L1Q

PCR I: The primer pair alpha11/alpha21aQ was used for the amplification of the alpha mating factor signal sequence from the cloning vector pPICZalpha (Invitrogen, CA). Primer alpha21aQ contains a sequence coding for the 5´end of the mature isoenzyme *Pa*HNL5 and mutation L1Q at the N-terminus of the expected mature secreted protein. At the same time, the *Eco*RI site at the end of the signal sequence from the template was removed. At the 3´end of the alpha factor signal sequence there was a Kex2 cleavage site and a GluAlaGluAla–sequence, that can be processed by proteases.

The PCR was performed in a 50 µL reaction (10 ng template, 0.1 µM of each primer, 0.2 mM dNTPs, 5 µL PCR buffer, 1 U *Pwo* polymerase from Roche). After denaturation for 2 min at 94 °C amplification was reached in 30 cycles (30 sec 94 °C, 60 sec 62 °C, 1 min 30 sec 72 °C) and a final step for 7 min at 72 °C.

PCR II: The *Pahnl*5 gene was amplified from plasmid pHILDPaHNL5a using the primer pair hnl5α11Q/hnl5α21. Primer hnl5α11Q contained a sequence, which was identical to the 3´ end of the product from PCR I.

This PCR was performed in a 50 µL reaction (10 ng template, 0.1 µM for each primer, 0.2 mM dNTPs, 5 µL PCR buffer, 1 U *Pwo* Polymerase from Roche (Mannheim, D). After denaturation for 2 min at 94 °C amplification was reached in 30 cycles (30 sec 94 °C, 60 sec 65 °C, 3 min 30 sec 72 °C) and a final step for 7 min at 72 °C. 3 µL of the products from PCR I und PCR II were used for combination of these two products by primerless PCR. Extension was made in a 45 µL reaction with 5 µL *Pwo* PCR buffer, 0.2 mM dNTPs and 1 Unit *Pwo* Polymerase (Roche). The reaction mix was heated to 94 °C for 2 min and then 10 cycles with 30 sec at 94 °C und 3 min at 72 °C were run in a thermocycler .

PCR III: For amplification of the overlap extension product by primers alpha11 and hnl5α21 the overlap extension PCR product was used as a template for amplification and 5 µL primer mix (3 µL water and 1 µL each of the primers alpha11 and hnl5α21, which had a concentration of 5 µM) were added. The product was amplified with 20 PCR cycles (30 sec 94 °C, 45 sec 62 °C, 4 min 72 °C) and a single temperature hold at 72 °C for 7 min. The PCR product was purified from an agarose gel and cloned into the expression plasmid pHILD2 (Invitrogen, San Diego, Ca). The resulting plasmid was named pHILDPaHNL5α_L1Q and transformed into *Pichia pastoris* GS115.

Primersequences:

oBT2835 alpha11: 5´- TCTTCGAAGAATTCACGATGAGATTTCCTTCAATTTTTACTGC- 3´
oBT2843 alpha21aQ: 5´- GAAGTATTGGCTTGAGCTTCAGCCTCTCTTTTCTCG-3´
oBT2842 hnl5α11Q: 5´- AGAGAGGCTGAAGCTCAAGCCAATACTTCTGCTCATG-3´
oBT2838 hnl5α21: 5´- ATGGTACCGAATTCTCACATGGACTCTTGAATATTATGAATAG-3´

Homology modeling and substrate docking

The structure of *Pa*HNL5 was modeled using the program Modeller v6.2[1] based on the known structure of *Pa*HNL1[2] (PDB-entry: 1ju2), which has a sequence identity of 76% with the target protein. The flavin cofactor as well as the bound substrate (*R*)-mandelonitrile, which was taken from the modeled complex[3], were included in the homology modeling procedure. The geometry of the resulting structure was further optimized using the program AMBER v7.0[4].

Molecular models of compounds (*R*)-**2b,** (*R*)-**2c** and (*R*)-**2d** were built and optimized using the program Sybyl v6.8 (Tripos Inc.) and were docked into the active site of PaHNL5 using the program AutoDock v3.0[4;5]. The protein was kept rigid, and the position and orientation of the substrates as well as two torsion angles (for the chlorophenyl- and the OH-group) were allowed to vary during the simulations. A hybrid genetic algorithm with phenotypic local search (designated

as a Lamarckian genetic algorithm[5]) was applied in all three cases. 25 independent simulations with populations consisting of 50 random structures were performed. The number of energy evaluations was limited to 500000 resulting in about 300 generations on average. The best individual of each generation automatically survived, the mutation and crossover rates were set to 0.02 and 0.80 respectively. The probability for performing a local search (up to 300 iterations) was 6%. A cluster analysis with an rmsd-cutoff of 1.5Å was performed and yielded energetically well separated optimal complex structures in all three cases.

Site directed mutagenesis

10 ng of plasmid pHILDPaHNL5α_L1Q as a template and 200ng of of each mutagenesis primer were used for site directed mutagenesis using the Quick Change XL Site Directed Mutagenesis kit from Stratagene, following the suppliers protocol. After denaturation for 1 min at 95 °C, 18 cycles with 50 sec at 95 °C, 50 sec at 60 °C and 20 min at 68 °C and a final extension step for 7 min at 68 °C, the template DNA was digested with *Dpn*I. 2 µL of the samples were used for transformation of *E. coli* XL10 Gold ultracompetent cells from Stratagene. The insert DNA sequence of the mutated plasmids was controlled by DNA sequencing. *Pichia pastoris* GS115 was transformed as described by the standard procedure from Invitrogen.

Primers for site directed mutagenesis:

oPaHNL5A111Gf: 5´-GTGGCACGACCATAATCAATGGAGGCGTCTACGCCAGAGCTAAC-3´

oPaHNL5A111Gr: 5´-GTTAGCTCTGGCGTAGACGCCTCCATTGATTATGGTCGTGCCAC-3´

oPaHNL5V317(A,G)f: 5´-TCCAATTGAAGCCTCTGTTGSAACTGTTTTAGGCATTAGAAGTG 3´

oPaHNL5V317(A,G)r: 5´ CTAATGCCTAAAACAGTTSCAACAGAGGCTTCAATTGGATTTGG 3´

S = G or C

Enzyme preparation

The best Mut^S clone of 100 transformants of each variant was selected and enzyme was produced by 250 mL shake flask cultures. Therefore 225 mL BM0.5G medium (0.2 M potassium phosphate, pH 6.0; 13.4 g /L Yeast Nitrogen Base; 5 g/L glycerol; 0.8 mg/L biotin) in 2 Liter baffled shake flasks were inoculated and shaked at 28 °C and 120 rpm. For induction 25 mL BMM10 medium (0.2 M potassium phosphate, pH 6.0; 13.4 g/L Yeast Nitrogen Base; 100 mL/L methanol; 0.8 mg/L Biotin) were added after ~60 h. 2.5 mL of pure methanol were added after further 10, 24 and 48 hours. After a total fermentation time of ~130 h and removal of the cell pellet after centrifugation, the culture supernatant was concentrated by ultrafiltration (30k) up to a concentration of ~1 mg/mL. For purification the procedure described by Dreveny et al.[2] was followed. The protein concentration was determined with the assay from Biorad, using *Pa*HNL from Sigma (M-6782 Lot 41H4016) as a reference.

Determination of enzyme activity and biocatalytic conversions

Benzaldehyde, 2-chloro-, 3-chloro- and 4-chlorobenzaldehyde were purchased from Aldrich – Benzaldehyde, 2-chlorobenzaldehyde and 3-chlorobenzaldehyde were distilled prior to use. Solvents and all other reagents were purchased from commercial sources and used without further purification.

Racemic cyanohydrins were prepared as described in the patent literature[6;7], via base catalyzed HCN addition to the aldhehyde.

The traditional chemical names as they are used in the communication and the corresponding CAS-names are listed in the following table:

Substance	Traditional name	CAS-name
2a	mandelonitrile	alpha-hydroxy benzeneacetonitrile
2b	2-Cl-mandelonitrile	2-chloro-alpha-hydroxy benzeneacetonitrile
2c	3-Cl-mandelonitrile	3-chloro-alpha-hydroxy benzeneacetonitrile
2d	4-Cl-mandelonitrile	4-chloro-alpha-hydroxy benzeneacetonitrile
3a	mandelic acid	alpha-hydroxy benzeneacetic acid
3b	2-Cl-mandelic acid	2-chloro-alpha-hydroxy benzeneacetic acid
3c	3-Cl-mandelic acid	3-chloro-alpha-hydroxy benzeneacetic acid
3d	4-Cl-mandelic acid	4-chloro-alpha-hydroxy benzeneacetic acid

The enzyme activity for cleavage of racemic mandelonitrile (2a) to 1a was determined at 25 °C by recording the kinetic slope at 280 nm photometrically. A substrate concentration of 12 mM 2a in 0.1 M citrate-phosphate buffer pH 5.0 was used. The enzyme was used in a dilution where a linear slope of ~0.1 per min was observed for the first 5 min. A molar extinction coefficient of 1.376 L $mmol^{-1}cm^{-1}$ was used for the calculation of specific enzyme activity.

Reactions on 15 mmol scale:
15 mmol scale reactions were used for all determinations of the turnover rates. Therefore 15 mmol substrate were dissolved in 2,1 mL tert-Butylmethylether (MTBE). 0.25 mg enzyme were diluted with 50 mM buffer (K_2HPO_4/citrate pH 3.4) to an end volume of 3.7 mL. Then the solution was set to a pH of 3.4 with citric acid again and mixed with the substrate/MTBE solution in 20 mL glass vials. The reaction mixture was cooled to 10°C, 1.2 mL HCN were added and the mixture was stirred with a magnetic stirrer at 700 rpm. After 5, 10, 15, 30, 60, 120 and 240 min 250 µL of the samples were taken, mixed with 100 µL MTBE and centrifuged. 50 µL of the organic phase were used for derivatization with 30 µL acetyl chloride in presence of 34 µL pyridine and 1 mL dichloromethane for 15 min at room temperature. The samples were analyzed by GC:

GC-Anaylsis of chiral Cyanohydrins:
Column: Chiraldex, Type: Chromopack, CP-Chirasil-Dex CB
Instrument: Perkin Elmer
Detector: FID

	Benzaldehyde	2-Chloro-benzaldehyde	3-Chloro-benzaldehyde	4-Chloro-benzaldehyde
Column length	25 m	5 m	25 m	25 m
Carrier Gas	Hydrogen	Helium	Hydrogen	Hydrogen
Pressure	74 kPa	74 kPa	74 kPa	74 kPa
T-Programme	1 min 100°C Ramp: 10°C/min to 160°C 3 min at 160°C	3 min 100°C Ramp: 10°C/min to 130°C	Isothermic 150°C	Isothermic 170°C
Aldehyde	0.92 min	0.55 min	1.61 min	1.33 min
R-Cyanohydrin-acetate	3.75 min	3.65 min	5.12 min	3.09 min
S-Cyanohydrin-acetate	4.37 min	3.81 min	5.81 min	3.7 min

The GC results were corrected with the response factors to give the conversion in mass %.

Preparative reactions were performed in a 100 mL Schmizo KPG reactor. 150 mmol substrate were dissolved in 21 mL MTBE. 5 mg enzyme were diluted to 37 mL with 50 mM buffer (K_2HPO_4/citrate pH 3.4) and set to pH 3.4 again. The substrate in MTBE was added and the mixture was stirred for 5 min. The temperature was kept at 10 °C, stirred with 900 rpm and over a time periode of 1 h HCN was pumped to the reaction mix. After completion of the reaction it was diluted with 140 mL MTBE, stirred for 5 min and the phases were separated after 10 min. The water phase was extracted with 40 mL MTBE again. The organic phases were pooled, stabilised by addition of citric acid and MTBE was evaporated. The yield and enantiomeric excess were determined by GC as described above.

Enzyme deglycosylation and determination of enzyme stability at low pH

Deglycosylation was performed using endoglycosidase H (EndoH) from New England Biolabs: 1 µL EndoH was added for 50 µg of protein, together with 1/10 volume of G5 buffer (0.5 M sodium citrate pH 5.5), and incubated for 16 hours at 37°C.

Samples were diluted in 0.03 M citrate/phosphate buffer pH 6.0 in a way, that after a further 70-fold dilution, they exhibited a slope of ~0.1 at 280 nm in the photometric determination of enzyme activity. 150 µL of this dilution were added to 900 µL of a 0.1 M citrate/phosphate buffer pH 2.6 and incubated at 25 °C. At given time-points 100 µl were used for photometric activity determination with mandelonitrile as described above but the buffer was 1 M.

References

[1.] A. Sali, T. L. Blundell, *J.Mol.Biol.* **1993,** *234* 779-815.

[2.] I. Dreveny, K. Gruber, A. Glieder, A. Thompson, C. Kratky, *Structure* **2001,** *9* 803-815.

[3.] I. Dreveny, C. Kratky, K. Gruber, *Protein Sci.* **2002,** *11* 292-300.

[4.] Case, D. A., Pearlman, D. A., Caldwell, J. W., Cheatham III, T. E., Ross, W. S., Simmerling, C. L., Darden, T. A., Merz, K. M., Stanton, R. V., Cheng, A. L., Vincent, J. J., Crowley, M., Ferguson, D. M., Radmer, R. J., Seibel, G. L., Singh, U. C., Weiner, P. K., and Kollman, P. A. AMBER. [7.0]. 2002. San Francisco, CA, University of California.

[5.] G. M. Morris, D. S. Goodsell, R. S. Halliday, R. Huey, W. E. Hart, R. K. Belew, A. J. Olson, *J.Comput.Chem.* **1998,** *19* 1639-1662.

[6.] Poechlauer, Peter, Wirth, Irma, Mayrhofer, Herbert, and Neuhofer, Rudolf. Method for the production of optically active cyanohydrins using R-oxynitrilase. DSM Fine Chemicals, Austria G. 2000-EP11753[0144487], 20. 25-11-2000. WO.

[7.] Schwab, H., Glieder, A., Kratky, C., Dreveny, I., Pöchlauer, P., Skranc, W., Mayerhofer, H., Wirth, I., Neuhofer, R., and Bona, R. Gene, enthaltend eine Hydroxynitrillyase codierende DNA-Sequenz, rekombinante Proteine mit Hydroxynitrillyase-Aktivität und deren Verwendung. 2002. Europa.