



シアン生成能を有する植物およびヤスデにおけるヒドロキシニトリ
ルリアーゼの産生並びにそれらの酵素化学

Occurrence and enzymology of hydroxynitrile lyases of cyanogenic plants
and millipedes

ニュイラート エム
Nuylert Aem

A Thesis Submitted in Partial Fulfillment of the Requirements for the
Degree of Doctor of Philosophy in Biotechnology
Toyama Prefectural University

2020年9月

CONTENTS

INTRODUCTION.....	1
CHAPTER I.....	5
Effect of glycosylation on the biocatalytic properties of hydroxynitrile lyase from the passion fruit, <i>Passiflora edulis</i> – a comparison of natural and recombinant enzymes	
CHAPTER II.....	22
Stabilization of hydroxynitrile lyases from two variants of passion fruits, <i>Passiflora edulis</i> Sims and <i>Passiflora edulis</i> forma <i>flavicarpa</i> by C-terminal truncation	
CHAPTER III.....	37
<u>Section I</u> : Effects of codon optimization and glycosylation on the high-level production of hydroxynitrile lyase from <i>Chamberlinius hualienensis</i> in <i>Pichia pastoris</i>	
<u>Section II</u> : Discovery and structural basis to improve the enzyme activity and enantioselectivity of hydroxynitrile lyase from <i>Parafontaria laminata</i> millipedes for (<i>R</i>)-2-chloromandelonitrile synthesis	
CHAPTER IV.....	74
Identification of saturated and unsaturated 1-methoxyalkanes from the Thai millipede <i>Orthomorpha communis</i> as potential “Raincoat Compounds”	
CONCLUSION.....	87
ACKNOWLEDGMENTS.....	90
REFERENCES.....	91
APPENDICES.....	99
PUBLICATIONS.....	110

INTRODUCTION

Hydroxynitrile lyases (HNLs, EC 4.1.2.10, 4.1.2.11, 4.1.2.46, and 4.1.2.47) act during the final step of the cyanohydrin biodegradation pathway.^{1, 2)} In cyanogenic plants, HNL plays a role in the defense mechanism of plants by acting decomposition of stored cyanogenic glycosides into aldehydes and hydrogen cyanide for protection against microbes and herbivores. In the cyanogenic millipedes, especially polydesmid species, their defense components are mainly produced by two enzymes [HNL^{3, 4)} and mandelonitrile oxidase (MOX)⁵⁾] from a mandelonitrile substrate stored in the reservoir of repugnatory glands. HNL is responsible for the production of HCN and benzaldehyde, which is further reduced to benzyl alcohol or oxidized to benzoic acid. MOX is responsible for the oxidation of mandelonitrile into benzoyl cyanide and the production of hydrogen peroxide as a by-product⁶⁾. Benzoyl cyanide reacts chemically with mandelonitrile to give mandelonitrile benzoate⁷⁾, and also with water to produce benzoic acid and HCN. As the reversible reaction of HNLs, the enzyme also catalyzes the synthesis of optically pure cyanohydrins which are important synthetic intermediates for fine chemical and pharmaceutical industries.⁸⁾

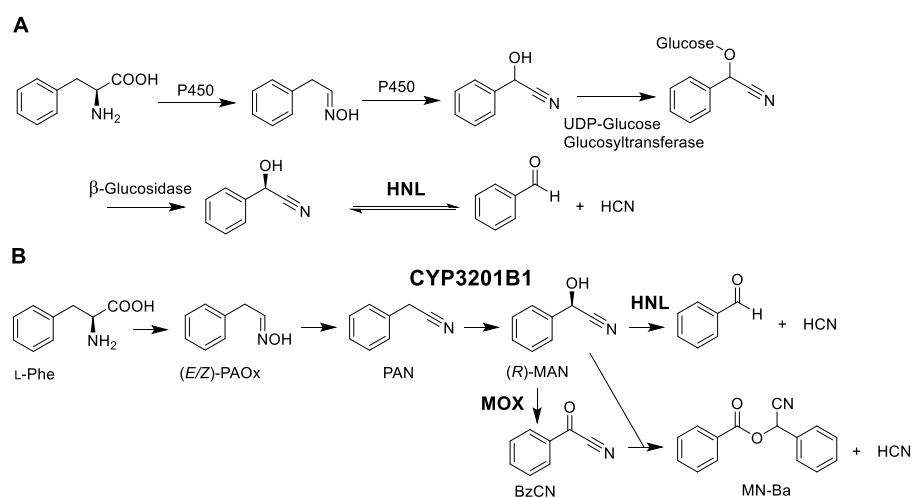


Figure 1. Cyanogenesis pathway of cyanogenic plants (A) and cyanogenesis pathway of millipedes, *Chamberlinius hualienensis* (B)²⁾. (A) Amino acids are converted into hydroxynitriles via aldoximes and nitriles by two cytochrome P450s. Hydroxynitriles are glycosylated to cyanogenic glycosides as a storage form. Cyanogenic glycosides are degraded to aldehydes or ketones and hydrogen cyanide (HCN) by the combination of β -glucosidase and hydroxynitrile lyase (HNL). (B) L-Phe is converted into (*R*)-mandelonitrile (MAN) via (*E/Z*)-phenylacetaldoxime (PAOx) and phenylacetone nitrile (PAN). The conversion of PAN into (*R*)-MAN is catalysed by CYP3201B1. HCN is produced from (*R*)-MAN via two pathways. One is the degradation of (*R*)-MAN to benzaldehyde and HCN by HNL. The other is the conversion of (*R*)-MAN into benzoyl cyanide (BzCN) by MOX. BzCN is spontaneously condensed with (*R*)-MAN to form mandelonitrile benzoate (MN-Ba) and HCN.

Mainly, HNLs have been discovered and characterized from cyanogenic plants, and they include the (*R*)-selective ones derived from *Prunus amygdalus* (*PaHNL*),⁹⁾ *Prunus serotina* Ehrh. (*PsHNL*),¹⁰⁾ *Prunus mume* (*PmHNL*),¹¹⁾ *Prunus armeniaca* L. (*ParsHNL*),¹²⁾ *Prunus communis* (*PcHNL*),¹³⁾ *Linum usitatissimum* (*LuHNL*),¹⁴⁾ *Phlebodium aureum* (*PhaHNL*),¹⁵⁾ *Arabidopsis thaliana* (*AtHNL*),¹⁶⁾ *Eriobotrya japonica* (*EjHNL*),¹⁷⁾ *Passiflora edulis* f. *flavicarpa* (*PeHNL*),¹⁸⁻²⁰⁾ *Davallia tyermannii* (*DtHNL*),²¹⁾ and *Nandina domestica* (*NdHNL*),²²⁾ and the (*S*)-selective ones derived from *Hevea brasiliensis* (*HbHNL*),²³⁾ *Manihot esculenta* (*MeHNL*),²⁴⁾ *Sorghum bicolor* (*SbHNL*),²⁵⁾ and *Baliospermum montanum* (*BmHNL*).²⁶⁾ Among the HNLs of plant origin, *SbHNL*,²⁵⁾ *EjHNL*,¹⁷⁾ *PeHNL*,²⁰⁾ *PaHNL*,²⁷⁾ *PcHNL*,¹³⁾ and *PmHNL*²⁸⁾ are glycosylated proteins, whereas, *LuHNL*, *BmHNL*, *AtHNL*, *NdHNL*, *HbHNL*, *MeHNL*, and *PhaHNL* are not glycosylated. Furthermore, *EjHNL*, *PaHNL*, *PcHNL*, and *PmHNL* contain flavin adenine dinucleotide (FAD), in contrast to the *LuHNL*, *PhaHNL*, *AtHNL*, *PeHNL*, *BmHNL*, *HbHNL*, *MeHNL*, *SbHNL*, and *NdHNL* which do not contain FAD. The cupin fold-containing HNLs derived have also been discovered from bacteria, *Granulicella tundricola* (*GtHNL*)²⁹⁾ and *Acidobacterium capsulatum* (*AcHNL*).³⁰⁾ Moreover, eleven new HNL genes were identified from five different genera of cyanogenic millipedes. Among the HNLs of arthropodal origin, the HNL from *Chamberlinius hualienensis*, henceforth referred to as *ChuaHNL*, was purified because it showed a high potential for industrial application^{3, 4)}. *ChuaHNL* showed the highest specific activity (7,420 U/mg) in the synthesis of (*R*)-mandelonitrile ((*R*)-MAN) as well as high enantioselectivity in the synthesis of various cyanohydrins.

Although the HNLs from natural one is the great potential as an industrial biocatalyst, it is difficult to collect sufficient amounts of natural resources and to purify HNL in industrial quantities. Thus, to enhance production yields, *Escherichia coli* and *Pichia pastoris* have been used as expression systems. Recombinant proteins from expression systems, however, vary in post-translational modifications such as glycosylation. N-Glycosylation is a eukaryote post-translational modification in which saccharides are covalently linked to the peptide backbone through a nitrogen atom of the asparagine residue in the tripeptide sequence Asn-X-Ser/Thr (where X can be any amino acid except proline). It has been known that the glycosylation played roles in improvement of enzymatic activity and molecular stability under various pH, temperature, and solvent conditions.

This dissertation was focused on the HNLs from plants and millipede origins. The smallest HNL from yellow passion fruit, *Passiflora edulis* forma *flavicarpa* (*PeHNL*-Ny) has discovered and

purified from the natural leaves. Base on N-terminal amino acid sequences of *PeHNL* analysis, the author cloned, expressed and compared the biocatalytic properties of recombinant *PeHNL* in *E. coli* and *P. pastoris* (Chapter I). *PeHNL* was also found in a purple one, *Passiflora edulis* Sims (*PeHNL*-Np) with only one amino acid difference, at position 107Thr in *PeHNL*-Ny, while it is Ala in *PeHNL*-Np. Consequently, *PeHNL*-Ny is glycosylated, but *PeHNL*-Np is not. The one glycosylation site located at high flexibility region C-terminus play a role on enzyme stability (Chapter II). Chapter III describe the study of HNL from two millipedes, *Chamberlinius hualienensis* (ChuaHNL) and a new discovering of HNL from *Parafontaria laminata* (PlamHNL). ChuaHNL was reported to be the first source of millipede HNL which showed highest specific activity. Since ChuaHNL could not be expressible in *E. coli* system, the focus of this chapter on ChuaHNL is develop of high-yield production system and characterized for recombinants ChuaHNL in *P. pastoris* (Chapter III; section I). The author also more discoveries of PlamHNL which showed high enantioselectivity in the synthesis of various cyanohydrins, especially toward (*R*)-2-chloromandelonitrile, a key intermediate for the anti-thrombotic agent clopidogrel which has a different characteristic than those of other millipede HNLs. PlamHNL was chosen to be a target of engineering to yield (*R*)-2-chloromandelonitrile with much higher enantiomeric excess, by using a computer aided substrate-docking simulation in the active site of PlamHNL, since an enantiomeric excess (*e.e.*) higher than 98% of the product was expected in the enzymatic process (Chapter III; section II).

The author searched for not only HNL, but also for the next target of the enzymatic research on the millipede. A new enzyme mandelonitrile oxidase (MOX) was found in ERATO studies by such an idea (Ishida et al., *Sci. Rep.* 2016). Like that, the author was looking for new metabolites other than benzaldehyde, benzoyl mandelate, a Schotten-Bauman reaction product. Dr. Kuwahara in ERATO has published on the detection of (2-nitroethyl) benzene and (*Z*)- and (*E*)-(2-nitroethenyl) benzenes in millipede and plant, and studied the biosynthesis of nitro compound from (*Z*)- and (*E*)-phenylacetaldoximes and phenylacetonitrile.

In the process of finding a new HNL in Thailand, the author has detected the HNL and MOX activities. The author isolated the DNA, mRNA, and did the cDNA synthesis there, but failed to clone the HNL gene, because primers made for Japanese millipedes did not work for selective cloning of HNL gene from Thailand. The author noticed that the smell of the millipedes was totally different from those from Japanese ones. The author extracted 10 bodies by hexane and brought the

extract to Japan and discovered the ether compounds. This dissertation newly reported a series of saturated and unsaturated wax-like components, other than conventional mixtures derived from mandelonitrile in the Thai millipede *Orthomorpha communis*. The wax-like components have not been reported in any other millipedes (Chapter IV).

CHAPTER I: Effect of glycosylation on the biocatalytic properties of hydroxynitrile lyase from the passion fruit, *Passiflora edulis* – a comparison of natural and recombinant enzymes

A hydroxynitrile lyase from the yellow passion fruit, *Passiflora edulis* (*PeHNL-N_y*) was isolated from the leaves and showed high stability in biphasic co-organic solvent systems for cyanohydrin synthesis. Cyanohydrins are important building blocks for the production of fine chemicals and pharmaceuticals. Thus, to enhance production yields of *PeHNL* for industrial applications, we cloned and expressed recombinant *PeHNL* in *E. coli* BL21 (DE3) and *P. pastoris* GS115 cells without a signal peptide sequence. The aim of the present study is to determine the effect of *N*-glycosylation on enzyme stability and catalytic properties in microbial expression systems. *PeHNL* from leaves (*PeHNL-N*) and *P. pastoris* (*PeHNL-P*) was glycosylated, whereas that expressed in *E. coli* (*PeHNL-E*) was not. The enzymes *PeHNL-N* and *PeHNL-P* showed much better thermostability, pH stability, and organic solvent tolerance than the deglycosylated enzyme *PeHNL-E* and the deglycosylated mutant N105Q from *P. pastoris* (*PeHNL-P-N105Q*). The glycosylated *PeHNL-P* also efficiently performed transcyanation of (*R*)-mandelonitrile with a 98% enantiomeric excess in a biphasic system with 2-isopropyl ether. These data demonstrate the efficacy of the present methods for improving enzyme expression and stability for industrial application based on *N*-glycosylation.

MATERIALS AND METHODS

Materials

Leaves of *P. edulis f. flavicarpa* were collected at the Botanic Gardens of Toyama (Toyama Prefecture, Japan) and stored at -20°C until use. Benzaldehyde (redistilled, 99.5%) and *racemic* mandelonitrile were purchased from Sigma Aldrich. All other chemicals were purchased from Wako Pure Chemical Industries (Osaka, Japan), Nacalai Tesque (Kyoto, Japan), or Kanto Chemical (Tokyo, Japan) and were used without further purification. Restriction endonuclease and the ligation reaction mixture were obtained from Takara Bio (Kusatsu, Japan).

HNL activity assays

Enzyme activities were measured using the methods reported in our previous study¹⁸⁾. Briefly, enzyme samples (0.5-5 U/mL) were added to 0.5 mL of 400 mM citrate buffer (pH 4) containing 25 mM benzaldehyde and 50 mM potassium cyanide, mixed and incubated at 25°C for 5 min. Subsequently, 100 μ L aliquots of reactant were transferred to 900- μ L mixtures of *n*-hexane:2-propanol with a volume ratio of 85:15, and mixed vigorously and centrifuged at $15000 \times g$ for 5 min at 4°C. Ten μ L aliquots of organic phase were then analyzed using an HPLC system (717 plus Autosample, Waters, Milford, MA USA; LC-6A liquid chromatograph pump, Shimadzu, Kyoto, Japan; column heater, Sugai U- 620, Japan; SPD-10A VP UV-vis detector, Shimadzu, Kyoto, Japan) equipped with a CHIRALCEL OJ-H column (particle size, 5 μ m; 4.6 mm i.d. \times 250 mm; Diacel, Osaka, Japan) under the following conditions: mobile phase, *n*-hexane:2-propanol with a volume ratio of 85:15¹⁹⁾; flow rate, 1 mL \cdot min⁻¹; absorbance, 254 nm; column oven temperature, 30°C. Benzaldehyde, (*R*)-mandelonitrile, and (*S*)-mandelonitrile retention times were 4.9, 10.2, and 12.7 min, respectively. Enzyme activity was calculated from the linear curve of the reaction in the first 5 min. One unit of HNL activity was defined as the amount of enzyme that produces 1 μ mol of optically active mandelonitrile from benzaldehyde and potassium cyanide per min.

Microorganisms

E. coli DH5 α (Toyobo, Osaka, Japan), JM 109 (Takara Bio), and BL21 (DE3),(Thermo Fisher Scientific, Waltham, MA, USA) cells, and *P. pastoris* GS115 cells were used as hosts for amplification of plasmid DNAs and recombinant protein expression.

Cloning of PeHNL cDNA

Young leaves were directly collected from plants, immediately frozen in liquid nitrogen, and ground in Plant RNA Reagent (Thermo Fisher Scientific). After isolation of total RNA, first-strand cDNA was synthesized using a SMART RACE cDNA Amplification kit, with a 5'-Full RACE Core Set (Takara Bio, Kusatsu, Japan), a GeneRacer Kit, and SuperScript II and III Reverse Transcriptases (Thermo Fisher Scientific). Degenerate and gene-specific primers were designed according to N-terminal amino acid sequences¹⁹⁾, and partial cDNA fragments were amplified using PCR (Appendix I). PCR products were gel-purified using a Wizard SV PCR and Gel Clean-Up System (Promega, Madison, WI, USA) and ligated into the *EcoRV* recognition site of pBluescript II

SK (+) (Agilent Technologies, Santa Clara, CA, USA). DNA sequences were determined using a 3500 Genetic Analyzer (Thermo Fisher Scientific), assembled and analyzed using ATGC and Genetyx Ver. 12 (Genetyx, Tokyo, Japan), respectively. To avoid PCR-derived sequence error, the full-length cDNA sequence was determined using 16 independent clones. The *N*-glycosylation site was predicted using NetNGlyc 1.0 software (<http://www.cbs.dtu.dk/>).

Construction of expression vectors of in recombinant PeHNL for expression in E. coli and P. pastoris

To express recombinant *PeHNL* in *E. coli* cells (*PeHNL*-E), insert DNA was PCR-amplified using 8F and 8R gene-specific primers (Appendix I) and *PeHNL* cDNA as a template. The products were then double-digested using *Sac*I and *Xba*I restriction enzymes, followed by gel-purification using a NucleoSpin (Macherey-Nagel, Germany). Purified insert DNA was ligated into the corresponding site of pCold I DNA (Takara Bio) using a DNA Ligation Kit (Mighty Mix; Takara Bio). The constructed vector was designated pColdI-*PeHNL*. To express recombinant *PeHNL* in *P. pastoris* GS115 cells, insert DNA was PCR-amplified using 9F and 9R gene-specific primers (Appendix I). After digestion using *Eco*RI and *Apa*I, gel-purified insert DNA was ligated into corresponding sites of pPICZ α A (Thermo Fisher Scientific). The constructed vector pPICZ α A-*PeHNL* was then linearized by digesting with *Sac* I and eletroporated into *P. pastoris* GS115 cells using an Easyselect Pichia Expression Kit (Thermo Fisher Scientific) according to the manufacturer's instructions. To determine the effects of glycosylation on the molecular and catalytic properties of the enzyme, the potential *N*-glycosylation site (asparagine residue at 105) was replaced with glutamine. Accordingly, the pPICZ α A-*PeHNL* plasmid was mutated using a QuikChange Lightning Site-Directed Mutagenesis Kit (Aligent Technologies) and 10F and 10R gene-specific primers (Appendix I), yielding the vector pPICZ α A-*PeHNL*-N105Q.

Expression and purification of recombinant PeHNLS in bacterial and Pichia expression systems

E. coli BL21 (DE3) cells harboring pColdI-*PeHNL* were inoculated into 5 mL of Luria-Bertani (LB) broth medium containing ampicillin (50 μ g mL⁻¹) and incubated overnight at 37°C with shaking at 300 rpm. Subsequently, 5-mL aliquots of starter culture were transferred into LB broth (500 mL) containing ampicillin (50 μ g mL⁻¹) in a 2-L Erlenmeyer flask and cultured at 37°C with shaking at 150 rpm for 6 h, followed by a cold-shock at 15°C. After 2 h, isopropyl- β -d-thiogalactopyranoside (IPTG) was added to a final concentration of 0.5 mM, and cells were cultured

at 15°C for 24 h with the same shaking rate. Cells were then harvested by centrifugation at $8500 \times g$ for 15 min and resuspended in 20 mM potassium-phosphate buffer (KPB; pH 7.4) containing 0.5 M sodium chloride and 25 mM imidazole. Cells were then lysed by sonication, and lysates were centrifuged at $15000 \times g$ for 15 min at 4°C to remove debris. Supernatants were loaded onto Ni Sepharose 6 Fast Flow (GE Healthcare, Little Chalfont, UK) columns (i.d. 25 mm; column volume, 30 mL) and eluted with a linear gradient of 25–500 mM imidazole in 20-mM KPB (pH 7.4) containing 0.5 M sodium chloride at a flow rate of $0.5 \text{ mL} \cdot \text{min}^{-1}$. Fractions containing the highest enzymatic activity were pooled, dialyzed, concentrated, and loaded onto a MonoQ 5/50 GL (GE Healthcare) HPLC column. Recombinant enzyme was eluted with a linear gradient of 0–200 mM sodium chloride in 20 mM KPB (pH 6.0) at a flow rate of $0.5 \text{ mL} \cdot \text{min}^{-1}$. Expression of the recombinant enzyme was monitored using SDS-PAGE³¹⁾ and enzyme activity was measured as described above.

Single *P. pastoris* colonies were transformed with pPICZ α A-*PeHNL* or pPICZ α A-*PeHNL*-N105Q and inoculated into 5 mL of YPD (1% yeast extract, 2% peptone, and 2% dextrose) broth containing Zeocin ($100 \mu\text{g} \cdot \text{mL}^{-1}$; Thermo Fisher Scientific). Inoculums were then incubated at 30°C overnight with shaking at 300 rpm. Five mL starter cultures were then transferred to YPD broth (500 mL) and cultivated at 30°C in 2-L baffled flasks. After 48 h incubation, cells were harvested by centrifugation and resuspended in 500 mL expression medium, which comprised buffered minimal methanol medium (BMMH) containing 100-mM potassium phosphate buffer (pH 6.0) 1.34% yeast nitrogen base without amino acid, $4 \times 10^{-5}\%$ biotin, 0.004% histidine, and 0.5% methanol as an inducer every 24 h. After 96 h of induction, cells were then harvested by centrifugation at $5000 \times g$ for 15 min, resuspended in 20-mM KPB (pH 6.0), and lysed using a Multi-Bead Shocker (Yasui Kikai, Osaka, Japan) with 0.5 mm diameter glass beads. After centrifugation at $15000 \times g$ for 15 min at 4°C to remove debris, supernatants were fractionated using 20%–60% saturated ammonium sulfate. The recombinant enzyme was then harvested by centrifugation at $18800 \times g$ for 20 min at 4°C and dissolved in 120 mL of 20-mM KPB (pH 6.0). After addition of ammonium sulfate to 30% saturation, solutions containing crude enzyme were centrifuged to remove debris and loaded onto a Toyopearl Butyl-650M (Tosoh; i.d., 25 mm; column volume, 40 mL). Proteins were eluted with a gradient of ammonium sulfate from 30% to 0% in the same buffer and the active fraction was dialyzed against 20-mM KPB (pH 6.0), loaded onto a Toyopearl DEAE-650M (Tosoh; i.d. 25 mm; column volume, 20 mL) column, and eluted with a

linear gradient of sodium chloride from 0 to 200 mM at a flow rate of 0.5 mL min⁻¹. After addition of ammonium sulfate to 30% saturation, active fractions were centrifuged at 15000 × g for 15 min at 4°C and supernatants were loaded onto a Resource Phe column (volume, 1 mL; GE Healthcare). Proteins were eluted with a gradient of ammonium sulfate from 30% to 0% in the same buffer at a flow rate of 0.5 mL·min⁻¹. Active fractions were pooled, concentrated, and desalted using a centrifugal filtration device (Amicon Ultra-15; 10,000 NMWL, EMD Millipore, Billerica, MA, USA), and proteins were loaded onto a MonoQ 5/50 GL column and eluted with a linear gradient of sodium chloride from 0 to 200 mM in 20 mM KPb (pH 6.0) at a flow rate of 0.5 mL·min⁻¹. The positive fraction was re-purified using a Resource Phe as described above. Protein elution was monitored at 280 nm in each purification step, and protein fractions were analyzed using SDS-PAGE. All purification steps were performed at 4°C.

Purification of PeHNL from leaves

Leaves of *Passiflora edulis* (1500 g wet weight) were collected at the Botanic Garden of Toyama and native PeHNL was purified as described previously.¹⁹⁾ Proteins were fractionated with 20%–40% saturated ammonium sulfate and collected by centrifugation at 18800 × g for 20 min at 4°C. Harvested proteins were dissolved in 100 mL of 20 mM KPb (pH 6.0) and dialyzed against the same buffer three times. The resulting fractions were loaded onto a DEAE Toyopearl-650M column (i.d., 25 mm; column volume, 50 mL) and eluted with a linear gradient of sodium chloride from 0 to 100 mM in 20 mM KPb (pH 6.0) at a flow rate of 0.5 mL min⁻¹. After addition of ammonium sulfate to 30% saturation, positive fractions were loaded onto a Butyl-Toyopearl (Tosoh; i.d., 25 mm; column volume, 25 mL) column. Proteins were then eluted stepwise with 1330, 860, 420, and 0 mM ammonium sulfate in 20 mM KPb (pH 6.0) at a flow rate of 0.5 mL·min⁻¹. Active fractions were pooled, concentrated, and desalted using a centrifugal filtration device (Amicon Ultra-15; 10,000 NMWL, EMD Millipore) and then loaded onto a MonoQ 10/100 GL and eluted with a linear gradient of sodium chloride from 0 to 150 mM in 20 mM KPb (pH 6.0) at a flow rate of 1.0 mL min⁻¹. Positive fractions were repurified using Resource Phe as described above. Subsequently, active fractions were concentrated in fresh 10 mM KPb (pH 6.0) using a centrifugal filtration device (Amicon Ultra-15; 10,000 NMWL, EMD Millipore) and loaded onto a Ceramic Hydroxyapatite type I column (i.d., 7 mm; column volume, 2 mL; BioRad, Hercules, CA, USA). Proteins were then eluted with a gradient of 10 to 500 mM KPb (pH 6.0) at a flow rate of 0.5 mL min⁻¹. After concentration and replacement of buffer with 20 mM Tris-HCl (pH 7.4) containing

0.5-M sodium chloride using a centrifugal filtration device, enzymes were loaded onto a HiTrap Con A 4B (1 ml; GE Healthcare) and eluted with a linear gradient of 0–100 mM α -D-methylglucoside in the same buffer at a flow rate of 0.3 mL min⁻¹.

Periodic acid-Schiff (PAS) staining

After separation of each *PeHNL* using SDS-PAGE, glycosylation of purified *PeHNLs* was detected using periodic acid-Schiff (PAS) staining with a GelCode Glycoprotein Staining Kit (Thermo Fisher Scientific) according to the manufacturer's instructions.

Measurement of kinetic parameters

The initial velocity of the enzymatic synthesis of benzaldehyde to (*R*)-mandelonitrile was assayed in 400 mM sodium citrate buffer, pH 4.0, according to method as described above at various substrate concentrations.

Measurement of pH and temperature stability

To determine pH stability, each *PeHNL* was preincubated at 30°C in the range of pH 3.5 to 10 (40 mM) for 1 h. On the other hand, in determination of thermostability the enzyme was preincubated in 20 mM potassium-phosphate buffer, pH 6.0 in the range of 30°C to 80°C for 1 h. The remaining activity of the enzyme was measured as described above.

Measurement of organic solvent stability

Each organic solvent (0.75 mL) ethyl acetate (EA), diethyl ether (DEE), methyl-*t*-butyl ether (MTBE), 2-isopropyl ether (DIPE), dibutyl ether (DBE), and hexane (HEX) was mixed with equal volumes of 400 mM citrate-phosphate buffer (0.75 mL; pH 4.0) and equilibrated with shaking at 1500 rpm for 60 min. Native or purified recombinant *PeHNLs* were added to the citrate-phosphate buffer phase, gently mixed without disturbing the interface between the aqueous and organic phases. The enzyme activity at time zero were measured under the saturated solvent in aqueous phase and then incubated for 12 h at 30°C with shaking at 1500 rpm. Remaining enzyme activities were analyzed as described above.

To investigate transcyanation reactions in biphasic systems, organic phases containing benzaldehyde (250 mM) were mixed with 400 mM citrate buffer (pH 4.0) containing 2.34 U of each

partially purified *PeHNL* in a total volume of 1.5 mL in 2.0-mL micro-tubes. Reactions were initiated by adding acetone cyanohydrin (900 mM), and the mixture was incubated at 10°C with shaking at 1500 rpm in an incubator shaker (BioShaker M-BR-022UP, Taitec Corporation, Tokyo, Japan). Aliquots of sample (50 μ L) were collected from the organic phase at various times and analyzed using HPLC with a chiral column as described above, and ee were determined according to relative concentrations (mM) of the two enantiomers using Eq. (1):

$$ee (\%) = \frac{[R]-[S]}{[R]+[S]} \times 100 \quad (1)$$

Reusability

In the studies on the stability and transcyanation of *PeHNL* in biphasic system, MTBE was selected for *PeHNL*-N while DIPE was chosen for *PeHNL*-P, *PeHNL*-P-N105Q and *PeHNL*-E. All enzymes were used in the batch transcyanation reaction of benzaldehyde (250 mM) and acetone cyanohydrins (900 mM) in biphasic system of 400 mM citrate buffer, pH 4.0 (0.75 mL) and MTBE or DIPE (0.75 mL) in 2.0-mL micro-tube. Reactions were initiated by addition of each *PeHNL* (5.0 U) into each biphasic system and incubated at 10°C with shaking at 1500 rpm. After 3 h incubation, the buffer phase containing the enzyme was recovered and dialyzed against the 400 mM citrate buffer, pH 4.0 (100 mL of buffer; 3 h at 4°C) and then was re-used for the next batch reaction under the same condition. Amount of (*R*)-mandelonitrile produced was measured as described above.

RESULTS AND DISCUSSION

Cloning of HNL cDNA from Passiflora edulis

Prior to expression of recombinant enzymes, full-length *PeHNL* cDNA was cloned and its sequence was determined using RACE with degenerate primers that were designed on the basis of N-terminal amino acid sequences of the purified enzyme, and with gene-specific primers and cDNA from a young leaf. The open reading frame of *PeHNL* comprised 444 bp encoding 147 amino acid residues including a 26 amino acid signal peptide (Accession No. LC115048; Figure 1-1). The predicted N-terminal amino acid sequence from the cDNA was identical to that of purified *PeHNL* from leaves.¹⁹⁾ Moreover, the predicted mature protein (no signal peptide) had a molecular mass of 14066.92 Da, an isoelectric point of 5.20. The mass of *PeHNL* from leaves was confirmed using

SDS-PAGE analysis (Figure 1-3A).¹⁹⁾ (*R*)-*PeHNL* is smaller than (*R*)-HNLs from *Eriobotrya japonica*, *Prunus amygdalus*, *Linum usitatissimum*, and *Phlebodium aureum*, which have molecular masses ranging from 62 to 108 kDa.³²⁻³⁶⁾ Although *PeHNL* has low sequence homology with previously characterized HNLs, it has 37%, 32%, and 28% amino acid identity with boiling-stable proteins from aspen *Populus tremula* (accession number, 1SI9_A)³⁷⁾, putative At3q1720 from the *Arabidopsis thaliana* genome (accession number, 1Q4R_A)³⁸⁾, and olivetolic acid cyclase from *Cannabis sativa* (accession number, 5B08A)³⁹⁾, respectively (Figure 1-2). The predicted mature enzyme had an *N*-glycosylation site on the asparagine residue at 105 (Figure 1-1).

```

M R N P G E T F L I K H F L V L S L F L 20
ATGAGAAACC CAGGAGAGAC ATTTCTGATT AAGCATTTC CCTTTCTCTC 60

C A G T A H N P P E I V R H I V F N R Y 40
TGC GCAGGAA CGGCACACAA CCCTCCGGAG ATAGTGAGGC ACATAGTATT TAATCGATAC 120

K S Q L S L T P I D Q I I A D Y G N L Q 60
AAATCGCAAC TCTCTCTTAC ACCAATTGAT CAGATCATCG CAGACTATGG CAACCTGCAA 180

N I A P E M K E W K W G T D L G P A V E 80
AACATTGCC CCGAGATGAA GGAATGGAAG TGGGGTACAG ACTTGGGACC GGCGGTGGAG 240

D R A D G F T H A Y E S T F H S V A D F 100
GACAGAGCTG ATGGTTTCAC ACATGCATAT GAATCGACAT TTCACAGTGT CGCAGACTTT 300

L N F F Y S P P A L E F A K E F F P A C 120
CTCAATTTCT TTTATAGTCC TCCTGCCCTC GAATTTGCAA AAGAATTTT TCCTGC GTGT 360

E K I V V L N Y I I N E T F P Y T W A L 140
GAAAAAATCG TCGTGCTCAA CTACATAATC AACGAGACTT TTCCATACAC ATGGGCACTT 420

P N K Y V V T * 147
CCAAACAAAT ACGTAGTCAC CTGA 444

```

Figure 1-1. cDNA and deduced amino acid sequences of hydroxynitrile lyase from the passion fruit *P. edulis*. The underlined portion is the 26 amino acid signal peptide. The prediction of one *N*-glycosylation site is marked with a thick triangle.

```

PeHNL      1 NPPEIVRHIVFNRYKSQLSQKQIDQIIADYGNLQNIAPEMKEWKWGTDLG 50
PtreBSP    1 --PKLVKHTLLTRKDEITREQIDNYINDYTNLLDLIPSMKSFNWGTDLG 48
At3q1720   1 ----VVRHVLVAKFKQEI SQEETEKLIKGYANLVNLIPPMKELHWGTELG 46
CstaOAC    1 ----VRHIVIAKFKDEISQEKIDELIKGYANLVNLIPPMKSFHWGTDVS 45

PeHNL      51 PAVEDRADGFTHAYESTFHSVADFLNFFYSPPALEFAKEFFPACEKIVVL 100
PtreBSP    49 MESAELNRGYTHAFESTFESKSGLQEYLD SAALAAFAEGFLPTLSQRLVI 98
At3q1720   47 LAN--MDQGFTHLESTFESTEGVAEYIAHPAHVAFANIFLPALEKFLVI 94
CstaOAC    46 AEN--LHQGFTHVFSNFESVEAIAEYVAHPAHVEYANLFLSNLEKVI AI 93

PeHNL      101 NYIINETFPYTWALPNKYVVT 121
PtreBSP    99 DYFLY----- 103
At3q1720   95 DY----- 96
CstaOAC    94 DY----- 95

```

Figure 1-2. ClustalW alignment of *PeHNL* and homologous proteins; PtreBSP, boiling stable protein from aspen, *Populus tremula* (accession number, 1S19_A); At3q1720, predicted protein from *Arabidopsis thaliana* genome (accession number, 1Q4R_A); CstaOAC, olivetolic acid cyclase from *Cannabis sativa* (accession number, 5B08_A). Identical residues among 4 and 3 proteins are indicated by black and gray boxes, respectively.

Purification of *PeHNLs*

To investigate the properties of *PeHNL*, native *PeHNL* was purified from leaves (*PeHNL*-N) and compared with recombinant *PeHNL* without posttranslational modification (*PeHNL*-E), recombinant *PeHNL* with posttranslational modification (*PeHNL*-P), and *PeHNL* with an asparagine to glutamine substitution at amino acid 105 (*PeHNL*-P-N105Q), which were expressed and purified using bacterial and *Pichia* expression systems. The purification steps for natural HNL from the leaves of *P. edulis* (*PeHNL*-N) are summarized in Table 1-1. A total of 12200 activity units of crude *PeHNL*-N enzyme were extracted from 1500 g of leaves. Subsequently, *PeHNL*-N was purified to 13.4-fold with a 0.04% recovery yield and specific activity of 134 U·mg⁻¹ after ammonium sulfate fractionation and chromatography through six columns (DEAE-Toyopearl, Butyl-Toyopearl, MonoQ 10/100 GL, Resource Phe, Hydroxyapatite type I, and Con A Sepharose 4B).

To express proteins in *E. coli*, cDNA fragments were cloned into the pColdI vector, which adds an N-terminal His-tag. The resulting recombinant enzyme from *E. coli* (*PeHNL*-E) was purified to 22.6 fold with a 1.17% recovery yield using a Ni Sepharose 6 Fast Flow column followed by monoQ HR10/100. The specific activity of *PeHNL*-E was 150 U·mg⁻¹ (Table 1-1), and was slightly higher than those of *PeHNL*-N, *PeHNL*-P, and *PeHNL*-P-N105Q.

Table 1-1. Purification summary of *PeHNLs*

Purification step	Total activity (U)	Specific activity (U/mg)	Recovery (%)	Purification fold
<i>PeHNL-N</i>				
Crude (leaves 1,500g)	12200	9.96	100	1.00
20-40% (NH ₄) ₂ SO ₄	5320	11.9	43.6	1.20
DEAE Toyopearl	4220	12.2	34.5	1.23
Butyl Toyopearl	2560	16.8	32.3	1.69
MonoQ HR10/100	360	39	4.5	3.91
Resource PHE	42.2	52.4	0.53	5.26
Hydroxyapatite type II	14.7	58.7	0.18	5.90
Con A Sepharose 4B	3.4	133.6	0.04	13.4
<i>PeHNL-P</i>				
Cell-free extract (1L)	1065	0.74	100	1.00
20-60% (NH ₄) ₂ SO ₄	540	1.06	50.7	1.44
Butyl Toyopearl	527	2.37	49.5	3.20
DEAE Toyopearl	146	17.6	13.7	23.8
Resource PHE 1 st	63.5	36.8	5.97	49.8
MonoQ HR10/100	14.7	59.6	1.40	80.0
Resource PHE 2 nd	3.62	134	0.34	182
<i>PeHNL-P-N105Q</i>				
Cell-free extract (1L)	263	0.22	100	1.00
20-60% (NH ₄) ₂ SO ₄	118	0.94	44.8	4.30
Butyl-Toyopearl	86.5	3.25	32.8	14.7
DEAE-Toyopearl	13.8	16.4	5.25	74.5
Resource PHE	2.10	140	0.80	636
<i>PeHNL-E</i>				
Cell-free extract (1L)	5440	6.64	100	1.00
Ni sepharose	2270	47.8	51	7.2
MonoQ HR10/100	64.1	150	1.17	22.6

Molecular masses and glycosylation statuses of native PeHNL, recombinant PeHNLs, and mutant PeHNL

The proteins *PeHNL-N*, *PeHNL-P*, *PeHNL-P-N105Q*, and *PeHNL-E* had molecular masses of about 14, 17, 13, and 13 kDa, respectively, in SDS-PAGE analyses (Figure 1-3A). In subsequent experiments, *PeHNL-N* and *PeHNL-P* were reacted with periodic acid-Schiff (PAS) reagent and eluted through a Con A Sepharose 4B column (Figure 1-3B). These experiments indicated total carbohydrate contents of 7% and 30% in *PeHNL-N* and *PeHNL-P*, respectively. These data were consistent with the higher molecular mass of *PeHNL-P* compared with *PeHNL-N* (Figure 1-3A), likely reflecting the high mannose-type glycosylation that is commonly observed in *Pichia pastoris*.⁴⁰⁾ In contrast, PAS staining experiments did not show glycosylation of *PeHNL-P-N105Q*, suggesting that the Asn105 in *PeHNL* is the only glycosylation site in the present *Pichia* expression system.

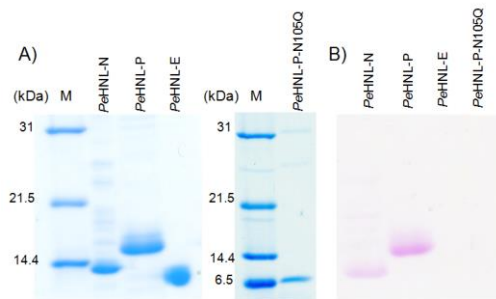


Figure 1-3. SDS-PAGE analysis and periodic acid-Schiff staining of native and recombinant *PeHNL*s; A, Coomassie Brilliant Blue staining; B, Periodic acid-Schiff (PAS) staining; M, molecular marker; *PeHNL-N*, native *PeHNL* isolated from leaves; *PeHNL-P*, recombinant *PeHNL* expressed in the *Pichia* expression system; *PeHNL-P-N105Q*, recombinant *PeHNL* lacking the glycosylation site at Asn 105 and expressed in the *Pichia* expression system; *PeHNL-E*, recombinant *PeHNL* expressed in the bacterial expression system.

Kinetic parameters of native PeHNL, recombinant PeHNLs, and mutated PeHNL

To investigate enzyme kinetics of *PeHNL-N* and recombinant *PeHNL*s, the synthesis of (*R*)-mandelonitrile from benzaldehyde and potassium cyanide was monitored in the presence of native and recombinant *PeHNL*s using HPLC-chiral column analyses. In these experiments, K_m , V_{max} , K_{cat} , and K_{cat}/K_m values of *PeHNL-N* were 11.2 ± 0.0 mM, 217 ± 2.7 $\mu\text{mol min}^{-1} \text{mg}^{-1}$, 50.9 ± 2.6 s^{-1} , and 4.56 ± 0.36 $\text{mM}^{-1} \cdot \text{s}^{-1}$, respectively, and were comparable to those of *PeHNL-P*. However, K_m , V_{max} , K_{cat} , and K_{cat}/K_m values of *PeHNL-E* were 26.1 ± 0.3 mM, 164 ± 1.8 $\mu\text{mol min}^{-1} \text{mg}^{-1}$, 38.4 ± 0.4 s^{-1} , and 1.47 ± 0.05 $\text{mM}^{-1} \cdot \text{s}^{-1}$, respectively, and these values were similar to those of *PeHNL-P-N105Q*, which lacks the potential *N*-glycosylation site (Table 1-2). K_m values of *PeHNL-E* and *PeHNL-P-N105Q* were two times higher than those of *PeHNL-N* and *PeHNL-P*. Moreover, V_{max} , K_{cat} , and K_{cat}/K_m of *PeHNL-N* and *PeHNL-P* were 1.3, 1.3, and 2.5 times higher, respectively, than those of *PeHNL-E* and *PeHNL-P-N105Q* (Table 1-2). These results indicate decreased K_m and increased V_{max} , K_{cat} , and K_{cat}/K_m values following glycosylation of *PeHNL*, suggesting that glycan participates in protein conformation changes that increase activity without reducing substrate affinity for benzaldehyde and KCN. Accordingly, previous studies suggested that glycans interact and form covalent bonds with the protein structure.^{41, 42} However, future protein structure and glycan analyses are warranted to investigate the related active site mechanisms.

Table 1-2. Kinetic parameters of native and recombinant *PeHNLs*

	K_m (mM)	V_{max} ($\mu\text{mol min}^{-1} \text{mg}^{-1}$)	k_{cat} (s^{-1})	k_{cat}/K_m ($\text{mM}^{-1} \cdot \text{s}^{-1}$)
<i>PeHNL-N</i>	11.2 \pm 0.0	217 \pm 2.7	50.9 \pm 2.6	4.56 \pm 0.36
<i>PeHNL-P</i>	13.0 \pm 0.1	222 \pm 1.7	52.1 \pm 1.3	4.00 \pm 0.17
<i>PeHNL-P-N105Q</i>	19.8 \pm 0.0	169 \pm 0.4	39.7 \pm 1.6	2.00 \pm 0.16
<i>PeHNL-E</i>	26.1 \pm 0.3	164 \pm 1.8	38.4 \pm 0.4	1.47 \pm 0.05

pH, temperature, and organic solvent stabilities of native PeHNL and recombinant PeHNLs

Industrial biotechnology generally requires stable enzymes with sustained activity under varying pH, temperature, and organic solvent conditions.⁴³⁻⁴⁵⁾ Previous studies indicate that glycosylation improves enzyme stability.⁴⁶⁻⁴⁸⁾ Accordingly, glycosylation of β -glucuronidase in *P. pastoris* provides greater stability between pH 4.0 and 7.0 than that in *E. coli*.^{42, 49)} Thus, to determine the effects of glycosylation on *PeHNL* stability, we determined enzyme activities under various conditions of pH, temperature, and organic solvent. In analyses of the effects of pH, activities of all *PeHNLs* were stable over the range from pH 3.5 to 10.0, although the activity of *PeHNL-E* was reduced to 60% and 70% at pH 7.0 and 10.0, respectively (Figure 1-4A). These results suggest that glycosylation does not increase pH stability of *PeHNL*. Whereas, the one glycosylation site of hydroxynitrile lyase isoenzyme 5 from almonds (*PaHNL5*) played a key role in stability at low pH when expressed in *P. pastoris*.⁵⁰⁾ To avoid nonselective synthesis of chiral cyanohydrins, low pH conditions are necessary for enantioselective synthesis of cyanohydrins.⁴³⁾ Thus, the present data warrant further consideration of *PeHNLs* as suitable biocatalysts for synthesis of pharmaceutical and agrochemical compounds. This tolerance of lower pH may allow control over the synthesis of unstable cyanohydrins by reducing spontaneous chemical degradation.

In further analyses, all *PeHNLs* showed 100% activity at temperatures of 30°C to 50°C following 1 h preincubation. However, enzyme activities of *PeHNL-E* and *PeHNL-P-N105Q* were decreased at 55°C and lost at 60°C. Although both *PeHNL-P* and *PeHNL-N* had 90% activity at 55°C, *PeHNL-P* activity was reduced to 60% at 60°C and lost at 70°C, whereas *PeHNL-N* had 80% activity at 60°C and enzyme products remained detectable at 80°C (Figure 1-4B). In direct comparisons at 60°C, *PeHNL-P* and *PeHNL-N* had between 60% and 80% remaining activity, whereas *PeHNL-E* and *PeHNL-P-N105Q* were almost inactive (Figure 1-4B). In addition, *PeHNL-N* had 40% remaining activity at 70°C, whereas *PeHNL-P* did not. These results suggest that *N*-glycosylation contributes to the thermostability of *PeHNL-P* and *PeHNL-N*. Similarly, the

hyperthermostability of *PeHNL-N* may reflect specific glycan compositions of passion fruit enzymes, which differ from those in *PeHNL-P* from the *P. pastoris* expression system. Meldgaard and Svendsen analyzed irreversible thermal denaturation using *Bacillus amyloliquefaciens* (1,3-1,4)- β -glucanases, *B. macerans* (1,3-1,4)- β -glucanases, and chimeric enzymes. In agreement with the present data, thermostabilities of these enzymes increased with glycosylation, although these effects were multifactorial.⁵¹⁾ Fonseca *et. al.* also showed that glycoproteins are more stable than their corresponding non-glycosylated forms, owing to steric interactions between sugar residues and protein structures and effects of glycosylation on the protein energy landscape.⁵²⁾

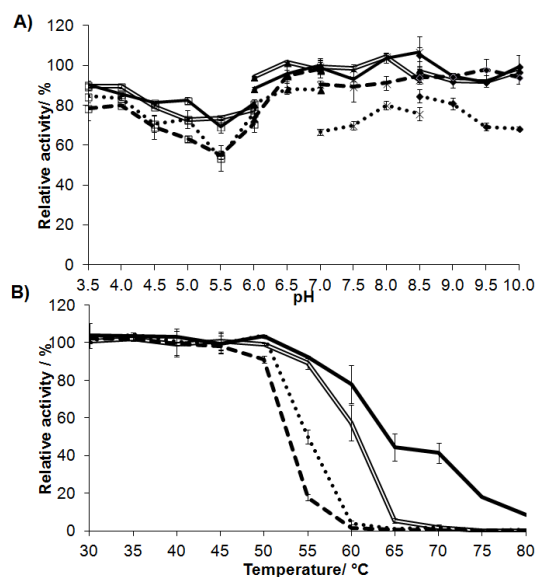


Figure 1-4. pH (A) and temperature stability (B) of native and recombinant *PeHNLS*; Solid line *PeHNL-N*; double line, *PeHNL-P*; dashed line, *PeHNL-P-N105Q*; dotted line, *PeHNL-E*; □, citrate buffer (40 mM); ▲, phosphate buffer (40 mM); ×, Tris-HCl buffer (40 mM); ◆, glycine-sodium hydroxide buffer (40 mM). After incubation of enzymes at various pH or temperatures for 1 h, the remaining enzyme activity was estimated by monitoring the synthesis of (*R*)-mandelonitrile from benzaldehyde and potassium cyanide using HPLC with a chiral column. Data are presented as means \pm SD (n = 3).

According to our previous work, pH and temperature were important parameters for controlling the enantiomeric purity of cyanohydrin products, and the choice of organic solvent was essential to activity and enantioselectivity of enzymes.¹⁹⁾ Thus, in the present study, we investigated stabilities of native and recombinant *PeHNLS* following fractionation using organic solvents in biphasic systems comprising 400 mM citrate buffer (pH 4.0) with ethyl acetate (EA), diethyl ether (DEE), methyl-*t*-butyl ether (MTBE), diisopropyl ether (DIPE), dibutyl ether (DBE), or *n*-hexane (Hex) at a volume ratio of 50:50. Following incubation at 10°C for 12 h, *PeHNL-N* and *PeHNL-P*

stabilities were comparable, with remaining enzyme activities of 60%–94%. In contrast, *PeHNL-P-N105Q* and *PeHNL-E* activities varied between 70% and 0% in all biphasic systems apart from those with DEE and DIPE. Interestingly, both of non-glycosylated *PeHNL-E* and *PeHNL-P-N105Q* lost the activities by 50% in aqueous buffer with shaking at 1,500 rpm at 10 °C for 12 h (Figure 1-5) indicating that the presence of glycan could prevent the self-degradation at 10 °C of this enzyme. In the aqueous system, water is such major deleterious reactions as deamidation of Asn/Gln residues and hydrolysis of peptide bonds.⁵³⁾ Therefore, the enzyme could be more stable in organic solvent than in water.⁵⁴⁾ This hypothesis is supported by the evidence that the non-glycosylated enzyme, *PeHNL-P-N105Q* and *PeHNL-E* were found to be more stable in the presence of Hex or DBE than in the absence of an organic solvent. The stability of the enzyme in organic solvents depends on the hydrophobicity of the organic solvent. In general, enzymes only need a thin layer of water on the surface of the protein to retain their catalytically active conformation. Moreover, specific activities of the glycosylated enzymes *PeHNL-N* and *PeHNL-P* were significantly higher than those of the unglycosylated enzymes *PeHNL-P-N105Q* and *PeHNL-E* under conditions of buffer alone and with EA or MTBE. However, no differences in remaining activity were identified between the 4 HNLs in biphasic systems containing DBE or Hex (Figure 1-5). To investigate the relationship between glycan composition and thermostability and stability of *PeHNL* in organic solvents, further mass spectrometric analyses of pure *PeHNL-N* and *PeHNL-P* glycan structures are required after digestion using endoglycosidases and endopeptidases.

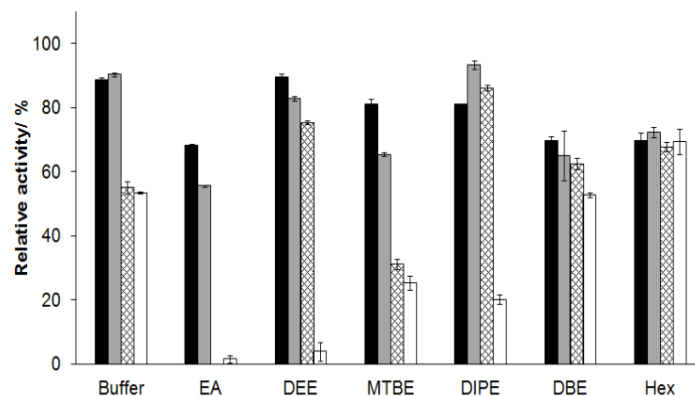


Figure 1-5. Effects of organic solvents on the stability of *PeHNLS* (2.34 U of each partially purified); After 12-h incubation at 10°C in each biphasic system at a citrate buffer/organic solvent ratio of 50:50, remaining enzyme activities of each *PeHNL* were measured using HPLC with a chiral column. Solid bar, *PeHNL-N*; Grey bar, *PeHNL-P*; Cross bar, *PeHNL-P-N105Q*; Open bar, *PeHNL-E*; EA, ethyl acetate; DEE, diethyl ether; MTBE, methyl-*t*-butyl ether; DIPE, 2-isopropyl ether; DBE, dibutyl ether; HEX, hexane. Data are presented as means \pm SD (n = 3).

Effect of biphasic systems on (a) enantiomeric excess of synthesized cyanohydrin and (b) transcyanation by native and recombinant PeHNLs

PeHNL-N was reportedly stable in biphasic systems of water and organic solvent, and catalyzed the synthesis of (*R*)-mandelonitrile from benzaldehyde and cyanide, with a 98.6% enantiomeric excess.¹⁹⁾ All of organic solvent used in this study have the positive values of the logarithm of the partition coefficients ($\log P$) which gave a high ee (88-99%) for the products, while the negative $\log P$ yielded the product cyanohydrins with low ee (2-30%).⁵⁵⁾ The biphasic system with ethyl acetate (EA; $\log P$ 0.67), diethyl ether (DEE; $\log P$ 0.85), methyl-*t*-butyl ether (MTBE; $\log P$ 1.4), 2-isopropyl ether (DIPE; $\log P$ 1.9), dibutyl ether (DBE; $\log P$ 2.9) and hexane (HEX; $\log P$ 3.5) have been studied on HNLs from *Hevea brasiliensis* (HbHNL),⁵⁶⁻⁵⁸⁾ *Manihot esculenta* (MeHNL),^{57, 58)} *Sorghum bicolor* (SbHNL),^{57, 58)} *Prunus amygdalus* (PaHNL)⁵⁹⁾ and *Eriobotrya japonica* (EjHNL).⁶⁰⁾ Thus, to further investigate the effects of glycosylation on enzyme activities, we determined enantiomeric excesses (*ee*) and transcyanation of (*R*)-mandelonitrile by PeHNLs in biphasic systems. PeHNLs in biphasic systems containing EA, DEE, MTBE, DIPE, or DBE and an equal volume of 400 mM citrate buffer pH 4.0 produced (*R*)-mandelonitrile from benzaldehyde and acetone cyanohydrin with ee of 94%, whereas enzymes from biphasic systems containing citrate buffer and Hex produced (*R*)-mandelonitrile with ee of 50%–35%. The lack of clear differences in ee between the other biphasic systems (Figure 1-6) suggest that glycosylation of PeHNL does not effect chiral specificity.

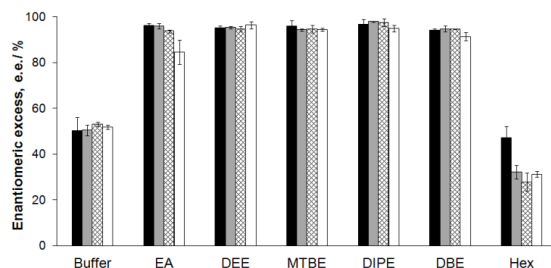


Figure 1-6. Effects of organic solvents on ee in (*R*)-mandelonitrile products at 12 h after the start of incubation with PeHNLs (2.34 U of each partially purified); (*R*)-mandelonitrile was synthesized using partially purified PeHNLs from benzaldehyde (250 mM) and acetone cyanohydrin (900 mM) in biphasic systems at a citrate buffer/organic solvent ratio of 50:50. Reactions were performed at 10°C to avoid decomposition of product. Solid bar, PeHNL-N; Gray bar, PeHNL-P; Cross bar, PeHNL-P-N105Q; Open bar, PeHNL-E; EA, ethyl acetate; DEE, diethyl ether; MTBE, methyl-*t*-butyl ether; DIPE, 2-isopropyl ether; DBE, dibutyl ether; HEX, hexane. Data are presented as means \pm SD (n = 3).

In further experiments, the relationship between cyanohydrin production and glycosylation of *PeHNL* was investigated by incubating substrates with each *PeHNL*, and periodically collecting and measuring quantities of (*R*)-mandelonitrile. Since an enantiomeric purity or ee of 98% of target chiral compound is minimum acceptable level for commercial process⁶¹, we expected that to achieve the highest ee of (*R*)-mandelonitrile and highest productivity of reaction. In most biphasic systems, (*R*)-mandelonitrile production increased with incubation times. Specifically, tested *PeHNL*s produced 6, 12, and 11 μM (*R*)-mandelonitrile after incubation for 12 h in 50 μL of buffer/EA, buffer/MTBE, and buffer/DBE biphasic systems, respectively (Figures 1-7A, C, and E). In 50- μL biphasic buffer/DEE and buffer/DIPE systems, *PeHNL*-P preferentially synthesized 5- and 15- μM (*R*)-mandelonitrile, respectively, after incubation for 12 h (Figures 1-7B and D), strongly indicating the influence of organic solvent on *PeHNL* activity. Moreover, MTBE was a suitable solvent for *PeHNL*-N, as indicated by its stability relative to other *PeHNL*s in this system, whereas HNL from *Prunus amygdalus* (*PaHNL*) had the lowest initial catalytic activity in the buffer/MTBE biphasic system.⁵⁹ DEE, DIPE, and DBE were reportedly suitable solvents that allowed high production of cyanohydrins by HNLs from *Eriobotrya japonica* (*EjHNL*)⁶⁰, *Arabidopsis thaliana* (*AtHNL*)¹⁶, and wild type *PeHNL*.¹⁹ The glycosylated enzyme can improve the stability of *PeHNL* when used in the biphasic system, however the productivity of (*R*)-mandelonitrile was depended on the type of organic solvent.^{56, 59} Hexane was not likely an appropriate organic solvent for the biphasic system under the present experimental conditions

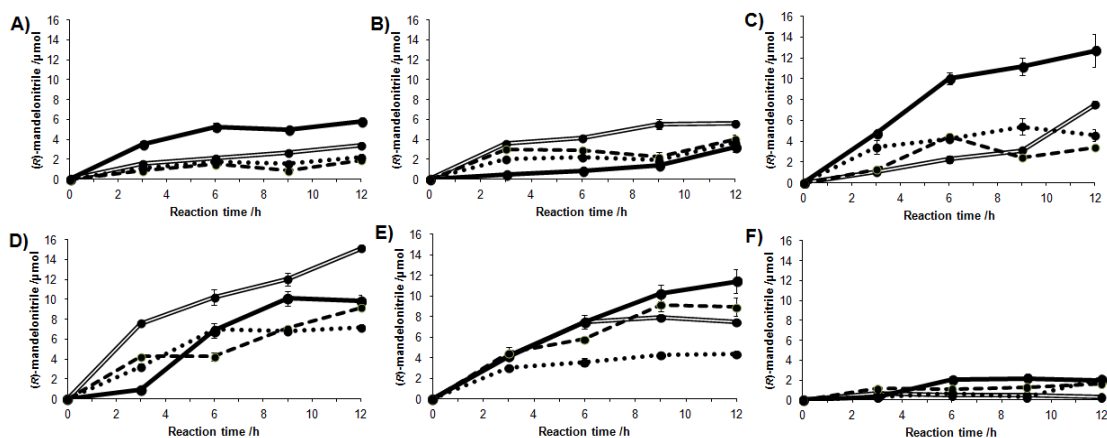


Figure 1-7. Time course of (*R*)-mandelonitrile synthesis following transcyanation reactions with *PeHNL*s in biphasic systems; reactions and analyses were performed as described in Figure 4. Products were collected at 3, 6, 9, and 12 h after the start of incubation; A), ethyl acetate; B), diethyl ether; C), methyl-*t*-butyl ether; D), 2-isopropyl ether; E), dibutyl ether; F), *n*-hexane; Solid line, *PeHNL*-N; double line, *PeHNL*-P; dashed line, *PeHNL*-P-N105Q; dotted line, *PeHNL*-E. Data are presented as means \pm SD (n = 3).

(Figure 1-7F). In our previous study, the production of cyanohydrin in a biphasic system was related to the partition coefficient for benzaldehyde between the buffer and the organic solvent. Non-enzymatic reactions are minimized by extraction and separation of benzaldehyde into the organic phase of biphasic systems with high partition coefficients, and the HCN remains in the aqueous phase and slowly releases acetone cyanohydrin.¹⁹⁾ In the present study, the Hex system gave the lowest partition value for benzaldehyde, likely accelerating the non-enzyme reaction significantly.⁶⁰⁾

Reusability of native and recombinant *PeHNLs*

The reusability among *PeHNLs* in synthesis of (*R*)-mandelonitrile using biphasic systems were compared. The suitable organic solvent for *PeHNL-N* is MTBE while that for *PeHNL-P*, *PeHNL-P-N105Q* and *PeHNL-E* are DIPE (Figure 1-5 and 1-7). The ee value of (*R*)-mandelonitrile was greater than 98.5% in each experimental condition (Figure 1-6). After four cycles of reusability experiment, the glycosylated *PeHNLs*, *PeHNL-N* and *PeHNL-P*, showed the production of (*R*)-mandelonitrile in the range of 30.8-45 μM and 26.7-31 μM , respectively (Figure 1-8A and D). In contrast, the non-glycosylated *PeHNLs*, *PeHNL-P-N105Q* and *PeHNL-E*, in the first cycle of reusability experiment showed the production of 20 μM and 15 μM of (*R*)-mandelonitrile, respectively, which were lower than that by the glycosylated *PeHNLs*. After the two cycles, the

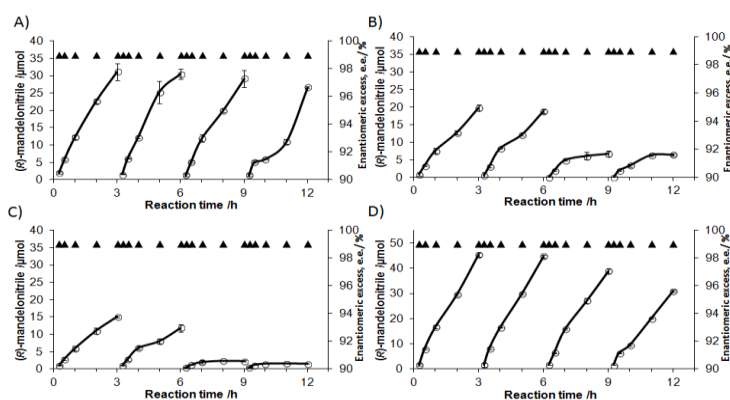


Figure 1-8. The reusability of *PeHNL-P* (A), *PeHNL-P-N105Q* (B), *PeHNL-E* (C) and *PeHNL-N* (D) in synthesis of (*R*)-mandelonitrile from benzaldehyde and acetone cyanohydrin. \circ , amount of (*R*)-mandelonitrile. \blacktriangle , enantiomeric excess, %. The reaction were performed in biphasic system of buffer (pH 4.0; 50% v/v) and DIPE or MTBE (50% v/v) at 10°C containing benzaldehyde (250 mM), acetone cyanohydrin (900 mM), and enzyme (5 U).

activity of non-glycosylated *PeHNLs* was significantly decreased in biphasic system of DIPE (Figure 1-8B and C). These results is likely caused by the organic solvent stability of *PeHNL-P* and *PeHNL-N*, which showed the remaining activity of 90% and 80% in biphasic systems of DIPE and MTBE for 12 h incubation at 10°C, respectively (Figure 1-5.), while *PeHNL-P-N105Q* and *PeHNL-E* lost 15% and 80% of the activity in biphasic systems of DIPE at the same condition, respectively.

CHAPTER II: Stabilization of hydroxynitrile lyases from two variants of passion fruits, *Passiflora edulis* Sims and *Passiflora edulis* forma *flavicarpa* by C-terminal truncation

Because the synthesis of chiral compounds generally requires a broad range of substrate specificity and stable enzymes, screening for better enzymes and/or improvement of enzyme properties through molecular approaches is necessary for sustainable industrial development. Herein, the discovery of unique hydroxynitrile lyases (HNLs) from two species of passion fruits, *Passiflora edulis* forma *flavicarpa* (yellow passion fruit, *PeHNL*-Ny) and *Passiflora edulis* Sims (purple passion fruit, *PeHNL*-Np), isolated and purified from passion fruit leaves is reported. These are the smallest HNLs (comprising 121 amino acids). Amino acid sequences of both enzymes are 99% identical; there is a difference of one amino acid in a consensus sequence. *PeHNL*-Np has an Ala residue at position 107 and is nonglycosylated at Asn105. Because it was confirmed that natural and glycosylated *PeHNL*-Ny showed superior thermostability, pH stability, and organic tolerance to that of *PeHNL*-Np, it has been speculated that protein engineering around the only glycosylation site, Asn105, located at the C-terminal region of *PeHNL*-Ny, might contribute to the stabilization of *PeHNL*. Therefore, the focus is on improved stability of the nonglycosylated *PeHNL* by truncating its C-terminal region. The C-terminal-truncated *PeHNL*D107 was obtained by truncating 15 amino acids from the C terminus followed by expression in *E. coli*. *PeHNL*D107 expressed in *E. coli* was not glycosylated, and showed improved thermostability, solvent stability, and reusability similar to that of the wild-type glycosylated form of *PeHNL* expressed in *P. pastoris*. These data reveal that the lack of the high-flexibility region at the C terminus of *PeHNL* might be a possible reason for improving the stability of *PeHNL*.

MATERIALS AND METHODS

Materials and chemicals

Leaves of *Passiflora edulis* Sims and *Passiflora edulis* forma *flavicarpa* were collected from the Thung Roeng Royal Project Development Center in Chiang Mai, Thailand, and the Botanic Garden of Toyama, Japan, respectively, and stored at -20 °C. Benzaldehyde (redistilled,

99.5 %) and (*rac*)-mandelonitrile were purchased from Sigma-Aldrich. Other chemicals used in the experiments were purchased from chemical sources and used without further purification.

Assays to determine the synthetic activity of HNL

Enzymatic activity was quantified by determining (*R*)-mandelonitrile synthesis using benzaldehyde and KCN as described in the chapter I.

Purification of PeHNLS from leaves

The PeHNL from *P. edulis* forma *flavicarpa* was purified as described previously²⁰. Here we describe a procedure to purify PeHNL-Np. Briefly, *P. edulis* Sims leaves were frozen using liquid nitrogen and ground with mortar and pestle. Extraction and protein fractionation with ammonium sulfate was performed in the same manner as previously reported.²⁰ After dialysis of the active ammonium sulfate fraction, proteins from the extract were loaded onto the Toyopearl DEAE -650M column (Tosoh, 50 mm i.d., column volume 150 mL) and eluted with a linear gradient of sodium chloride from 0 to 500 mM in potassium phosphate buffer (KPB, pH 6.0, 20 mM). The active fractions were added ammonium sulfate (30 % saturated) and loaded onto a Butyl-Toyopearl column (Tosoh, 25 mm i.d., column volume 50 mL). Bound proteins were eluted stepwise with ammonium sulfate (1330, 860, 420, and 0 mM) in KPB (pH 6.0, 20 mM). Active fractions were combined, dialyzed thrice against KPB (pH 6.0, 20 mM) and then loaded onto the 2nd Toyopearl DEAE-650M column (25 mm i.d., column volume 20 mL). Bound protein was pre-washed with 50 mM sodium chloride in KPB (pH 6.0, 20 mM) and eluted with a linear gradient of sodium chloride from 50 to 100 mM in the same buffer. Active fractions were pooled, concentrated, desalted by using a centrifugal filtration device (Amicon Ultra-15, 10000 NMWL, EMD Millipore) and then loaded onto a Mono Q 5/50 GL column (1 mL, GE Healthcare). Proteins were then eluted with a linear gradient of sodium chloride from 0 to 100 mM in KPB (pH 6.0, 20 mM) at a flow rate of 0.8 mL.min⁻¹. Active fractions were pooled, concentrated and further purified using the Superdex 200 10/300 GL column (GE Healthcare). Purified PeHNL-Np was monitored using sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE)³¹ and glycosylation was detected using Periodic acid-Schiff (PAS)⁶² staining using a GelCode Glycoprotein Staining Kit (Thermo Fisher Scientific) according to the manufacturer's instructions.

Cloning of PeHNL-Np cDNA

Total RNA was isolated from a young leaf of *P. edulis* Sims after 3 months of cultivation and reverse transcribed. As purified PeHNL-Np showed a molecular mass of 14000 Da which is almost the same as PeHNL-Ny, we first tried to identify the PeHNL-Np gene from *P. edulis* Sims using degenerate and gene-specific primers as chapter I. DNA sequences were determined using a 3500 Genetic Analyzer (Thermo Fisher Scientific) and assembled and analyzed using ATGC and Genetyx Ver. 12 (Genetyx, Tokyo, Japan), respectively.

B-factor and root mean squared deviation analysis

The B-factor of wild-type PeHNL (PDB ID:5XZQ) was extracted from the pdb structure using the B-FITTER software.⁶³⁾ The root mean squared deviation (RMSD) of the similarity between the structures of the two PeHNLS, wild-type PeHNL and PeHNL Δ ₁₀₇ (PDB ID: 5XZT) have been calculated by the Molecular Operating Environment (MOE, Chemical Computing Group Inc Canada).

Bacterial strain and construction of the C-terminal truncated PeHNL

E. coli JM 109 (Takara Bio) and BL21 (DE3) (Thermo Fisher Scientific) cells were used as hosts for amplification of plasmid DNAs and recombinant protein expression, respectively. PeHNL from *P. edulis* forma *flavicarpa* was cloned and expressed using the pColdI expression vector (pColdI-PeHNL) as chapter I. Wild-type pColdI-PeHNL was truncated at the C-terminus of each variant using QuikChange® Lightning Mutagenesis Kit (Agilent Technologies, USA) and the truncation primers shown in Appendix II. PCR products were treated with 10 U of *DpnI* at 37 °C for 1 h, and then used for transformation of *E. coli* JM 109 using the heat shock method. Sequences were confirmed using the 3500 Genetic Analyzer.

Gene expression, purification, and determination of molecular mass of wild-type and C-terminal truncated PeHNL

E. coli BL21(DE3) cells harboring recombinant plasmids were cultured in Luria–Bertani (LB, 5 mL) broth containing ampicillin (50 μ g ml⁻¹) and incubated overnight at 37 °C with shaking at 300 rpm. The pre-culture starter was transferred into LB broth (500 mL) containing ampicillin (50 μ g ml⁻¹) in an Erlenmeyer flask (2 L) and cultured at 37 °C with shaking at 150 rpm until *A*₆₀₀ reached 0.6, followed by a cold-shock at 16 °C. After 2 h, isopropyl- β -D-thiogalactopyranoside (IPTG) was added to a final concentration of 1.0 mM, and cells were cultured

at 16 °C for 24 h with the same rate of shaking. Protein extraction and purification were performed exactly as described previously.²⁰ Superdex™ 10/300 GL (GE Healthcare) was used to calculate the molecular mass of the purified recombinant wild-type *PeHNL* and its truncated form based on their relative mobility as compared with values for standard marker proteins as follows: glutamate dehydrogenase (290,000 Da), lactate dehydrogenase (142,000 Da), enolase (67,000 Da), adenylatekinase (32,000 Da), and cytochrome *c* (12,400 Da).

Measurement of kinetic parameters

Kinetic parameters of purified wild-type and C-terminal truncation mutants of *PeHNL* were assayed in sodium citrate buffer (pH 4.0, 400 mM), according to the method described above, at various substrate concentrations.

Effect of pH on enzyme activity and stability

Two natural purified *PeHNL*s were used to determine optimum pH for (*R*)-mandelonitrile synthesis. The reaction was performed at 25 °C between pH 2.5 to pH 6.0 (400 mM) for 5 min. To determine enzyme stability, each *PeHNL* was preincubated at 30 °C between the range of pH 3.5 to pH 10.0 (40 mM) for 1 h. The remaining activity of the enzyme was measured as described above.

Effect of temperature on enzyme activity and stability

Optimum temperature required for enzyme activity was examined by incubating the reaction mixture at temperatures ranging from 5 °C to 45 °C in citrate buffer (pH 4.0, 400 mM) for 5 min. For determination of thermostability, the enzyme was preincubated in potassium phosphate buffer (pH 6.0, 20 mM) over the range of 30 to 80 °C for 1 h. The remaining activity of the enzyme was measured as described above.

Effect of organic solvent on stability and cyanohydrin synthesis in a biphasic system

Organic solvent stability of two natural *PeHNL*s were compared using different organic tests as described in detail in chapter I. In the study on cyanohydrin synthesis of C-terminal truncation *PeHNL* mutants, diisopropyl ether (DIPE) was selected for synthesis of cyanohydrins in a biphasic system, organic solvent phases containing various aldehyde (250 mM) and mixed with citrate buffer (pH 3.5, 400 mM) containing each purified C-terminal truncated *PeHNL* mutants (10 U) in total volumes of 1.0 mL in 2.0 mL microcentrifuge tubes. Reactions were initiated by adding

potassium cyanide (600 mM), and the mixtures were incubated at 10 °C with shaking at 1500 rpm in an incubator shaker (BioShaker M-BR-022UP, Taitec Corporation, Tokyo, Japan). Aliquots of samples (10 µL) were collected from the organic phase at various time-points and analyzed using HPLC at condition as described in detail by Dadashipour *et al.*³⁾ Alternatively, for chiral high performance liquid chromatography (HPLC) analysis of the conversion of 2-chlorobenzaldehyde, 3-chlorobenzaldehyde and 4-chlorobenzaldehyde were performed with a CHIRALCEL OD-H column (Diacel, Osaka, Japan) under the following condition: mobile phase n-hexane/propan-2-ol/trifluoroacetic acid with a volume ratio of 96:4:0.2, flow rate 1 mL min⁻¹, absorbance 228 nm, column oven temperature 25 °C. *ee* values were determined according to relative concentration (mM) of the two enantiomers using Equation (1) as the same as chapter I.

Reusability

The reusability of recombinant purified *PeHNL*, wild-type (*PeHNL*_{121y}), and C-terminal truncated *PeHNL* Δ ₁₀₇ expressed in *E. coli* was compared in the cyanohydrin synthesis reaction batch containing benzaldehyde (250 mM) and KCN in biphasic systems of citrate buffer (pH 3.5, 400 mM, 0.5 mL) and DIPE (0.5 mL) in a microcentrifuge tube (2 mL). A reaction was also started by addition of each *PeHNL* (5 U) the same as described above. After 3 h of incubation, the buffer phase containing the enzyme was recovered and dialyzed against citrate buffer (pH 3.5, 400 mM) using a centrifugal filtration device. The exchanged buffer containing *PeHNL* was then reused for the next batch of reactions under the same conditions. The amount of (*R*)-mandelonitrile produced was estimated as described above.

RESULTS AND DISCUSSION

PeHNL is present in two species of passion fruits

The HNL from *P. edulis* f. *flavicarpa* (yellow passion fruit, *PeHNL*-Ny) has been discovered, characterized, and gene coding for it was successfully cloned and expressed to produce the recombinant *PeHNL*-Ny.¹⁸⁻²⁰⁾ Recently, we solved the *PeHNL*-Ny structure and proposed its catalytic mechanism.⁶⁴⁾ Here, we report the discovery of HNL from another commercial passion fruit species, *P. edulis* SIM (purple passion fruit, *PeHNL*-Np), which is cultivated on a large scale

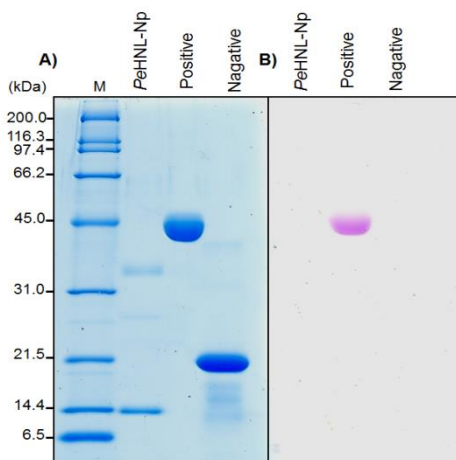


Figure 2-1. A) SDS-PAGE analysis, and B) Periodic acid-Schiff (PAS) stained of native purple *PeHNL* (*PeHNL*-Np). Positive control: horseradish peroxidase (HRP), negative control: soybean trypsin inhibitor

in Thailand. A large amount of *P. edulis* SIM leaves (5000 g) was collected at a plantation in Chaing-Mai, Thailand.

The HNL activity from the leaf extract was determined by (*R*)-mandelonitrile ((*R*)-MAN) synthesis using HPLC equipped with a chiral column. *PeHNL*-Np levels (2.42 U/g leaf sample) detected in leaves were approximately 3.3-fold less than in *P. edulis* f. *flavicarpa* (8.13 U/ g leaf sample), possibly due to differences in origin and maturity of plants. *PeHNL*-Np was also purified to near homogeneity using several rounds of column chromatography as summarized in Table 2-1. The specific activity for the synthesis of (*R*)-mandelonitrile from benzaldehyde and potassium cyanide (KCN) was 120

U/mg, which was marginally lower than that from *PeHNL*-Ny (136 U/mg). *PeHNL*-Np has a molecular mass of 14 kDa, same as *PeHNL*-Ny, as revealed by SDS-PAGE; moreover, they have 99% identity. At position 107, *PeHNL*-Np from purple passion fruit has an Ala residue, whereas that from yellow passion fruit has a Thr residue (Appendix III). Therefore, the purple passion fruit *PeHNL* cannot be glycosylated on the nitrogen atom of the Asn residue of nascent polypeptide in the recognition tripeptide sequence Asn-X-Ser/Thr.⁶⁵ Periodic acid Schiff (PAS) staining of the purified *PeHNL*-Np confirmed that it is not a glycoprotein (Figure 2-1). Using various approaches such as the use of *N*-glycosylation and *N*-glycan-processing inhibitors and site-directed mutagenesis of *N*-glycosylation site, glycans in plant glycoproteins have long been shown to play a major role in the folding, processing, and secretion of proteins from the endoplasmic reticulum (ER) and the Golgi apparatus.⁶⁶ For example, treatment of suspension-cultured carrot (*Daucus carota*) cells with tunicamycin, an antibiotic which inhibits Asn-linked glycosylation, yielded unglycosylated carrot cell-wall β -fructosidase, and the unglycosylated enzyme was easily degraded during the last stages in the secretory pathway or immediately after its arrival in the cell wall.⁶⁷ The *N*-glycan is also known to protect the protein from proteolytic degradation and is responsible for thermostability, solubility, and biological activity of concanavalin A (Con A), the lectin from jack bean seeds.⁶⁸

The amount of glycosylated *PeHNL*-Ny in the whole leaves of the yellow passion fruit was 3.3-fold higher per mg of the total leaf protein than that of non-glycosylated *PeHNL*-Np in the leaves of purple passion fruits. It seems likely that the presence of glycan affected the transport of

PeHNL in the yellow passion fruit cells. The yellow passion fruit grows faster and has a greater resistance to soil fungi, whereas the purple one is more resistant to cold injury.⁶⁹⁾ Thus, yellow and purple passion fruits may have been adopted to their environment and evolved to control the amounts of HNLs glycosylation, according to their survival strategies.

Table 2-1. Purification summary of two natural *PeHNLs*

Purification step	Total activity	Specific activity (U/mg)	Recovery (%)	Purification fold
<i>PeHNL-Ny</i>				
Crude (1500 g)	12200	9.9	100	1.0
20-40% (NH ₄) ₂ SO ₄	5320	11.9	43.6	1.2
DEAE Toyopearl	4220	12.2	34.5	1.2
Butyl Toyopearl	2560	16.8	32.3	1.7
MonoQ HR10/100	360	39	4.55	3.9
Resource PHE	42.2	52.4	0.53	5.3
Hydroxapatite type I	14.7	58.7	0.18	5.9
Concanavalin A Sepharose	3.5	136	0.04	13.4
<i>PeHNL-Np</i>				
Crude (5000 g)	12100	1.56	100	1.0
20-40% (NH ₄) ₂ SO ₄	7560	1.98	62.5	1.3
DEAE Toyopearl 1 st	6200	6.45	51.2	4.1
Butyl Toyopearl	2510	7.17	20.7	4.6
DEAE Toyopearl 2 nd	1880	9.50	15.5	6.1
MonoQ HR10/100	350	97.3	2.9	62.3
Superdex 75 10/300 GL	15.2	120	0.12	77.0

Biochemical properties of two PeHNLs

To compare enzyme kinetics of two natural *PeHNLs*, the rate of synthesis (*R*)-mandelonitrile from benzaldehyde and KCN was measured using HPLC-chiral column analyses. The V_{\max} , k_{cat} , and k_{cat}/K_m values of glycosylated *PeHNL-Ny* were 220 $\mu\text{mol min}^{-1} \text{mg}^{-1}$, 50.4 s^{-1} , and 3.70 $\text{mM}^{-1} \cdot \text{s}^{-1}$, respectively, and these values were higher than those of non-glycosylated *PeHNL-Np* (150 $\mu\text{mol min}^{-1} \text{mg}^{-1}$, 35.7 s^{-1} , and 1.86 $\text{mM}^{-1} \cdot \text{s}^{-1}$, respectively,) (Table 2-2.). However, the K_m of *PeHNL-Ny* (13.7 mM) was slightly lower than that of *PeHNL-Np* (19.2 mM), which corresponds with our previous report that the glycan on *PeHNL* enhances the catalytic efficiency without reducing the substrate affinity for benzaldehyde and KCN.²⁰⁾ Also, we proved that those

Table 2-2. Kinetic parameters of native and C-truncated mutant *Pe*HNLs.

Enzyme	Spec.act. (U mg ⁻¹)	V _{max} (μmol min ⁻¹ mg ⁻¹)	K _m (mM)	k _{cat} (s ⁻¹)	k _{cat} /K _m (mM ⁻¹ .s ⁻¹)
<i>Pe</i> HNL-Ny	136	220±0.50	13.7±0.01	50.4±0.01	3.70±0.01
<i>Pe</i> HNL-Np	120	150±3.20	19.2±0.03	35.7±0.04	1.86±0.01
<i>Pe</i> HNL _{121p}	110	145±3.27	25.2±0.46	33.6±0.50	1.33±0.01
<i>Pe</i> HNL _{121y}	122	168±2.37	26.3±0.14	36.4±0.50	1.40±0.32
<i>Pe</i> HNLΔ ₁₁₇	134	228±0.67	23.4±0.14	53.4±0.16	2.30±0.52
<i>Pe</i> HNLΔ ₁₁₂	167	270±0.80	25.0±0.16	63.1±0.20	2.53±0.30
<i>Pe</i> HNLΔ ₁₀₉	118	167±0.37	26.1±0.11	39.0±0.10	1.50±0.05
<i>Pe</i> HNLΔ ₁₀₇	112	158±0.40	26.6±0.13	37.2±0.10	1.40±0.06
<i>Pe</i> HNLΔ ₁₀₅	60	103±0.20	33.0±0.12	24.1±0.04	0.73±0.01
<i>Pe</i> HNLΔ ₁₀₂	n.d.	n.d.	n.d.	n.d.	n.d.

n.d. = not determined

kinetic parameters of natural *Pe*HNL were caused by the presence of glycan. We determined the kinetic parameters using both purified recombinant *Pe*HNLs lacking glycosylation expressed in *E. coli*. The K_m, V_{max}, k_{cat}, and k_{cat}/K_m values of *Pe*HNL_{121y} were estimated to be 26.3 mM, 168 μmol min⁻¹ mg⁻¹, 36.4 s⁻¹, and 1.85 mM⁻¹.s⁻¹, respectively, and these values were similar to those of *Pe*HNL_{121p} (25.2 mM, 145 μmol min⁻¹ mg⁻¹, 33.6 s⁻¹, and 1.33 mM⁻¹.s⁻¹, respectively) (Table 1). Thus, the difference of one amino acid at position 107 of *Pe*HNL appears not to affect the affinity and catalytic efficiency of *Pe*HNL.

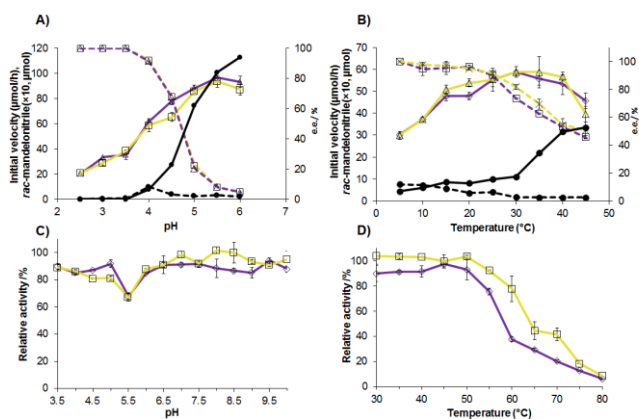


Figure 2-2. The comparison of biochemical properties of two *Pe*HNLs; *Pe*HNL-Ny (squares) and *Pe*HNL-Np (triangles). A) Effect of pH and B) temperature on initial velocity (solid line) and enantiomeric excess (dashed line). C) pH stability and D) temperature stability after incubation of enzymes at various pH or temperatures for 1 h. Remaining activity was estimated by monitoring the synthesis of (*R*)-mandelonitrile from benzaldehyde and potassium cyanide using HPLC with a chiral column. Data are presented as mean ± SD (n = 3)

The effect of pH and temperature on the synthesis (*R*)-mandelonitrile was studied using the purified natural *Pe*HNLs. *Pe*HNL-Ny and *Pe*HNL-Np exhibited the highest (*R*)-MAN synthetic activity at pH 4.0 with >95% *ee* and no stereoselectivity above pH 6.0 (Figure 2-2A). The *ee* of 99% (*R*)-MAN was obtained at pH 2.5-3.5 for both *Pe*HNLs. The pH stability analyses revealed that both natural *Pe*HNLs were broadly stable over the range of 3.5 to 10 after preincubation at 30°C for 1 h without

substrate (Figure 2-2C). The optimum temperature range of both *PeHNL*s for the (*R*)-MAN synthetic activity was 20 to 25 °C (Figure 2-2B). The *ee* of the product by the two natural *PeHNL*s were more than 95% at low temperatures between 5°C to 20°C, and low *ee* products were obtained when the temperature was increased, due to the acceleration of the non-enzymatic reaction. In addition, the activity of *PeHNL*-Ny remained above 80% after 1 h incubation at 60°C (Figure. 2-2D), while *PeHNL*-Np retained only 40% of activity under the same condition. Obviously, the thermostability was significantly improved by glycosylation in *PeHNL*-Ny. Furthermore, glycosylated *PeHNL*-Ny was more stable in the presence of an organic solvent in a biphasic system than non-glycosylated *PeHNL*-Np toward all the organic solvents tested (Appendix IV). This evidence supports that the glycosylated *PeHNL* is more stable than non-glycosylated *PeHNL*.²⁰⁾ Recently, we solved the structure of *PeHNL* and discussed that a glycosylation site on Asn105 of *PeHNL* might stabilize the high flexibility region at C-terminus, which could improve *PeHNL* stability.⁶⁴⁾ Thus, we hypothesized that reduction in the high flexibility region of *PeHNL* by the deletion of part of C-terminal region increases the stability of non-glycosylated *PeHNL*.

Identification of flexible regions in PeHNL

The secondary structure of wild-type *PeHNL* (PDB ID:5XZQ) consists mainly of four antiparallel β -sheets (β 1, 6-14 aa; β 2, 44-47 aa; β 3, 62-68 aa; β 4, 94-103 aa), three α -helices (α 1, 20-36 aa; α 2, 71-79 aa; α 3, 81-93 aa), and a special loop at C-terminal that mainly contains hydrophobic amino acid residues (₁₀₄INETFPYTWALPNKYVVT₁₂₁).⁶⁴⁾ During structural analysis, we missed the C-terminus region (residues 110-121) because of the poor electron density, indicating that this region has high flexibility and does not form a stable structure.⁶⁴⁾ Generally, the B-factor, which can evaluate the flexibility of an individual residue in a protein, is used for explaining the behavior of atomic electron densities from X-ray data.⁶³⁾ The available C-terminal region of wild-type *PeHNL* crystal structure showed remarkably high B-factor values at the residues E106, T107, F108, and P109, which were 90, 70.5, 72.6, and 82.5, respectively (Figure 2-3A). In agreement with the root mean squared deviation (RMSD) of the similarity between two superimposed atomic coordinates of the *PeHNL* wild-type and *PeHNL* Δ ₁₀₇ (PDB ID: 5XZT)⁶⁴⁾ structure at C-terminal region that also showed a high RMSD, with an average value above 1.281 Å (Figure 2-3B). The

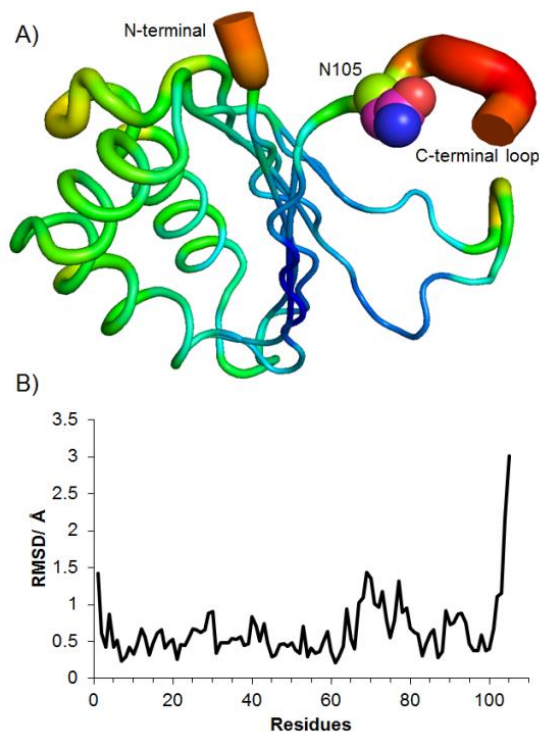


Figure 2-3. Flexibility in *PeHNL*. A) A B-factor putty drawing of wild-type *PeHNL* (PDB ID:5XZQ) structure in the nucleocapsid-like particle. Regions colored red also have a large diameter, which corresponds to higher B factors and are associated with higher structural flexibility. A consensus sequence (N105) for *N*-glycosylation site on *PeHNL* is shown as cyan spheres. B) Root mean square deviation (RMSD) between the crystal structure of wild-type and *PeHNL* Δ_{107} (PDB ID: 5XZT).

RMSD value of this region is much higher than that of other regions, indicating that the C-terminus region is highly flexible, which might be responsible for the instability of *PeHNL*.

Expression and purification of C-terminal truncation mutants

To further study the roles of C-terminal loop, truncation mutants were constructed lacking the chosen amino acid residues from this region. Five to 18 residues (4.1 to 14% of the total *PeHNL* length) were deleted from the C-terminus. Up to 17 amino acid residues of *PeHNL* could be removed (*PeHNL* Δ_{105}) without abolishing its activity. A 6xHis tag was attached to the N-terminus of truncation mutants and they were heterologously expressed in *E. coli* BL21 (DE3) cells. The yields of all C-terminal truncation mutants were similar, except *PeHNL* Δ_{105} , which showed the lowest *PeHNL* production. The truncation mutants were purified using Ni-

spharose 6 Fast Flow and a Mono Q 5/50 column. The purified *PeHNL* Δ_{117} , *PeHNL* Δ_{112} , *PeHNL* Δ_{109} , *PeHNL* Δ_{107} , and *PeHNL* Δ_{105} had molecular masses of approximately 14.9, 14.4, 14.3, 14.2, and 14.2 kDa, respectively, in SDS-PAGE analyses (Appendix V). The molecular masses of these mutants as determined by high-performance gel chromatography on Superdex 200 10/30 were 27.6, 26.8, 26.8, 26.6, and 26.4 kDa, respectively, indicating that the native enzyme was active as a homodimer.

Analysis of specific activity and kinetic parameters of C-terminal truncation mutants

The C-terminal truncation mutants affected the mandelonitrile synthesis activity of *PeHNL* as shown in Table 2-2. The specific activity of *PeHNL* Δ_{117} and *PeHNL* Δ_{112} were 134 U mg⁻¹ and 167 U mg⁻¹, respectively, which were marginally higher than that of wild-type *PeHNL*₁₂₁. In

contrast, the specific activity of *PeHNL* Δ_{105} was 60 U mg⁻¹, significantly (2-fold) lower than that of wild-type *PeHNL* Δ_{121} ; *PeHNL* Δ_{102} activity could not be detected. These results indicated that the 15 amino acid deletion from C-terminus did not affect *PeHNL* activity. Next, we determined kinetic parameters of the C-terminal truncation mutants for the synthesis of mandelonitrile from benzaldehyde and KCN. The V_{\max} , k_{cat} , and k_{cat}/K_m of *PeHNL* Δ_{117} and *PeHNL* Δ_{112} were 228±0.67 and 270±0.80 $\mu\text{mol min}^{-1} \text{mg}^{-1}$, 53.4±0.16 and 63.1±0.20 s⁻¹, 2.30±0.52 and 2.53±0.30 mM⁻¹.s⁻¹, respectively, while the V_{\max} , k_{cat} , and k_{cat}/K_m of other C-terminal truncation mutants were similar to those of the wild-type, except for *PeHNL* Δ_{105} , which had the lowest values. The K_m values of wild-type, *PeHNL* Δ_{117} , *PeHNL* Δ_{112} , *PeHNL* Δ_{109} , and *PeHNL* Δ_{107} were 26.3±0.14, 23.4±0.14, 25.0±0.16, 26.1±0.11, and 26.6±0.13 mM, respectively (Table 2-2). Thus, truncation of 5-15 amino acid residues from C terminal did not affect the affinity of *PeHNL* for benzaldehyde.

C-terminal truncation mutant *PeHNL* Δ_{107} improved *PeHNL* stability

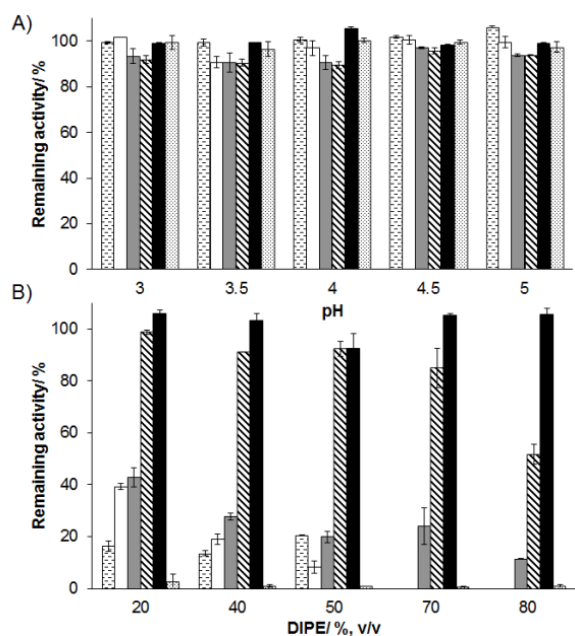


Figure 2-4. The stability profile of wild-type and C-terminal truncated *PeHNL* variants. A) Stability at low pH. After incubation in reaction mixtures with varying pH values for 1 h, the remaining activity was measured as described in Figure 1. B) Solvent stability, the remaining activities of *PeHNL* mutants were measured after incubation in each biphasic system in a citrate buffer (200 μL)/organic solvent (various solvent phase content) at 10 $^{\circ}\text{C}$ for 12 h using HPLC with a chiral column. Dashed bar: *PeHNL* Δ_{121y} . Open bar: *PeHNL* Δ_{117} . Gray bar: *PeHNL* Δ_{112} . Cross bar: *PeHNL* Δ_{109} . Solid bar: *PeHNL* Δ_{107} . Dot bar: *PeHNL* Δ_{105} . Data are presented as means \pm SD (n = 3)

To prove that the flexible region at C-terminus of *PeHNL* contributes to enzyme stability, we further investigated the resistance of the wild-type enzyme and truncation mutants to low pH, temperature, and organic solvent. A low pH is necessary for the enantioselective synthesis of cyanohydrins because of the spontaneous unselective addition of HCN to the carbonyl group occurs as a background reaction at elevated pH.^{11, 56, 70, 71)} We analyzed the stability of C-terminal truncated mutants of *PeHNL* at low pH (3.0 to 5.0). All *PeHNL*s were stable over the pH range tested (Figure 2-4A).

Although stability under most pH condition was comparable, a marginal difference in the stability was observed among the *PeHNL*s. Wild-type, *PeHNL* Δ_{117} , and

PeHNL Δ_{107} showed > 95% remaining activity at low pH, which was slightly higher than the remaining activities of *PeHNL* Δ_{112} , and *PeHNL* Δ_{109} (90%) under same pH conditions for 1 h incubation. It appears that the deletion of 5-17 amino acid residues from C-terminus does not increase the stability of *PeHNL* under the pH range tested.

Thermostability analyses revealed that the partial deletion of 5 to 13 amino acid (*PeHNL* Δ_{117} , *PeHNL* Δ_{112} , and *PeHNL* Δ_{109}) from C-terminus decreased the thermostability at 45°C compared with that of wild-type for 1 h preincubation. Surprisingly, only *PeHNL* Δ_{107} exhibited >80% and >60% remaining activity at 50°C and 55°C, respectively, which was highest for 1 h preincubation (Figure 2-5A). *PeHNL* Δ_{107} also showed >40% remaining activity after 2 h incubation at 50°C, in contrast to the wild-type and other C-terminal truncated variants that were rapidly inactivated (Figure 2-5B). These results suggest that the deletion of high flexibility region of 15 amino acid at C-terminal likely contributes to the thermostability of *PeHNL*.

Previously, we found that diisopropyl ether (DIPE) is a suitable organic solvent for cyanohydrin synthesis by non-glycosylated *PeHNL* expressed in *E. coli* when used in a biphasic system; however, the wild-type non-glycosylated *PeHNL* expressed in *E. coli* has low stability in an organic solvent.²⁰ In this study, we compared the solvent stability of the wild-type and C-terminus truncated mutants of *PeHNL* in a biphasic system of buffer and different concentration of DIPE. The stabilities of the variants were monitored after incubation at 10 °C for 12 h with shaking at 1,500 rpm and the residual activity of enzyme were compared with the initial activity in the aqueous

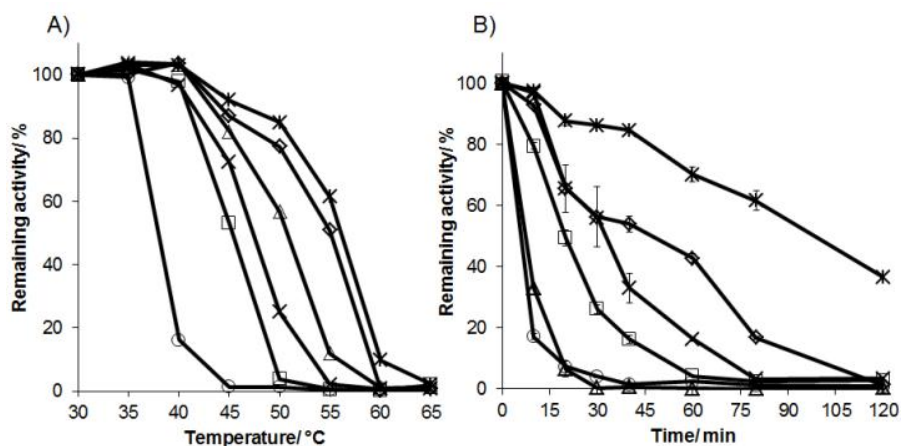


Figure 2-5. A) Thermostability at different temperatures, B) thermostability at 50 °C of C-terminal truncated variants of *PeHNL* Δ_{121y} (◇), *PeHNL* Δ_{117} (□), *PeHNL* Δ_{112} (Δ), *PeHNL* Δ_{109} (×), *PeHNL* Δ_{107} (⋈) and *PeHNL* Δ_{105} (○). Activity of the enzyme was detected by incubating the enzyme for 1 h at different temperatures and then the remaining activity was measured as described in Figure 1. Data are presented as means \pm SD (n = 3)

phase saturated with organic solvents before incubation. The highest solvent stability was shown by *PeHNL* Δ ₁₀₇ in all concentrations of DIPE (20–80%), with more than 90% residual activity. On the other hand, activity of the wild-type, *PeHNL* Δ ₁₁₇, *PeHNL* Δ ₁₁₂, and *PeHNL* Δ ₁₀₅ enzymes reduced to 90% to 50% as the DIPE concentration increased from 20% to 50% v/v, and these enzymes were almost inactivate at a high DIPE concentration of 70% to 80% v/v (Figure 2-4B). These results demonstrated that deleting 15 amino acid from the C-terminus (*PeHNL* Δ ₁₀₇) improved the thermostability and solvent stability of *PeHNL*.

PeHNL Δ ₁₀₇ catalyzes the synthesis of cyanohydrins

We are tried to improve the stability of non-glycosylated *PeHNL* expressed in *E. coli* by truncation of its C-terminus, it showed potential for cyanohydrin synthesis in a biphasic system with an organic solvent. It has been reported that biphasic systems could improve the enantiomeric excess of the cyanohydrin product and minimized the non-enzymatic cyanation reaction upon extraction and separation of an aldehyde into the organic phase in a biphasic system.^{19, 20, 72, 73} In this study, we synthesized cyanohydrins in the presence of either the wild-type form *PeHNL* or *PeHNL* Δ ₁₀₇ (10 U) in biphasic systems containing several aldehydes at high concentration (250 mM) in DIPE (0.75 mL) and an equal volume of KCN (500 mM) in citrate buffer (400 mM, pH 3.5) used as a substrate. *PeHNL* showed a preference for several aromatic aldehydes other than bulky aromatic aldehydes such as 3-methoxybenzaldehyde (1f), 1-naphthaldehyde (1i) and 2-naphthaldehyde (1j) which proved to be poor substrates for *PeHNL* (Table 2-3). The maximum conversion and more than 99% *ee* were obtained in less than 30 min for the synthesis of (2*S*)-thiophen-2-yl-hydroxyacetonitrile (2g). Compared to the wild-type *PeHNL*_{121y}, *PeHNL* Δ ₁₀₇ showed good *ee* (>90%) and the over 90% conversion after 3 h, which are higher than that of the wild-type *PeHNL*_{121y} for the synthesis of (*R*)-mandelonitrile (2a), (*R*)-2-methylmandelonitrile (2b), (*R*)-2-chloromandolonitrile (2d) and (*R*)-3-chloromandolonitrile (2e). The substrate specificity toward aldehyde or substituted benzaldehydes for *PeHNL* is the major intermediate cyanohydrins for the preparation of bioactive drugs such as duloxetine (2g, used for treatment of depression and general anxiety disorder),⁷⁴ semi-synthetic cephalosporins (2a, antibacterial agents),⁷⁵ and clopidogrel (2d, anticoagulant).⁷⁶ Thus, *PeHNL* is a good potential biocatalyst that can be used for industrial application.

Table 2-3. Conversion and enantioselectivity of purified wild type *PeHNL*₁₂₁ and C-truncated *PeHNL* Δ ₁₀₇ toward different aldehydes.

Sub. 1[R=]		t=0.5 h		t=3 h		t=24 h	
		Conv. [%]	ee [%]	Conv. [%]	ee [%]	Conv. [%]	ee [%]
1a	<i>PeHNL</i> ₁₂₁	72.2	95.5	87.7	90.7	90.5	60.5
	<i>PeHNL</i> Δ ₁₀₇	76.4	97.1	92.2	93.4	94.1	75
1b	<i>PeHNL</i> ₁₂₁	48.3	92.0	81.2	82.7	95.6	65.4
	<i>PeHNL</i> Δ ₁₀₇	55.5	93.5	91.1	91.0	96.5	78.8
1c	<i>PeHNL</i> ₁₂₁	28.0	53.3	46.0	30.1	87.4	15.6
	<i>PeHNL</i> Δ ₁₀₇	29.5	67.5	59.0	58.1	90.3	38.5
1d	<i>PeHNL</i> ₁₂₁	60.8	76.6	85.6	60.0	99.5	45.5
	<i>PeHNL</i> Δ ₁₀₇	69.1	95.1	93.7	90.7	99.5	90.5
1e	<i>PeHNL</i> ₁₂₁	46.5	2.5	50.4	2.5	97.2	7.23
	<i>PeHNL</i> Δ ₁₀₇	54.2	99.8	89.0	99.4	97.4	98.9
1f	<i>PeHNL</i> ₁₂₁	30.7	22.0	72.5	9.10	96.6	5.70
	<i>PeHNL</i> Δ ₁₀₇	32.6	30.5	72.2	26.0	96.6	16.8
1g	<i>PeHNL</i> ₁₂₁	82.2	95.5	97.7	90.7	n.t.	n.t.
	<i>PeHNL</i> Δ ₁₀₇	86.4	99.1	99.2	96.4	n.t.	n.t.
1h	<i>PeHNL</i> ₁₂₁	52.4	90.3	56.5	70.0	64.1	26.0
	<i>PeHNL</i> Δ ₁₀₇	54.0	93.6	58.0	78.0	66.8	17.0
1i	<i>PeHNL</i> ₁₂₁	30.3	16.4	34.4	4.8	60.1	1.6
	<i>PeHNL</i> Δ ₁₀₇	30.6	18.3	34.2	13.6	60.1	3.4
1j	<i>PeHNL</i> ₁₂₁	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.
	<i>PeHNL</i> Δ ₁₀₇	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.

n.d. = not determined; n.t. = not tested

Reusability of *PeHNL* Δ ₁₀₇

The reusability of *PeHNL* Δ ₁₀₇ has been investigated for the production of (*R*)-MAN synthesis in biphasic batch reaction systems. Wild-type *PeHNL* and *PeHNL* Δ ₁₀₇ showed no significant differences in the production of (*R*)-MAN in the range of 50-55 μ M for the first cycle of reaction. After four cycles in the reusability experiment, the production of (*R*)-MAN of wild-type *PeHNL* was significantly decreased in the DIPE biphasic system, whereas the production of *PeHNL* Δ ₁₀₇ was almost the same as that in the first cycle of the reaction (Figure 2-6). These results are in agreement with the high solvent stability of *PeHNL* Δ ₁₀₇ in contrast to that of the wild-type form which lost more than 50 % activity upon incubation in a biphasic system for 6 h (Appendix VI). Thus, the C-terminal truncated *PeHNL* Δ ₁₀₇ without glycosylation which was expressed in *E.*

coli could improve the stability and reusability similar to that shown by the glycosylated wild-type *PeHNL* expressed in *P. pastoris*.²⁰⁾

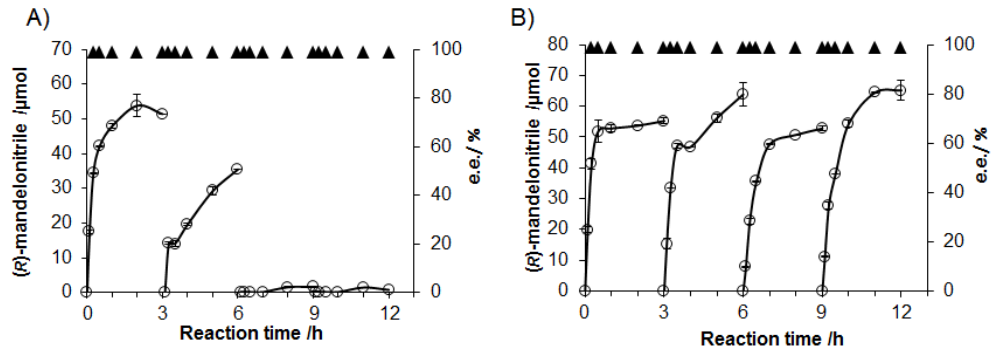


Figure 2-6. Reusability of A) *PeHNL*₁₂₁, B) *PeHNL* Δ ₁₀₇ during the production of (*R*)-mandelonitrile from benzaldehyde and potassium cyanide. ○: Amount of (*R*)-mandelonitrile. ▲: Enantiomeric excess [%].

CHAPTER III Occurrence and enzymology of hydroxynitrile lyases of cyanogenic millipedes
Section I: Effects of codon optimization and glycosylation on the high-level production of hydroxynitrile lyase from *Chamberlinius hualienensis* in *Pichia pastoris*

To date, HNLs have been mainly isolated from various plants^{18, 77}, some bacteria (with low activities)^{30, 78, 79}, and arthropods^{3, 4}. In our laboratory, eleven HNL genes were identified from five different genera of cyanogenic millipedes. Among the HNLs of arthropodal origin, the HNL from *Chamberlinius hualienensis*, henceforth referred to as ChuaHNL, was purified because it showed a high potential for industrial application^{3, 4}. ChuaHNL showed the highest specific activity (7,420 U/mg) in the synthesis of (*R*)-mandelonitrile ((*R*)-MAN) as well as high enantioselectivity in the synthesis of various cyanohydrins. However, the preparation of ChuaHNL is not suitable for applications, because collection of the millipede is quite limited to the season and the place, and ChuaHNL concentrations in the millipede body are low (26.6 U/g of millipede). Therefore, the development of a high-yield production system for recombinant ChuaHNL is necessary for industrial scale production and further utilization of the enzyme. ChuaHNL expression has been investigated using *P. pastoris*³, *E. coli* BL21 (DE3), *E. coli* Shuffle T7, and insect cell (Sf9) systems⁴. The recombinant ChuaHNL was produced only in the eukaryotic expression systems of *P. pastoris* and Sf9 cells, which produced 305 ± 9.8 and 400 ± 100 U/L culture, respectively. The productivity of both methods was still not enough for industrial use. The Sf9 cell system is difficult to handle and is an expensive system to use for enzyme production. Therefore, we further investigated *P. pastoris* as an alternative expression system for the high-level production of a recombinant ChuaHNL. It has been reported that ChuaHNL is a glycoprotein³, and all millipede HNLs have the eight conserved cysteine residues that might be involved, via disulfide bonds, during the folding process to generate the active form of HNL⁴. In addition, we revealed that the ChuaHNL gene exhibited a high level of codon usage bias in *P. pastoris*. Therefore, we considered that these factors should be the limiting in the high-yield production of ChuaHNL. In this study, we examined the effects of codon-optimization for *ChuaHNL*, co-expression of *PDI*s, and a combination of co-expression and codon-optimized *ChuaHNL* and *PDI*s on the yield of His-ChuaHNL expression in *P. pastoris*. Furthermore, we analyzed the effects of glycosylation on the production of His-ChuaHNL in the *P. pastoris* expression system.

MATERIALS AND METHODS

Strains, media, and cultivation conditions

E. coli HST08 (Takara Bio, Otsu, Japan) was incubated at 37°C in Luria-Bertani (LB) medium (1% tryptone, 0.5% yeast extract, and 0.5% NaCl) containing ampicillin (50 mg/ml) or zeocin (25 mg/ml) for plasmid propagation. *P. pastoris* GS115 was used as a host for expression. Each transformed *P. pastoris* was grown at 30°C on an agar plate of YPDS medium (2% glucose, 2% peptone, 1% yeast extract, and 1M sorbitol) containing 100 µg/mL zeocin (Invitrogen, Carlsbad, CA, USA). Next, the yeast strain was pre-cultured at 30°C in a buffered minimal glycerol medium (BMG medium; 100 mM potassium phosphate buffer, pH 7.0, 1.34% yeast nitrogen base without amino acids, 4×10^{-5} % biotin and 1% glycerol) or BMGH medium (BMG medium plus 0.004% histidine). The grown cells were aseptically, collected, and transferred into induction medium, BMM medium (1% glycerol of BMG medium was replaced by 1% methanol) or BMMH medium (BMM medium plus 0.004% histidine). Cultivation was performed at 30 or 28 °C under aerobic conditions with reciprocal shaking (150 rpm), and the cell growth was monitored by measuring the optical density at 600 nm.

Enzyme activity

The enzyme activities for MAN synthesis and MAN degradation were assayed by HPLC and spectrophotometric methods, respectively, as described previously³⁾. Briefly, ChuaHNL mediated MAN synthesis was assayed using 50 mM benzaldehyde and 100 mM KCN as substrates in 400 mM sodium citrate buffer, pH 4.0 (total reaction volume; 0.5 mL). The reaction was initiated by addition of KCN at 25°C for 3 min, and it was stopped by mixing 100 µL of the reaction mixture with 900 µL of organic solvent (85% *n*-hexane, 15% isopropanol by volume). After the mixture was centrifuged (15,000 ×*g*, 10 min), (*R*)-MAN was quantified by HPLC, injecting 10 µL of the organic phase. One unit of HNL activity was defined as the amount of the enzyme that catalyzes the synthesis of 1 µmol of (*R*)-MAN per min. The MAN degradation activity was performed at 25°C as mentioned here. The reaction was started by adding enzyme sample to 0.5 mL of 100 mM sodium citrate buffer, pH 5.5, containing 2 mM racemic MAN. The reaction velocity of benzaldehyde formation was followed by monitoring absorbance at 280 nm for one min using a UV-2600 UV-Vis spectrophotometer (Shimadzu, Kyoto, Japan).

Construction of strain expressing codon-optimized ChuaHNL

The oligonucleotide primers used in this study are listed in Appendix VII. ChuaHNL cDNA was amplified by PCR with KOD-Plus-Neo-DNA polymerase (Toyobo, Osaka, Japan) using His-ChuaHNL-Fw and His-ChuaHNL-Rv as primers and ChuaHNL cDNA (GenBank accession no. LC203137)³⁾ as a template. The resulting PCR product also contained a Kex2 signal, an 8x His-tag, and TEV protease cleavage sites (ALEKRAHHHHHHHHENLYFQGS, 22 aa) attached at the N-terminus. The product was digested with *XhoI* and *XbaI* (Takara Bio) and cloned into a *P. pastoris* expression vector, pPICZ α A (Thermo Fisher Scientific, Waltham, MA, USA), under the *AOXI* promoter, yielding pPICZ α A-His-Ori-ChuaHNL.

The codon optimization of the ChuaHNL gene was carried out by Life Technology Japan (Tokyo, Japan). Synthetic ChuaHNL with N-terminal His tag (His-Syn-ChuaHNL) was ligated to pPICZ α A, using an In-Fusion HD Cloning Kit (Clontech Laboratories, Palo Alto, CA, USA), yielding pPICZ α A-His-Syn-ChuaHNL. Both plasmids, pPICZ α A-His-Ori-ChuaHNL and pPICZ α A-His-Syn-ChuaHNL, were linearized by *SacI* (Takara Bio) and used to transform *P. pastoris* GS115 strain using the *Pichia* EasyComp Transformation kit (Invitrogen, Karlsruhe, Germany). The resulting strains were named His-Ori-ChuaHNL/GS115 and His-Syn-ChuaHNL/GS115, respectively.

Construction of strains expressing protein disulfide isomerases (PDIs) and ChuaHNL

To evaluate the effect of PDIs, the strains containing different PDI genes (one from *P. pastoris*, *PpPDI*, and the other two from *C. hualienensis*, *ChuaPDI1* and *ChuaPDI2*) were constructed as mentioned here. The *PpPDI* (GenBank accession no. AJ302014) was amplified by PCR using KOD-Plus-Neo-DNA polymerase (Toyobo) and the primers PpPDI-InFu-Fw and PpPDI-InFu-Rv (Appendix VII). The genomic DNA of *P. pastoris* was used as a template DNA. The cDNA of PDI1 from *C. hualienensis* (*ChuaPDI1*) was amplified by PCR using KOD-Plus-Neo-DNA polymerase and the primer pair ChuaPDI1-InFu-Fw and ChuaPDI1-InFu-Rv (Appendix VII). The cDNA prepared from the paraterga of *C. hualienensis* was used as template DNA. The cDNA of PDI2 from *C. hualienensis* (*ChuaPDI2*) was also amplified by PCR using the primers ChuaPDI2-InFu-Fw and ChuaPDI2-InFu-Rv (Appendix VII) by the same procedure as described above. The resulting DNA fragments were cloned into the *EcoRI* restriction site of the *P. pastoris* expression empty vector pAO815 under the *AOXI* promoter using In-Fusion HD Cloning Kit and *EcoRI* (Takara Bio), and the junction was confirmed by sequencing. The resulting plasmids and

empty pAO815 were linearized by *StuI* (Takara Bio) and used to transform the His-Ori-ChuaHNL/GS115 strain or the His-Syn-ChuaHNL/GS115 strain. The resulting strains were named His-Ori-ChuaHNL/PpPDI/GS115, His-Syn-ChuaHNL/pAO815/GS115, His-Syn-ChuaHNL/PpPDI/GS115, His-Syn-ChuaHNL/ChuaPDI1/GS115, and His-Syn-ChuaHNL/ChuaPDI2/GS115.

Construction of strains expressing deglycosylated ChuaHNL

It was predicted that ChuaHNL has three glycosylation sites. Accordingly, the asparagine residues in the tripeptide sequences Asn-X-Ser/Thr (where X can be any amino acid except proline) at N99, N109, and N123 were targeted for mutagenesis. The mutants with a deglycosylated ChuaHNL gene were constructed using pPICZ α A-His-Syn-ChuaHNL as a template. The plasmids with a deglycosylation site at N99, N109 or N123 were prepared by mutation using a Quick-Change Lightning Site-Directed Mutagenesis Kit (Aligent Technologies, Santa Clara, CA, USA) and gene-specific primer (Appendix VII). The vectors of pPICZ α A-His-Syn-ChuaHNL-N99Q, pPICZ α A-His-Syn-ChuaHNL-N109Q, and pPICZ α A-His-Syn-ChuaHNL-N123Q correspond to a mutation of N99Q, N109Q, or N123Q, respectively. The pPICZ α A-His-Syn-ChuaHNL-N99Q/N109Q, pPICZ α A-His-Syn-ChuaHNL-N99Q/N123Q, pPICZ α A-His-Syn-ChuaHNL-N109Q/N123Q, and pPICZ α A-His-Syn-ChuaHNL-N99Q/N109Q/N123Q were also obtained. The vector carrying each mutant was transformed into *E. coli* HST08 competent cells and selected on a low-salt LB agar plate with Zeocin (25 μ g/mL). DNA sequences were determined with a 3500 Genetic Analyzer (Thermo Fisher Scientific) and assembled and analyzed by using ATGC and Genetyx Ver. 12 (Genetyx, Tokyo, Japan). *SacI* (Takara Bio) was used to linearize all vectors, and the resulting products were electroporated into *P. pastoris* cells harboring PpPDI/GS115 using an Easysselect Pichia Expression Kit (Thermo Fisher Scientific), according to the manufactures instructions.

Expression of His-ChuaHNL gene in transformants

P. pastoris transformants harboring pPICZ α A-His-Ori-ChuaHNL or pPICZ α A-His-Syn-ChuaHNL were pre-cultured in BMGH medium (500 mL) at 30°C for 48 h in baffled flasks (2L) with shaking at 150 rpm. Next, the cells were harvested by centrifugation (6,500 \times g, 10 min) and resuspended into BMMH medium (500 mL) followed incubating at 28°C with reciprocal shaking for 6 days. The co-expression strains, His-Ori-ChuaHNL/PpPDI/GS115, His-Syn-ChuaHNL/pAO815/GS115, His-Syn-ChuaHNL/PpPDI/GS115, His-Syn-

ChuaHNL/ChuaPDI1/GS115, His-Syn-ChuaHNL/ChuaPDI2/GS115, and His-Syn-ChuaHNL/PpPDI/GS115 also containing the deglycosylated enzyme gene were pre-cultured in BMG and transferred into BMM medium for enzyme induction. The cultivation was performed under the same condition as described above.

Purification of His-ChuaHNL from transformants

Each of the supernatant solutions containing the expressed His-ChuaHNL was collected by centrifugation at 3,000 ×g for 5 min. The pH of the supernatant was adjusted to 7.5 with NaOH and directly applied to an Ni sepharose 6 fast flow column (GE Healthcare, Buckinghamshire, UK), which was equilibrated with buffer A (20 mM potassium phosphate buffer (KPB), pH 7.5, plus 300 mM NaCl) containing 20 mM imidazole. The column was washed with buffer A containing 20 mM imidazole, and absorbed enzyme was eluted with buffer A containing 300 mM imidazole. The active fractions were concentrated, and the buffer was changed to 20 mM KPB, pH 7.0. The enzymes produced by His-Ori-ChuaHNL/PpPDI/GS115, Syn-ChuaHNL/PpPDI/GS115, and corresponding deglycosylated mutant strains were further purified by MonoQ 5/50 GL column chromatography (GE Healthcare), in which the absorbed enzyme was eluted by a linear gradient with 20 mM KPB, pH 7.5, and 20 mM KPB, pH 7.5, containing 400 mM NaCl. The active fractions were also concentrated, and the buffer was changed to 20 mM KPB, pH 7.0. The purity of these His-ChuaHNLs was analyzed by sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) ³¹.

Stability against temperature, low pH, and solvents

The thermal stability of wild-type and deglycosylated enzymes was analyzed after the enzymes were incubated with 20 mM KPB, pH 7.0, in the range of 30°C to 90°C for 1 h without substrate. The stability at low pH was analyzed after the enzymes were incubated at 25°C for 24 h with 50 mM sodium citrate buffer, pH 4.0, without substrate. To analyze the solvent stability, methyl-*tert*-butyl ether (MTBE) was mixed with 50 mM KPB, pH 7.0, (v/v; 50/50) and equilibrated with shaking at 1500 rpm for 60 min. The purified wild-type or deglycosylated enzyme (each 0.33 µg) was added to the buffer phase and gently mixed at 10°C for 12 h without disturbing the interface between the aqueous and organic phase. After these treatments, the remaining activity was measured under standard assay conditions of MAN synthesis.

(R)-Mandelonitrile synthesis

The efficiency of (*R*)-MAN synthesis for wild-type and deglycosylated His-ChuaHNLs was measured using 250 mM benzaldehyde and 500 mM KCN in 1.5 mL of a biphasic system consisting of 400 mM citrate buffer, pH 4.0, and MTBE (each 0.75 mL). The reaction was started by adding 1.0 U of wild-type or deglycosylated His-ChuaHNL and incubating at 10°C for 3 h with shaking at 1,500 rpm. The (*R*)-MAN quantities were determined by the HPLC method for the measurement of (*R*)-MAN synthesis activity. The analysis of the enantiomeric excess (*e.e.*) or (*R*)-MAN synthesis was performed using 20 U of His-ChuaHNL and determined according to relative concentrations (mM) of the two enantiomers by using Equation (1) as describe in chapter I.

Western blotting analysis

The protein samples were separated by SDS-PAGE and transferred onto a polyvinylidene difluoride (PVDF) membrane (ATTO, Tokyo) using a powerBLOTmini (ATTO). The His-tagged proteins on the membrane were immunoblotted with Anti-His-tag mAb-HRP-Direct (MBL) after blocking with a 5% solution of skim milk (Wako, Tokyo Japan). Immunodetection was performed with an Amersham ECL Prime Western Blotting Detection Reagent (GE healthcare), according to the instruction of the manufacturers.

Protein concentrations

Protein concentration was determined by the method of Bradford⁸⁰⁾ using a Bio-Rad protein assay kit (Bio-Rad Laboratories, Hercules, CA, U.S.A.) and bovine serum albumin as a protein standard.

Nucleotide sequence accession numbers

Sequence alignments were done using the ClustalW. The nucleotide sequences of Synthetic *ChuaHNL*, *ChuaPDII*, and *ChuaPDI2* were deposited in DDBJ under accession no. LC422783, LC422784, and LC422785, respectively.

RESULTS

Effect of codon optimization on ChuaHNL production

When the His-Ori-ChuaHNL/GS115 strain was incubated in BMMH medium at 28°C for 6 days, the production of His-ChuaHNL was 22.6 ± 3.8 U/L of culture in (*R*)-MAN synthesis activity, which is not enough for industrial uses. Next, we analyzed the DNA codon sequence of ChuaHNL and found that the codons of the ChuaHNL gene exhibited a high codon usage bias for *P. pastoris*. Since it was presumed that the remarked difference in codons is one of limiting factors in the efficient production of His-ChuaHNL in the *P. pastoris* expression system, we optimized the codons of ChuaHNL gene to *P. pastoris* preferred codons⁸¹⁾ and obtained a synthetic *ChuaHNL*. In this codon optimization, a total of 100 codons from the original *ChuaHNL* were substituted, which accounted for 61.7 % of the total amino acid residues ([Appendix VIII](#)). We next evaluated the effect of codon optimization on His-ChuaHNL production using the synthetic *ChuaHNL* in *P. pastoris* expression system (His-Syn-ChuaHNL/pAO815/GS115 strain). As expected, the production of His-ChuaHNL significantly increased to $1,950 \pm 86.5$ U/L of culture at 6 days of incubation, which is 86 times higher than that by His-Ori-ChuaHNL/GS115 strain (Table 3-1). These results suggest that the translation efficiency of the codons is one of key factors responsible for increasing His-ChuaHNL production in the *P. pastoris* expression system.

Table 3-1. Effects of codon optimization of *ChuaHNL* and *PpPDI* on the production of His-ChuaHNL

Strain	Enzyme activity ^a (U/ml)	Protein (mg/ml)	Specific activity (U/mg)	Enzyme productivity (U/L of culture)
His-Ori-ChuaHNL/GS115	0.023	N.D.	N.D.	22.6 ± 3.8
His-Syn-ChuaHNL/pAO815/GS115	1.95	0.014	139	$1,950 \pm 86.5$
His-Ori-ChuaHNL/PpPDI/GS115	0.458	0.011	41.8	458 ± 12.6

[a] The HNL activity of (*R*)-MAN synthesis was assayed by a HPLC method. Values are the means \pm SD; n=3. N.D., not detected.

Isolation of PDI gene candidates from C. hualienensis

Since it has been reported that ChuaHNL has eight cysteine residues, we further investigated the effect of PDIs on His-ChuaHNL production in the *P. pastoris* expression system. At first, the sequence of *PpPDI* was obtained from GenBank (accession no. AJ302014). To obtain

PDIs from *C. hualienensis*, a homology search against *PpPDI* was performed using BLAST against the RNA sequence data of *C. hualienensis*⁸²⁾. As a result, two candidate PDI genes, *ChuaPDI1* and *ChuaPDI2*, were isolated. *ChuaPDI1* and *ChuaPDI2* consist of 501 and 493 amino acid residues, respectively, and both candidates were 29.2% identical to each other at the amino acid sequence level. The deduced amino acid sequence of *ChuaPDI1* and *ChuaPDI2* was 29.6% and 27.9% identical to that of *PpPDI*, respectively. Both *ChuaPDI* candidates had thioredoxin domains with putative two active sites⁸³⁾, which show high sequence similarity to the corresponding regions of *PpPDI*. Two cysteine residues, which are important in PDIs, were also conserved in both *ChuaPDI* candidates (a and a', in Figure 3-1).

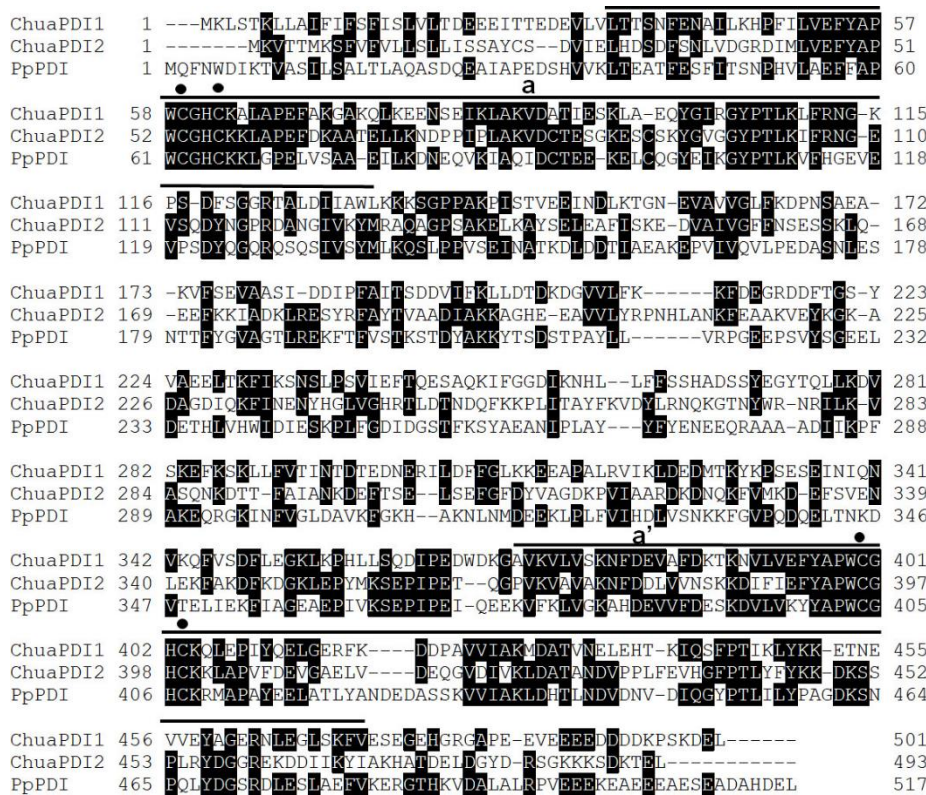


Figure 3-1. Alignment of the deduced amino acid sequences of *PpPDI*, *ChuaPDI1*, and *ChuaPDI2*. White letters indicate identical residues between those of *ChuaPDI1* and *ChuaPDI2*. The two active site domains (a and a') are over-lined and the closed circles indicated the predicted cysteine residues in the active site

Effect of PDI genes on His-ChuaHNL production

To evaluate the effect of PDI genes on His-ChuaHNL production, we investigated the co-expression with *PpPDI* in the *P. pastoris* expression system using the His-Ori-ChuaHNL/*PpPDI*/GS115 strain. By this method, His-ChuaHNL production increased to 458 ± 12.6 U/L of culture at 6 days of incubation (Table 3-1). Thus, co-expression of *PpPDI* was effective to improve His-

ChuaHNL production in the *P. pastoris* expression system. However, the His-ChuaHNL production by this method was lower than that of the above codon-optimized method.

Since it was demonstrated that the codon optimization of *ChuaHNL* was much better than the co-expression with *PpPDI*, following studies were performed using the synthetic *ChuaHNL* to obtain high production of His-ChuaHNL. Three transformed strains, His-Syn-ChuaHNL/*PpPDI*/GS115, His-Syn-ChuaHNL/*ChuaPDI1*/GS115 and His-Syn-ChuaHNL/*ChuaPDI2*/GS115 were constructed, and the His-ChuaHNL productivity of each strain was compared. All transformants showed a similar growth rate and reached to stationary phase after 7-8 days of induction (Figure 3-2a). Thus, the cell growth was not affected by co-expression of codon-optimized *ChuaHNL* and *PDI*s in BMM medium. The ChuaHNL activity of MAN degradation increased by prolonging the cultivation time, reaching a maximum at 7 days of incubation in all transformants. Notably, MAN degradation activity was remarkably different among *PDI* species (Figure 3-2b). We also assayed the HNL activity of MAN synthesis at 6 days of cultivation to confirm the correct MAN synthesis values. The His-ChuaHNL productivity by His-Syn-ChuaHNL/*PpPDI*/GS115, His-Syn-ChuaHNL/*ChuaPDI1*/GS115, and His-Syn-ChuaHNL/*ChuaPDI2*/GS115 was $3,170 \pm 144.7$ U/L, $2,060 \pm 43.7$ U/L, and $2,290 \pm 40.7$ U/L of culture, respectively (Figure 3-2c). The productivity of these strains was 6.9, 4.5, and 5.0-fold higher than that of the His-Ori-ChuaHNL/*PpPDI*/GS115 strain, which expressed non-codon optimized *ChuaHNL* and *PpPDI*. Thus, co-expression of codon-optimized *ChuaHNL* and *PDI*s was more effective than the expression of codon-optimized *ChuaHNL* without *PDI*s. These results also

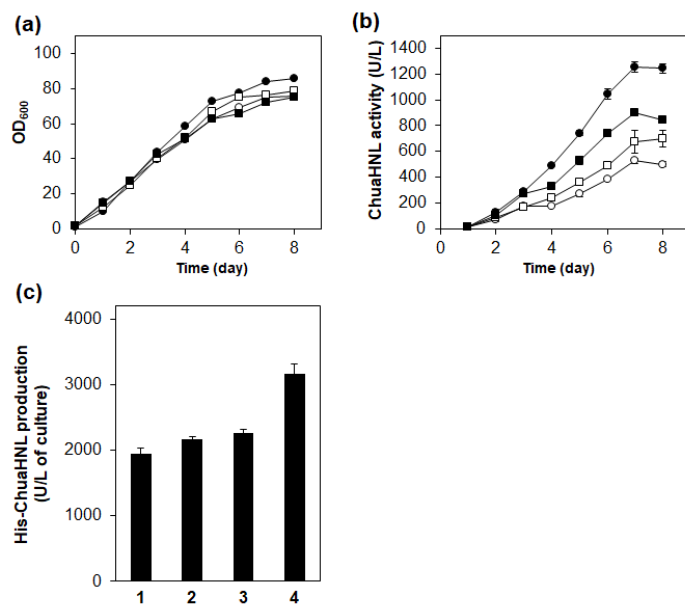


Figure 3-2. Effects of *PpPDIs* on His-ChuaHNL production. Three transformants with codon optimized *ChuaHNL* and different *PDI*s were cultured in BMM medium. (A) Cell growth. (B) HNL activities of MAN degradation to benzaldehyde were assayed every day. In (A) and (B), open circles, His-Syn-ChuaHNL/*pAO815*/GS115 strain; closed circles, His-Syn-ChuaHNL/*PpPDI*/GS115 strain; open squares, His-Syn-ChuaHNL/*ChuaPDI1*/GS115 strain; closed squares, His-Syn-ChuaHNL/*ChuaPDI2*/GS115 strain. (C) His-ChuaHNL productivity after 6 days of induction. The HNL activity of (*R*)-MAN synthesis was measured by a HPLC method. 1, His-Syn-ChuaHNL/*pAO815*/GS115 strain; 2, His-Syn-ChuaHNL/*ChuaPDI1*/GS115 strain; 3, His-Syn-ChuaHNL/*ChuaPDI2*/GS115 strain; 4, His-Syn-ChuaHNL/*PpPDI*/GS115 strain. Values are the means \pm SD; n=3

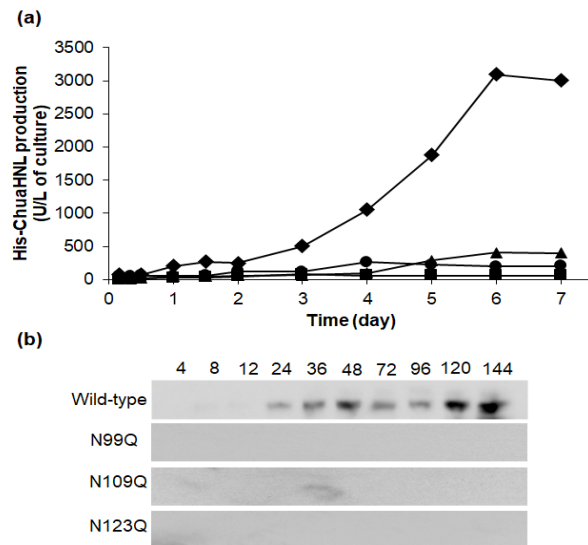


Figure 3-3. Effects of glycosylation sites on His-ChuaHNL production. (A) The time course of His-ChuaHNL production was assayed by (R)-MAN synthesis activity after 4, 8, 12, 24, 36, 48, 72, 96, 120, and 144 h of cultivation. (B) Western blot analysis at the same cultivation time. Enzyme activities of His-ChuaHNL-WT, His-ChuaHNL-N99Q, His-ChuaHNL-N109Q, and His-ChuaHNL-N123Q are shown in diamonds, triangles, squares, and circles, respectively

indicated that *PDI* from *P. pastoris* is more effective than *PDI*s from *C. hualienensis* at increasing His-ChuaHNL productivity in the co-expression system using codon-optimized genes.

Effect of glycosylation on His-ChuaHNL production

Since it has been predicted that ChuaHNL has three N-glycosylation sites (N99, N109 and N123), the effects of sugar chains on the production of His-ChuaHNL was investigated using three mutants, N99Q, N109Q and N123Q, in the above co-expression system. The production of His-ChuaHNL-N99Q, His-ChuaHNL-N109Q, and His-ChuaHNL-N123Q were 404 ± 6.6 ,

58 ± 15 , and 210 ± 28 U/L of culture at 6 days of incubation, which are 7.6, 53, and 15 times lower than that of His-ChuaHNL-WT ($3,029 \pm 50$ U/L by His-Syn-ChuaHNL/PpPDI/GS115 strain), respectively (Figure 3-3a). Thus, enzyme production was decreased by the deglycosylation of ChuaHNL, while the growth of the strains was similar to that of the His-Syn-ChuaHNL/PpPDI/GS115 strain. The secretion of deglycosylated His-ChuaHNLs was also not detected by western blotting analysis (Figure 3-3b). Furthermore, the production of deglycosylated enzymes within the cells was remarkably lower than that of intracellular His-ChuaHNL-WT (intracellular enzyme production of His-ChuaHNL-WT, His-ChuaHNL-N99Q, His-ChuaHNL-N109Q, and His-ChuaHNL-N123Q were 285, 88, 5, and 87 U/L, respectively ([Appendix IX](#))). The enzymes deglycosylated at two positions (N99Q/N109Q, N99Q/ N123Q, or N109Q/N123Q) and all three positions were not produced in the culture broth (data not shown). These results indicated that sugar chains are important for the production of His-ChuaHNL.

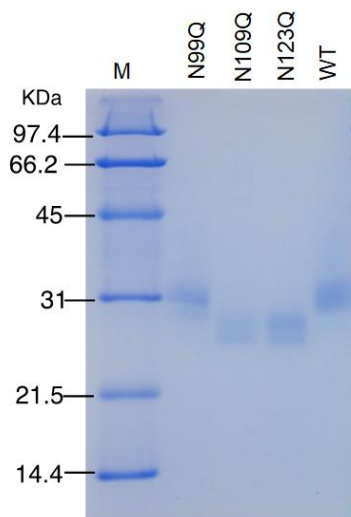


Figure 3-4. SDS-PAGE analysis of purified His-ChuaHNL produced by His-Syn-ChuaHNL/PpPDI/ strain. The enzyme was purified by the procedure described in the Materials and Methods and analyzed by 15% SDS-PAGE

Molecular mass of deglycosylated enzymes

The molecular mass of wild-type and the three deglycosylated His-ChuaHNLs were analyzed by SDS-PAGE. The molecular mass of His-ChuaHNL-N99Q was similar to that of His-ChuaHNL-WT (30-31.3 kDa) and His-ChuaHNL-N109Q, while His-ChuaHNL-N123Q was smaller than that of His-ChuaHNL-WT (27.4-29.1 kDa) (Figure 3-4). These results indicated that the N109 and N123 are glycosylated, but the degree of glycosylation at residue N99 would be very small.









Effects of glycosylation on substrate specificity and kinetic properties

Since it was reported that ChuaHNL preferentially acts on various aromatic aldehydes³⁾, we analyzed the effects of glycosylation sites on (*R*)-cyanohydrin synthesis using His-ChuaHNL-WT and the three deglycosylated His-ChuaHNLs. The wild-type and the three deglycosylated enzymes showed (*R*)-cyanohydrin synthesis activity toward all of aromatic aldehydes tested (Table 3-2). However, the specific activities of the deglycosylated mutants were remarkably different from that of His-ChuaHNL-WT. Among the deglycosylated enzymes, N109Q exhibited the lowest activity toward all of the aldehydes tested, except 4-biphenyl carboxaldehyde, and N123Q exhibited the highest activity toward almost all of the aldehydes tested. The enzyme activity of His-ChuaHNL-N99Q was also lower than that of His-ChuaHNL-WT for all the aldehydes tested, but slightly higher than that of His-ChuaHNL-N109Q for all the aldehydes tested, except 4-biphenyl carboxaldehyde. Thus, the sugar at N109 exhibited an activating effect and that at N123 exhibited inhibitory effect. This is because the His-ChuaHNL-N109Q enzyme keeps glycosylation at N99 and N123, while the His-ChuaHNL-N123Q enzyme keeps the glycosylation at N99 and N109. The K_m values of these four enzymes for benzaldehyde were determined next.

Table 3-3 Kinetic values for benzaldehyde of wild-type and glycosylated His-ChuaHNLs in the (*R*)-MAN synthesis reaction. The K_m values for benzaldehyde in (*R*)-MAN synthesis were assayed under standard assay conditions of MAN synthesis reaction using 0.5–50 mM benzaldehyde and 100 mM KCN at pH 4.0

	K_m (mM)	V_{max} ($\mu\text{mol min}^{-1} \text{mg}^{-1}$)	k_{cat} (s^{-1})	k_{cat}/K_m ($\text{mM}^{-1} \text{s}^{-1}$)
His-ChuaHNL-WT	4.65±0.35	5,580±200	1,950±70	422±47
His-ChuaHNL-N99Q	5.30±0.06	3,490±130	1,220±44	230±11
His-ChuaHNL-N109Q	4.13±0.15	1,705±13	595±4.5	144±5
His-ChuaHNL-N123Q	3.10±0.11	2,330±10	815±3.0	263±8

Table 3-2 Effects of His-ChuaHNLs glycosylation sites on cyanohydrins synthesis activity

Substrate	Specific activity (U/mg)				
	His-WT	His-N99Q	His-N109Q	His-N123Q	ChuaHNL ^a
	5,280±105 (100)	3,090±43 (58.5)	1,310±10 (24.8)	1,985±57 (37.6)	8,186±171
	1,120±30 (100)	336±26 (30.0)	225±13 (20.1)	890±17 (79.5)	N.D.
	10,280±190 (100)	4,860±70 (47.3)	4,520±15 (44.0)	17,200±150 (167.3)	16,639±1491
	10,320±906 (100)	5,450±78 (52.8)	4,620±370 (44.8)	16,070±286 (155.7)	13,338±611
	2,130±186 (100)	840±52 (39.4)	880±46 (41.3)	3,140±67 (147.4)	4,060±211
	2,220±54 (100)	753±27 (33.9)	335±32 (15.1)	1,610±76 (72.5)	8,394±373
	3,370±68 (100)	1,480±52 (43.9)	1,050±26 (31.2)	3,280±30 (97.3)	3,605±204
	95±12 (100)	38.5±3 (40.5)	70±14 (73.7)	273±16 (287.4)	118.7±8.4

[a] The V_{max} values of native ChuaHNL (Dadashipour *et al.*, 2015). N.D., not described. The values in parentheses are relative activity to the enzyme activity of His-ChuaHNL-WT

The K_m values of His-ChuaHNL-WT, His-ChuaHNL-N99Q, His-ChuaHNL-N109Q, and His-ChuaHNL-N123Q were 4.65 ± 0.35 , 5.30 ± 0.06 , 4.13 ± 0.15 , and 3.10 ± 0.11 , respectively (Table 3-3). Thus, the sugars at three positions did not affect the affinity for benzaldehyde, whereas the catalytic efficiency of MAN synthesis was remarkably expedited by the presence of the sugars.

Effects of glycosylation on enzyme stability

When His-ChuaHNL-WT and the three glycosylated His-ChuaHNLS were incubated at pH 7.0 for 1 h at a temperature below 55°C, more than 90% of the original activities remained in all enzymes. His-ChuaHNL-WT retained more than 90% of its activity by incubation at pH 7.0 for 1 h at 65°C, whereas His-ChuaHNL-N99Q, His-ChuaHNL-N109Q, and His-ChuaHNL-N123Q lost 20%, 40%, and 40% of enzyme activity, respectively (remaining activities are 80%, 60%, and 60%, respectively). The His-ChuaHNL-WT was also more stable than the deglycosylated enzymes when incubated at higher temperatures (Figure 3-5a). Thus, glycosylation contributes to enzyme stability. Specifically, the sugars at N109 and N123 play important roles for thermostability.

Since it has been reported that the reaction toward cyanohydrin synthesis is optimal in the acidic region ³⁾, the pH stability was analyzed by incubating at pH 4.0 and 25°C for 24 h. The His-ChuaHNL-N99Q enzyme gradually lost activity after 3 h of incubation, and the remaining activity was reduced to 80% after 24 h of incubation, whereas the other two deglycosylated and wild-type enzymes retained more than 90% of the original activity (Figure 3-5b). These results indicated that the sugar at N99 contributes to stability in a lower pH region.

In our preliminary study, we found that MTBE is a suitable organic solvent for cyanohydrin synthesis by His-ChuaHNL when used in a biphasic system. So, we compared the solvent stability of the wild-type and deglycosylated enzymes in a biphasic system consisting of MTBE and 50 mM KPB, pH 7.0 (v/v; 50/50). His-ChuaHNL-WT and His-ChuaHNL-N99Q retained their original activities after incubation at 25°C for 12 h. In contrast, His-ChuaHNL-N123Q lost 20% of its activity after 12 h incubation in the biphasic system (Figure 3-5c). These results indicated that the sugar chain at N123 plays an important role on solvent stability in the biphasic system.

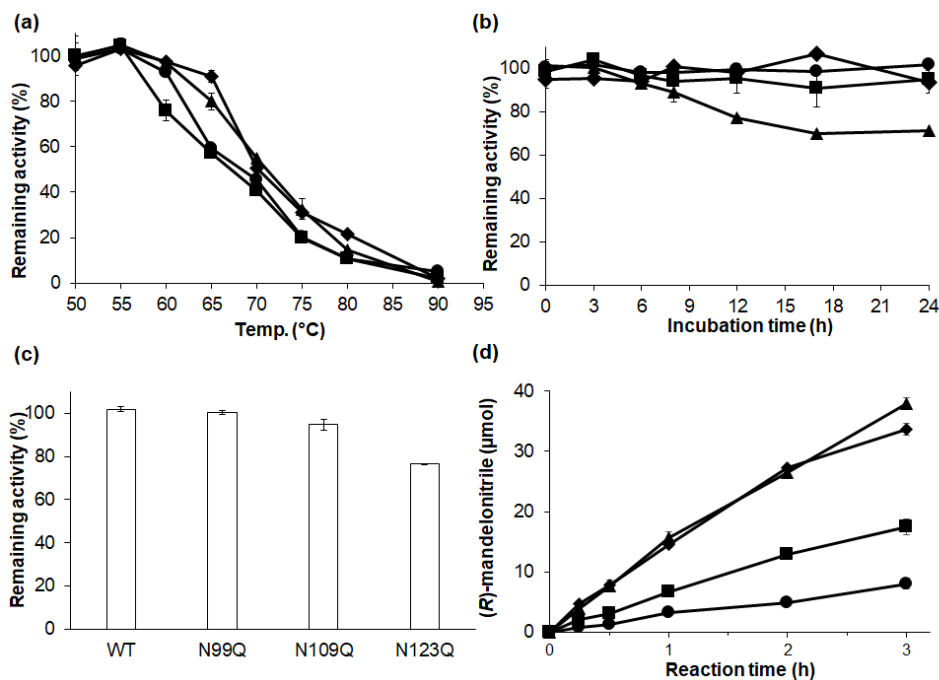


Figure 3-5. Effects of glycosylation sites on enzyme stability. (A) Thermostability. (B) Stability at pH 4.0. (C) Solvent stability at pH 7.0. (D) Synthesis of MAN in biphasic systems with MTBE and citrate buffer, pH 4.0 (50% v/v). His-ChuaHNL-WT (diamonds), His-ChuaHNL-N99Q (triangles), His-ChuaHNL-N109Q (squares), His-ChuaHNL-N123Q (circles)

Effects of glycosylation on the synthesis of (R)-MAN

The synthesis of (R)-MAN was performed using the aforementioned enzymes (each 1.0 U) in the biphasic system. His-ChuaHNL-WT and His-ChuaHNL-N99Q produced 33.5–38.0 μM of (R)-MAN after 3 h of incubation, whereas His-ChuaHNL-N109Q and His-ChuaHNL-N123Q produced 17.5 and 8.0 μM of (R)-MAN, respectively. Thus, the production of (R)-MAN by the deglycosylated enzymes at N109 or N123 was lower than that of His-ChuaHNL-WT and His-ChuaHNL-N99Q (Figure 3-5d). These results indicate that sugar components at N109 and N123 are effective for the efficient synthesis of (R)-MAN in a biphasic system composed of MTBE and buffer. Next, 50 mM benzaldehyde and 100 mM KCN or 250 mM benzaldehyde and 500 mM KCN were incubated with His-ChuaHNL-WT (20 U) in 1.5 mL of the biphasic system at 10°C for 60 min, and the *e.e.* of (R)-MAN synthesis was analyzed. In this reaction, (R)-MAN was synthesized with 99% and 95% *e.e.*, respectively (data not shown). Thus, His-ChuaHNL-WT produced by our expression system was also useful for (R)-MAN production with a high *e.e.* in the biphasic system.

DISCUSSION

Recently, a novel HNL with high enantioselectivity of (*R*)-cyanohydrins synthesis was discovered in a cyanogenic millipede, *C. huaiensis*, and its unique characteristics were revealed³⁾. The enzyme is a glycoprotein and contains eight cysteine residues that could be involved in the formation of disulfide bonds. The enzyme exhibits the highest specific activity, pH stability, and thermostability among all the characterized HNLs. These enzymatic characteristics indicated that ChuaHNL is an excellent candidate for industrial (*R*)-cyanohydrin synthesis. More recently, Yamaguchi et al. isolated 10 HNL genes from four other millipedes and revealed that these millipede HNLs also showed high activities for (*R*)-MAN synthesis, but the specific activities were lower than that of ChuaHNL. Their deduced amino acid sequences showed 45-66% similarities to that of ChuaHNL and all eight cysteine residues were conserved, whereas the N-glycosylation sites were not⁴⁾.

Yamaguchi et al. also investigated the heterologous expression of millipede HNLs using *E. coli* BL21 (DE3), *E. coli* Shuffle T7, and Sf9 cells. All of the above millipede HNLs could not be produced by *E. coli* BL21 (DE3) strain. While five millipede HNLs were produced in the *E. coli* Shuffle T7 strain, ChuaHNL was not. All millipede HNLs were produced by Sf9 cells, but ChuaHNL production was very low at 400 ± 100 U/L of culture⁴⁾. We investigated the production of ChuaHNL and His-ChuaHNL using *P. pastoris*, which was used as a host for plant HNL production^{20, 28, 50, 84)}. Both ChuaHNL and His-ChuaHNL were produced by *P. pastoris*, but the production level was very low (305 ± 9.8 U/L of culture and 22.6 ± 3.8 U/L of culture, respectively). Since it has been reported that codon optimization was effective on heterologous protein expression in *P. pastoris*⁸⁵⁻⁸⁷⁾, we first analyzed the effect of codon-optimization of the ChuaHNL gene on His-ChuaHNL production. It has been predicted that ChuaHNL has eight cysteine residues and three N-glycosylation sites in the subunit molecule. Therefore, we also evaluated co-expression of *PDI*s and glycosylation sites in this chapter.

In the analyses of these factors, we revealed that the His-ChuaHNL production by the codon optimized *ChuaHNL* was 86 times higher than that by the original *ChuaHNL*, whereas that by the co-expression of *PpPDI* was 20 times higher (Table 3-1). We also revealed that the His-ChuaHNL production by the co-expression of codon-optimized *ChuaHNL* and *PpPDI*, *ChuaPDI1*, or *ChuaPDI2* was 146, 91, and 101 times higher, respectively, than that of His-Ori-ChuaHNL alone. Thus, the co-expression of codon-optimized *ChuaHNL* and *PDI*s was more effective than the codon

optimization of *ChuaHNL* or co-expression of *PpPDI* to improve His-ChuaHNL productivity in the *P. pastoris* expression system, and *PDI* from *P. pastoris* was better than *PDI*s from *C. hualienensis* among the *PDI*s tested. From these results, it was presumed that the high-level production of His-ChuaHNL might be affected by the translation efficiency of the ChuaHNL gene and the correct formation of the S-S bridges in the *P. pastoris* expression system.

Next, we analyzed effects of glycosylation on His-ChuaHNL production, and the enzyme activities and stabilities using three mutant enzymes, N99Q, N109Q and N123Q, all predicted sugar-binding residues in ChuaHNL. The secretion of these three deglycosylated enzymes dramatically decreased, and the production of deglycosylated His-ChuaHNLs was remarkably lower than that of the wild type His-ChuaHNL (Figure 3-3). Thus, the glycosylation at the above three positions was essential for a high-level production of His-ChuaHNL. We also analyzed the V_{\max} value of the recombinant enzymes and revealed that the V_{\max} value of the wild type His-ChuaHNL was higher than that of three deglycosylated His-ChuaHNLs (Table 3-3), but lower than that of native ChuaHNL from *C. hualienensis*³⁾. These results indicated that the sugar chain molecules also affect to the enzyme activity. The molecular masses of wild type His-ChuaHNL and the N99Q enzyme were 30–31.3 kDa, whereas that of the mutant enzymes, N109Q and N123Q was 27.4–29.1 kDa (Figure 3-4). Thus, the sugar contents of the recombinant enzymes were higher than those of native ChuaHNL from *C. hualienensis* (molecular mass of native ChuaHNL was 25 kDa on SDS-PAGE and the calculated of molecular mass of His-ChuaHNL was 20.95 kDa). The native ChuaHNL might reflect specific glycan compositions of millipede enzymes, differing from those expressed in *P. pastoris* that are often N-glycosylated with high proportion of mannoside⁴⁰⁾.

In the analysis of the substrate specificity of deglycosylated His-ChuaHNLs, it was revealed that the deglycosylated enzymes also catalyzed cyanohydrin synthesis reactions. However, the reaction velocities were remarkably different according to the deglycosylation sites. The His-ChuaHNL-N99Q and His-ChuaHNL-N109Q enzymes exhibited lower activity toward all the substrates tested as compared to His-ChuaHNL-WT, whereas the His-N123Q enzyme exhibited a higher activity toward most of the substrates tested as compared to His-ChuaHNL-WT (Table 3-2). Thus, the sugar chains at N99 and N109 exhibited an activating effect and that at N123 exhibited an inhibitory effect in cyanohydrin synthesis. The kinetic properties for benzaldehyde in the (*R*)-MAN synthesis reaction indicated that sugar at the three positions did not affect the affinity for benzaldehyde, but remarkably affected the catalytic efficiency of MAN synthesis (Table 3-3). The sugar chains at three positions also affected to the enzyme stability (Figure 3-5). The sugar chains at

N99 contributed to the pH stability, but not thermal or solvent stability. In addition, the sugar chains at N109 and N123 contributed to thermal stability and solvent stability. Thus, the synthesis of (*R*)-MAN in the solvent system was reduced remarkably by using the N109Q or N123Q enzyme. Taken together, the data indicate that the sugar chains affect to not only His-ChuaHNL production but also HNL activity and stability. Since our results suggested that S-S formation and the binding position of a sugar chain and its structure are important for efficient use of ChuaHNL, the analysis of the ChuaHNL crystal structure is now underway in our laboratory.

CHAPTER III Occurrence and enzymology of hydroxynitrile lyases of cyanogenic millipedes

Section II: Discovery and structural basis to improve the enzyme activity and enantioselectivity of hydroxynitrile lyase from *Parafontaria laminata* millipedes for (*R*)-2-chloromandelonitrile synthesis

An outbreak of *Parafontaria laminata* millipedes (Diplopoda: Xystodesmidae) has occurred during autumn in the mountainous area of central Japan every eight years.^{88, 89)} These millipedes belong to the order Polydesmida and are known as cyanogenic millipedes, which have a similar but species-specific mixture of defense compounds.⁷⁾ We expected that these millipedes could be a valuable source for a new HNL. In this study, we purified the HNL (designated as PlamHNL) from *P. laminata* and identified the gene from the mRNA by using the rapid amplification of cDNA ends (RACE) method. The cDNA gene for PlamHNLs was expressed in *Escherichia coli* and *Pichia pastoris*, and the enzymes were purified and their catalytic properties, glycosylation, and enzyme stability were characterized to clarify the effect of glycosylation on the enzyme produced in *P. pastoris*. The substrate specificity of wild type PlamHNL showed high enantioselectivity in the synthesis of various cyanohydrins, including 2-chloromandelonitrile (76% *e.e.*), a key intermediate for the anti-thrombotic agent clopidogrel. This characteristic is unique to other millipede HNLs that do not act on the compound. Thus, PlamHNL was chosen as a target for engineering to yield (*R*)-2-chloromandelonitrile with much higher enantiomeric excess using a computer-aided substrate-docking simulation in the active site of PlamHNL since an enantiomeric excess (*e.e.*) higher than 98% of the product was expected in the enzymatic process.

MATERIALS AND METHODS

Millipedes collection

Living *Parafontaria laminata* millipedes, which are often found individually under disintegrating leaves or embedded in the soil after heavy rain, were manually collected in Takaoka Castle Park, Toyama, Japan from June to August 2015 and subsequently reared in the laboratory on leaf litter collected from the site under a natural photoperiod with high humidity (ca. 100% RH) at 20 °C.

Extraction and purification of hydroxynitrile lyase from millipedes

Eight hundred living millipedes were extracted by hand shaking to induce the compounds and HNL for 2 min at room temperature before adding 10 mM phosphate-buffered saline (PBS) pH 7.4 buffer containing 0.133M NaCl. After confirmation of the smell like benzaldehyde, the crude extract was separated, and the supernatant was collected by centrifugation at $8000 \times g$ for 10 min at 4 °C to remove the insoluble material. Ammonium sulfate was added to the supernatant at 30% saturated concentration and loaded onto a Butyl-Toyopearl 650M (Tosoh) column chromatography (10×2.8 cm). The enzyme was eluted with a stepwise gradient of 10 mM PBS buffer (pH 7.4) containing 30% saturated ammonium sulfate (30%, 30 to 20%, 20 to 10%, and 10 to 0%, respectively, with 50 mL in each container) in the same buffer. The fractions containing the highest activity were pooled, dialyzed, concentrated, and loaded onto a Mono Q 5/50 GL anion exchange column (GE Healthcare) and eluted with a linear gradient of sodium chloride from 0 mM to 500 mM in the same buffer at a flow rate of 0.5 mL/min. The active fractions were pooled, and ammonium sulfate was added up to 30% for application to the Resource Phenyl HR column (GE Healthcare) and eluted with a linear gradient from 30% to 0% saturated ammonium sulfate in the same buffer at a flow rate of 0.5 mL/min. All purification steps were performed at 4 °C.

Protein sequencing

The internal amino acid sequencing was performed as previously described.⁴⁾ Briefly, the purified PlamHNL-N was separated using 15% SDS-PAGE and stained with Coomassie bright blue. The visible band was excised and combined for in-gel trypsin digestion (sequencing-grade modified trypsin; Promega, Madison, WI, USA). The internal peptides were analyzed using a nanoflow liquid chromatography-tandem mass spectrometry system. The digested peptides were loaded onto a nanoACQUITY UPLC Symmetry C18 Trap Column (particle size: 5 μm ; 180 μm i.d. \times 20 mm, Waters, Milford, MA, USA) and an ACQUITY UPLC Peptide CSH C18 nanoACQUITY Column (particle size: 1.7 μm ; 75 μm i.d. \times 200 mm, Waters), respectively. Mobile phases A and B were water containing 0.1% formic acid and acetonitrile containing 0.1% formic acid, respectively. The elution profile was a linear gradient of B (10%-80%) at a flow rate of 300 nL/min for 90 min. Mass spectrometry was simultaneously performed in positive electrospray ionization mode with a capillary voltage of 3 kV and cone voltage of 40 V using a SYNAPT G2-Si high-resolution

quadrupole time-of-flight mass spectrometer (Waters). [Glu1]-fibrinopeptide B was used as a mass calibrant. The MS/MS spectra were analyzed using BioLynx (Waters).

cDNA cloning of PlamHNL

Total RNA was prepared from dissected *P. laminata* paraterga immediately homogenized with TRIzol reagent (Thermo Fisher Scientific). First-strand cDNA was synthesized directly from total RNA by reverse transcription using a SMART RACE cDNA amplification kit (Clontech Laboratories) and SuperScript III (Life Technologies) as reverse transcriptases. Degenerate primers corresponding to the conserved regions were designed by our laboratory to obtain the internal peptide of HNL fragments with the following 1F and 1R primers (Appendix XIX.). The reaction mixture of PCR contained 0.2 μ L Takara ExTaq (5 unit/ μ L), 5 μ L of 10X ExTaq buffer, 4 μ L of dNTP mixture (2.5 mM each), 5 μ L of forward and reverse primers (10 pmol/ μ L), 1 μ L of the first-strand cDNA (50-100 ng/ μ L), and 29.8 μ L of sterilized water. The PCR program was used for amplification of the genes as follows: 98 °C for 1 min, followed by 30 cycles of 98 °C for 10 s, 45 °C for 30 s, 72 °C for 1 min, and finally 5 min at 72 °C prior to cooling to 4 °C. The HNL fragments were extended in both the 5' and 3' directions using gene-specific primers (primers 2-7; (Appendix XIX.) and the cDNA library as a PCR template. All resulting PCR were separated by agarose gel electrophoresis, purified with NucleoSpin (Macherey-Nagel, Germany), and ligated into a T-vector pMD20 (Clontech Laboratories) using a DNA Ligation Kit (Mighty Mix, Takara). DNA sequences were determined using a 3500 Genetic Analyzer (Applied Biosystems, Foster City, CA, USA). The obtained sequences were assembled and analyzed using ATGC and Genetyx Ver. 12 (Genetyx, Tokyo, Japan).

Construction of expression vectors of recombinant PlamHNL in E. coli and P. pastoris

The recombinant strain *E. coli* SHuffle T7 Express (New England Biolabs, Ipswich, MA, USA) was used as an expression host and constructed in the pET28a (+) vector (Novagen; Darmstadt, Germany) as an expression vector. The DNA insert was PCR-amplified using Primer 8F and 8R gene specific primers (Appendix XIX) and PlamHNL cDNA as template DNA. The reaction mixture of PCR contained 0.2 μ L Takara ExTaq (5 unit/ μ L), 5 μ L of 10X ExTaq buffer, 4 μ L of dNTP mixture (2.5 mM each), 5 μ L of forward and reverse primers (10 pmol/ μ L), 1 μ L of the first-strand cDNA (50-100 ng/ μ L), and 29.8 μ L of sterilized water. The PCR program was used for amplification of the genes as follows: 98 °C for 1 min, followed by 30 cycles of 98 °C for 10 s,

55 °C for 30 s, 72 °C for 1 min, and finally 5 min at 72 °C prior to cooling to 4 °C. PCR products were further digested with *Bam*HI and *Hind*III restriction enzymes and gel purification. The purified insert DNA was ligated to the corresponding site of the pET28a (+) vector using a DNA Ligation Kit (Mighty Mix, Takara). The constructed vector was designated as pET28a-PlamHNL.

To express the recombinant gene for PlamHNL in *P. pastoris* GS115 cells, a gene insert was amplified using Tks Gflex DNA polymerase and the gene-specific primers 9F and 9R (Appendix XIX.). Meanwhile, pPICZ α A was linearized by PCR amplification using an inverted PCR method (KOD-Plus-Mutagenesis Kit; TOYOBO). After PCR purification with NucleoSpin (Macherey-Nagel, Germany), both amplified products were ligated with an In-Fusion HD Cloning Kit (Clontech Takara Bio USA, Inc.), generating pPICZ α A-PlamHNL. Then, the coding sequence of 6x-His tag at the N-terminus and TEV protease (primers 10 and 11) recognition sequence was introduced by a second inverted PCR method designated as pPICZ α A-TEV-His PlamHNL. The constructed vector pPICZ α A-TEV-His PlamHNL was linearized by digestion with *Sac*I and electroporated into *P. pastoris* GS115 cells using an Easysselect Pichia Expression Kit (Thermo Fisher Scientific) according to the manufacturer's instructions.

Expression and production of recombinant PlamHNL in E. coli and P. pastoris

A single colony of *E. coli* SHuffle T7 Express harboring pET28a-PlamHNL was inoculated into 5 mL of Luria-Bertani (LB) broth containing kanamycin (80 μ g/mL) and incubated overnight at a shaking rate of 300 rpm at 30 °C. The starter culture (5 mL) was transferred to LB medium (500 mL) in 2 L Erlenmeyer flasks containing kanamycin (80 μ g/mL) and cultured at a shaking rate of 150 rpm at 30 °C. After 12 h, isopropyl- β -d-thiogalactopyranoside (IPTG) was added to a final concentration of 0.5 mM, and cells were cultured at 18 °C for 24 h with the same shaking rate. The cells were centrifuged (8500 \times g for 15 min) and resuspended in potassium phosphate buffer (KPB; 20 mM, pH 7.0) containing sodium chloride (0.5 M) and imidazole (25 mM). The resuspended cells were lysed by sonication and the lysate was centrifuged (15000 \times g for 15 min at 4 °C) to remove debris. The supernatant was loaded onto Ni Sepharose 6 Fast Flow (GE Healthcare, Little Chalfont, UK) columns (i.d. 25 mm; column volume, 20 mL), washed with 50 mM imidazole, and eluted with a linear gradient of 50–300 mM imidazole in KPB (20 mM, pH 7.0) containing sodium chloride (0.5 M) at a flow rate of 0.5 mL \cdot min⁻¹. The enzyme activity was measured by the method described above, and the active fractions were pooled, dialyzed, concentrated, and checked for purity by SDS-PAGE.³¹⁾

P. pastoris transformants were inoculated into 5 mL YPD (1% yeast extract, 2% peptone, and 2% dextrose) broth containing Zeocin ($100 \mu\text{g}\cdot\text{mL}^{-1}$). After overnight pre-culture at 30 °C with shaking at 300 rpm, the inoculum was transferred to buffered minimal glycerol medium (BMGH; 100-mM potassium phosphate buffer (pH 6.0), 1.34% yeast nitrogen base without amino acid, $4 \times 10^{-5}\%$ biotin, 0.004% histidine, and 1.0% glycerol) and grown at 30 °C with shaking at 200 rpm. After 48 h of incubation, cells were harvested by centrifugation and resuspended in 500 mL expression medium, buffered minimal methanol medium (BMMH; the same as BMGH but 0.5% (v/v) methanol was added instead of 1% (v/v) glycerol), and 0.5% methanol as an inducer every 24 h. As the maximum extracellular HNL activity was observed after 6 days of cultivation, the supernatant was harvested by centrifugation at $3000 \times g$ for 10 min and concentrated 25-fold with hollow fiber ultrafilter (Microza; Asahi KASEI Chemical). Ammonium sulfate (30% w/v) was added to the supernatant and then directly applied to a Toyopearl Butyl-650-M column (Tosoh, 25 mm i.d., column volume 30 mL) previously equilibrated with equilibration buffer (20 mM KPB (pH 7.0) and ammonium sulfate (30% w/v)). The absorbed enzymes were washed with 10 volumes of equilibration buffer and eluted with a gradient of ammonium sulfate from 30% to 0% in the same buffer, and the active fraction was dialyzed against 20 mM KPB (pH 7.0). All purification steps were performed at 4 °C.

Crystallization and structure determination

After cleavage by thrombin protease (GE Healthcare) to remove His-tagged protein, the recombinant PlamHNL-E was purified to homology using a HisTrap FF column (GE Healthcare). The purified PlamHNL-E without His-tag was dialyzed against KPB (10 mM, pH 7.5) using PD-10 (GE Healthcare) and concentrated to 10 mg/mL using an Amicon Ultra Centrifugal Filter Unites NMWL, 10 kDa (Merck Millipore, Billerica, MA, USA). The concentrated PlamHNL was used for crystal screening with commercial screening index HT (Hampton Research), PACT Suite (Qiagen), and Clear Strategy Screen I (Molecular Dimensions) using the vapor diffusion sitting drop method at 20 °C in 96-well Intelli-Plates (Art Robbins instruments, CA, USA). Crystals appeared within one day in Clear Strategy Screen I under the following conditions: 0.2 M potassium thiocyanate, 0.1 M sodium acetate pH 5.5, 10% w/v polyethylene glycol (PEG) 8000, and 10% w/v PEG 1000.

Finally, the best condition was obtained. Ligand-free crystals were obtained using the hanging drop vapor diffusion method at 20 °C for one day with a series of precipitant solutions: 0.1 M Bis-tris-HCl (pH 6.0–7.0), 0.15 M potassium thiocyanate, 8%–11% PEG 8000, and 8% PEG

1500. Droplets for crystallization were prepared by mixing 1 μL of PlamHNL solution and 0.5 μL of precipitant solution, and droplets were equilibrated against 500 μL precipitant solution. The crystallization conditions for the benzaldehyde-bound and 2-chlorobenzaldehyde-bound PlamHNL were slightly different. An aliquot (6 μL) of the PlamHNL solution was added to each ligand solution (3 μL of 10 mM Bis-tris-HCl pH 6.5 containing 50 mM benzaldehyde or 2-chlorobenzaldehyde). These solutions, PlamHNL/benzaldehyde or 2-chlorobenzaldehyde, were stored overnight at 4 $^{\circ}\text{C}$ and allowed to crystallize by the hanging drop vapor diffusion method using a series of precipitant solutions: 0.1 M Bis-tris-HCl (pH 6.0–6.5), 0.15 M potassium thiocyanate, 10% PEG 8000, and 8% PEG 1500 containing 50 mM benzaldehyde or 2-chlorobenzaldehyde. Droplets for crystallization were prepared by mixing 1.5 μL of complex solution and 0.5 μL of precipitant solution, and droplets were equilibrated against 500 μL precipitant solution. The ligand-free crystal soaked in each ligand (benzaldehyde, 2-chlorobenzaldehyde, or mandelonitrile) solution was also examined. The crystals were soaked in several solutions.

Prior to diffraction data collection, crystals were soaked in cryoprotectant solutions containing 25% glycerol and other reagents. Diffraction data sets were collected at 100 K at beamline BL-5A of KEK-PF (Tsukuba, Japan). Reflections were recorded with an oscillation range per image of 0.5° . Diffraction data were indexed, integrated, and scaled using programs XDS,⁹⁰ Pointless,⁹¹ and Scala.⁹¹ The structures of ligand-free and ligand-bound ones were solved by molecular replacement with program Molrep⁹² using coordinates of OgraHNL as an initial model, and finalized sets of coordinates were obtained after iterative rounds of model modification with the program Coot⁹³ and refinement with Refmac5.⁹⁴ Co-crystallization with 2-chlorobenzaldehyde and any soaking method was not able to provide good electron density for the ligands. A summary of the diffraction data and refinement statistics is shown in Appendix XII.

Docking simulation using Molecular Operating Environment (MOE) program

To predict the structure of PlamHNL complexed with (*R*)-2-chloromandelonitrile, we used the structure of OgraHNL complexed with (*R*)-2-chloromandelonitrile as a template for a docking simulation using MOE (Chemical Computing Group, Montreal, Canada). Homology modeling of PlamHNL complexed with (*R*)-2-chloromandelonitrile was computed using the Alanine and Residue Scanning functionality around 14 amino acid residues in the binding pocket to identify

critical residues for affinity by other 19 amino acids. Mutants that showed the lowest affinity were selected for experimental verification.

Site-directed mutagenesis

The PlamHNL mutant was prepared by site-directed mutagenesis using the QuikChange Site-Directed Mutagenesis Kit with forward and reverse primers (Appendix XIX) and pET28aPlamHNL as a template. The PCR reaction was performed for 18 cycles: denaturing (95 °C, 20 s; first cycle: 95 °C, 2 min), annealing at 52 °C for 20 s. PCR products were treated with *DpnI* (10 U) at 37 °C for 1 h and transformed into *E. coli* SHuffle T7.

RESULTS AND DISCUSSION

Purification of PlamHNL from P. laminata millipedes

In nature, millipedes excrete defensive compounds when they are near enemies. Cyanogenesis proceeds by mixing (*R*)-mandelonitrile in the reservoir of defensive glands with HNL, resulting in external benzaldehyde and hydrogen cyanide secretion through ozopores as defensive compounds.⁹⁵⁻⁹⁷⁾ The crude extract from *P. laminata* showed a specific activity of 40 U/mg and was purified to 33-fold with a 4.7% recovery yield, showing a specific activity equivalent to 1320 U/mg protein toward benzaldehyde as a model substrate in a cyanohydrin synthesis reaction (Appendix X). The native molecular mass of purified PlamHNL-N was determined to be 43 kDa using a gel filtration column. The enzyme gave a single band in denatured and reduced form on SDS-PAGE (Appendix XIII) with a calculated molecular mass of 21 kDa. These results indicate that PlamHNL-N is likely to be active as a homodimer. The purified enzyme was stained with the periodic acid-Schiff reagent (PAS) and showed a pink color, indicating that the PlamHNL from millipedes was glycosylated (Appendix XIII). In a previous report, we successfully purified the ChuaHNL from several kilograms of frozen *C. hualienensis* millipedes through a combination of column chromatography and column chromatography. ChuaHNL was characterized to have a molecular mass of 47.3 kDa with glycosylational modification and nearly six times higher specific activity (7420 U/mg protein) than PlamHNL-N.⁹⁷⁾ The specific activity of PlamHNL is almost equal to that of almond (*Prunus amygdalus*)⁹⁸⁾ but higher than those of known HNLs from plants, such as

rubber (*Hevea brasiliensis*),⁹⁹⁾ yellow passion fruit (*Passiflora edulis*),¹⁹⁾ loquats (*Eriobotrya japonica*)⁶⁰⁾ and sorghum (*Sorghum bicolor*).¹⁰⁰⁾

cDNA cloning and nucleotide sequencing

Since the transcription of the gene for HNL and the accumulation of the enzyme were localized in defensive secretory glands of the paraterga,⁹⁷⁾ the author cloned the full-length cDNA encoding PlamHNL from the paraterga of an adult millipede by rapid amplification of cDNA ends (RACE). Degenerate primers corresponding to conserved regions were designed by our laboratory⁴⁾ and based on the internal amino acid sequence to obtain the internal peptide of the HNL fragment. The 552 bp cDNA sequence encoded 183 amino acid residues, including a 20 amino acid-long signal peptide. The predicted internal peptide sequences from the cDNA were identical to those of purified PlamHNL from millipedes (Appendix XIV). The predicted mature protein (no signal peptide) had a molecular mass of 20,855 Da and an isoelectric point of 5.13. Recently, we discovered and identified HNLs from 10 cyanogenic millipedes that could be categorized into two families.⁴⁾ As shown by the phylogenetic analysis based on the amino acid sequence of known HNLs from millipedes (Appendix XV), PlamHNL belongs to the Xystodesmidea family and was shown to have more than 80% identity with HNLs from *Parafontaria falcifera* (PflaHNL), *Par. tonominea* species complex 1, 2, and 3 (Pton1HNL, Pton2HNL, and Pton3HNL, respectively), and *Par. tokaiensis* (PtokHNL) as well as 67% identity with HNLs from *Riukiaria semicircularis* (RssHNL) and *Riukiaria* sp. (RspHNL) in the same family. The other family is Paradoxosomatidae, which exhibited 49%, 49%, 48%, and 49% identity with NtmHNL (*Nedyopus tambanus mangesinus*), NttHNL (*Nedyopus tambanus tambanus*), OgraHNL (*Oxidus gracilis*), and ChuaHNL, respectively. PlamHNL was clustered together in *Parafontaria* genus, indicating that genes encoding HNL likely evolved from one ancestral gene during the evolution of polydesmoid millipedes. Moreover, the predicted mature PlamHNL had one potential *N*-glycosylation site on the asparagine residue at 69.

Expression of PlamHNL gene in E. coli and P. pastoris

Although the HNL from *P. laminata* millipedes (PlamHNL-N) has a high specific activity and shows high potential as a biocatalyst, it is impractical to purify the enzyme from millipedes for applications. Molecular cloning of the gene for the enzyme and its expression in microorganisms, such as *E. coli*, enables an easier supply of the enzyme. In our previous report, HNLs from

millipedes have characteristic eight conserved Cys residues that functionally form disulfide bonds that play important roles in the folding, activity, and stability of the millipede HNLs.⁴⁾ Therefore, it was considered that the gene for the HNLs from millipedes could be better expressed in *E. coli* SHuffle T7 because they express disulfide bond isomerase DsbC, which promotes the correction of mis-oxidized proteins into their correct form.⁴⁾ Therefore, in this study, the author chose *E. coli* SHuffle T7 (PlamHNL-E) and *P. pastoris* GS115 (PlamHNL-P) as expression hosts for a comparison of the effect of glycosylation on HNL produced by two different systems. As expected, the genes for PlamHNL were successfully expressed in the two hosts as summarized in Appendix X. However, the expression level of PlamHNL in the *E. coli* host is low (1.75 mg/L culture), most of target protein is an insoluble form in the inclusion bodies. The soluble part of PlamHNL-E was easily purified by affinity chromatography using Ni-Sepharose fast flow, and the final step resulted in 115-fold purification with 55.7 % recovery. The specific activity of purified PlamHNL-E was 2,190 U/mg, which is slightly higher than that of PlamHNL-P. The production of PlamHNL-E (6,890 U/L) in *E. coli* was 20 times higher than that in PlamHNL-P (351 U/L) produced in *P. pastoris*. PlamHNL-E and PlamHNL-P had molecular masses of 20,000 Da and 24,000 Da, respectively (Appendix XIII-A). The molecular masses varied depending on the expression host, because the PlamHNL-P was glycosylated by cellular organelles. In contrast, PlamHNL-E from prokaryotic cells lacked such modifications, but it was folded and showed activity similar to those produced in eukaryotic cells.

Glycosylation analysis

To determine the glycosylation and non-glycosylation profiles of the wild type and recombinant PlamHNLs, the purified enzymes were stained with the periodic acid-Schiff reagent (PAS). Periodic acid oxidizes the sugars to aldehydes, which then condense with the Schiff reaction to form a pink color.¹⁰¹⁾ The PlamHNL-N and PlamHNL-P both displayed a pink color, indicating that the PlamHNL from both systems was glycosylated. In particular the enzyme expressed in *P. pastoris* showed a strong pink color, which may have been due to high sugar content. In contrast, the PlamHNL-E was not glycosylated as shown in Appendix XIII-B. On the other hand, the band of PlamHNL-P appeared as a broad and dark band, and the molecular mass was reduced after deglycosylation with PNGase or Endo H. The PNGase F cleaved between the innermost *N*-acetylglucosamine and Asn while Endo H hydrolyzed the glycosidic linkage between two *N*-acetylglucosamine residues. This result indicates that PlamHNL-P is an *N*-glycosylation type. The

PlamHNL-P digested with both enzymes dramatically reduced the molecular mass from 24,000 Da to 20,000 Da, which is the same as that of non-glycosylated PlamHNL-E (Appendix XIII-C).

Effect of pH, temperature, and metal ions on the activity and stability of recombinant PlamHNLs

The effect of pH and temperature on the activities of both purified recombinant PlamHNLs were compared with the same enzyme concentration (15 U/mL). The optimal pH optimum for (*R*)-mandelonitrile synthesis shown by PlamHNL-E and PlamHNL-P was 4.0, which exhibited the maximum initial activity (120-130 $\mu\text{mol/h}$) with a high enantiomeric excess of 99-100% of the product (Figure 3-6A and -B). However, the activity was significantly decreased at pH levels higher than 5.0 due to a rapidly increasing non-enzymatic reaction yielding the racemic mixture containing both of the enantiomers (Figure 3-6A and -B). To suppress the non-enzymatic reaction, a low pH condition is needed for chiral cyanohydrin synthesis. In the pH range of 2.5-3.5, PlamHNL-P was more stable than the one expressed in *E. coli* (55-60% retaining activity), displaying 80% remaining activity after 1 h incubation time. PlamHNL-P showed stable activity compared to PlamHNL-E in the pH range of 4.0-5.5. (Figure 3-6C). The effect of pH on the stability of PlamHNL-P may have been caused by post-translational glycosylation of the enzyme that is lacking in the *E. coli* system. This result corresponds to the recombinant HNL from almonds (*PaHNL5*) expressed in *P. pastoris* with the glycosylated enzyme by one glycosylation site and plays a role in the stability at low pH conditions.⁵⁰⁾

At the optimal pH (pH 4.0), the ideal temperature of both recombinant PlamHNLs was estimated to be 25-30 $^{\circ}\text{C}$, which resulted in initial activity for (*R*)-mandelonitrile synthesis and enantiomeric excess up to 120-140 $\mu\text{mol/h}$ and 99-100%, respectively (Figure 3-6D and -E). PlamHNL-P exhibited a slightly higher initial velocity than PlamHNL-E. At an elevated temperature from 35 to 50 $^{\circ}\text{C}$, the collision between enzyme and substrate molecules would likely be increased to form enzyme-substrate complexes, resulting in an increase in the initial synthesis velocity up to 200-240 $\mu\text{mol/h}$. However, it also accelerates the non-enzymatic reaction, which lowers the enantiomeric purity of the product below 90% *e.e.* (Figure 3-6E). Thus, 30 $^{\circ}\text{C}$ was selected as the optimum temperature for (*R*)-mandelonitrile synthesis, which enabled the yield of products with the highest % *e.e.*

To evaluate the thermostability of recombinant PlamHNLs, both enzymes were incubated at a temperature of 30 to 100 $^{\circ}\text{C}$ for 1 h. All of the PlamHNLs showed 100% remaining activity at a temperature of 30 to 55 $^{\circ}\text{C}$ after 1 h of preincubation. In direct comparisons at 80 $^{\circ}\text{C}$, PlamHNL-P

exhibited 100% remaining activity and displayed more stability than PlamHNL-E, which lost 50% of its activity under these conditions. The properties of HNL from *P. edulis*, a plant enzyme produced in *P. pastoris*, showed similar properties with the expressed HNL being more thermostable than that of *E. coli* because of the glycosylation.²⁰ However, PlamHNL-E was quite stable at temperatures up to 70 °C because it has Cys residues, which are known to make the protein more stable by forming disulfide bond(s) (Figure 3-6F).¹⁰²

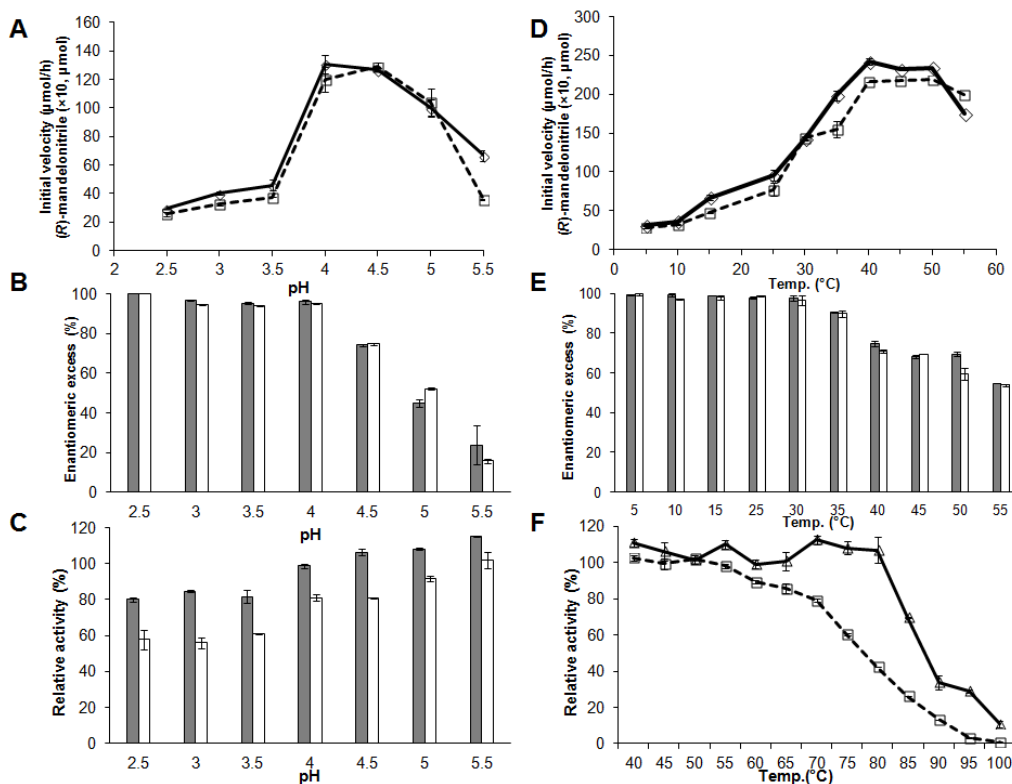


Figure 3-6. Effect of pH on initial velocity (A), enantiomeric excess (B), and enzyme stability (C). Effect of temperature on initial velocity (D), enantiomeric excess (E), and enzyme stability (F). The reaction was performed in citrate buffer (400 mM, pH 4.0), benzaldehyde (50 mM), and KCN (100 mM). Solid and gray bars indicate PlamHNL-P. The dashed line and open bar indicate PlamHNL-E.

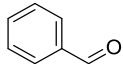
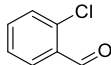
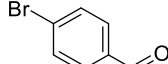
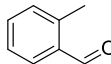
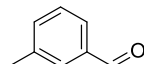
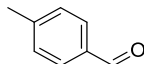
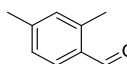
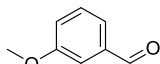
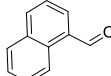
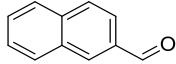
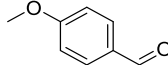
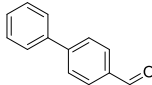
The effects of ions, inhibitors, and various additives on the recombinant PlamHNLS were determined. The metal ion or inhibitor sensitivity of non-glycosylated PlamHNL-E was not significantly different from that of glycosylated PlamHNL-P. The (*R*)-mandelonitrile synthesis activity of recombinant PlamHNLS was slightly inhibited by 1 mM of MgSO₄, MnSO₄, ZnSO₄, CdCl₂, CuSO₄, KCl, FeCl₃, LiCl, CoCl₂, and NaF caused 5-30% inhibition (Appendix XI). The activity of PlamHNL-E and PlamHNL-P was dramatically inhibited with cysteine- or histidine-modifying reagents, such as iodoacetamide (26.5% and 30.2% remaining activity, respectively),

iodoacetic acid (18.3% and 21.3% remaining activity, respectively), and diethylpyrocarbonate (68.4% and 62.5% remaining activity, respectively). These results indicate that cysteine and histidine are required for PlamHNL activity to be involved in the catalysis mechanism. Similar results to the cysteine- or histidine-modifying reagents have been observed in HNLs from *C. hualienensis*,⁹⁷⁾ *Eribotry japonica*,¹⁰³⁾ *Prunus mume*,²⁸⁾ *Phlebodium aureum*,¹⁰⁴⁾ and *Ximenia Americana*.¹⁰⁵⁾

Substrate specificity and kinetic parameters

According to our previous work, we found that the millipede HNL preferred aromatic aldehydes,⁹⁷⁾ analogs of benzaldehyde, which is the degraded product of the defensive secretion compound (*R*)-mandelonitrile.⁷⁾ Aromatic aldehydes were tested as substrates in kinetic analyses of recombinant PlamHNL in the synthetic reaction with a reaction time of 5 min at 25 °C. The Michaelis-Menten constants (K_m) of the purified enzyme toward benzaldehyde, 2-chlorobenzaldehyde, 4-bromobenzaldehyde, 2-methylbenzaldehyde, 3-methylbenzaldehyde, 4-methylbenzaldehyde, 2,4-dimethylbenzaldehyde, 3-methoxybenzaldehyde, 4-biphenylcarboxaldehyde, 2-naphthaldehyde, *p*-anisaldehyde, and 1-naphthaldehyde were calculated to be in the wider range of 1.6-66 mM (Table 3-3). The enzyme showed a broad spectrum toward aromatic aldehydes with the V_{max} value in the range of 164-33,300 $\mu\text{mol min}^{-1} \text{mg}^{-1}$. The maximum V_{max} , k_{cat} , and k_{cat}/K_m values of PlamHNL toward 3-methoxy benzaldehyde were calculated to be 33,300 $\mu\text{mol min}^{-1} \text{mg}^{-1}$, 13,800 s^{-1} , and 8,680 $\text{s}^{-1} \text{mM}^{-1}$, respectively. In contrast, the K_m value toward 3-methoxy benzaldehyde was the lowest among the aromatic aldehyde derivatives tested. Although the benzene aromatic substrate specificity profile of PlamHNL was similar to ChuaHNL from *C. hualienensis*,⁹⁷⁾ PlamHNL acted on the benzaldehyde substituted at the *ortho* position much better than ChuaHNL, and the *e.e.* of the products was higher. For example, the *e.e.* values of the cyanohydrin products of PlamHNL catalysis from 2-chlorobenzaldehyde and 2-methylbenzaldehyde were 76% and 86%, respectively, which were much higher than the corresponding products synthesized by ChuaHNL with *e.e.* values of 11% and 44.9%, respectively. Interestingly, PlamHNL catalyzed the synthesis of (*R*)-2-chloromandelonitrile from 2-chlorobenzaldehyde. The compound is used as a precursor for (*R*)-2-mandelic acid, which is a key intermediate for a potent oral anticoagulant.¹⁰⁶⁾ We used this compound as a model for further study to improve the activity and enantioselectivity of PlamHNL based on structural analyses.

Table 3-3. Kinetic parameters of PlamHNL for aldehyde in synthesis reaction of cyanohydrins

Substrate	Structure	<i>e.e.</i> (%)	K_m (mM)	k_{cat} (s ⁻¹)	k_{cat}/K_m (s ⁻¹ mM ⁻¹)	V_{max} ($\mu\text{mol min}^{-1}$)
Benzaldehyde		95	2.33±0.2	694±19	298±1.9	2660±12
2-chlorobenzaldehyde		76	59.3±2.6	378±19	6.38± 2.2	909 ±35
4-Bromo benzaldehyde		99	36.7±1.2	416±25	11.4±0.4	1000±90
2-Methyl benzaldehyde		86	15.8±0.5	520±0.6	32.8±1.1	1250±67
3-Methyl benzaldehyde		94	15.1±1.1	417±7.3	27.6±2.3	1000±125
4-Methyl benzaldehyde		99	32.1±0.8	320±22	10±0.6	770±48
2,4-Dimethyl benzaldehyde		99	13.6±1.0	85±2.3	6.2±1.1	204±22
3-Methoxy benzaldehyde		91	1.6±0.1	13800±125	8680±0.5	33300±475
1-Naphthaldehyde		88	2.2±0.2	68.3±9.3	31.3±0.6	164±36
2-Naphthaldehyde		90	12.1±1.1	220±20	18.2±1.2	526±8.8
<i>p</i> -Anisaldehyde		99	66±0.8	378±14	5.7±0.1	909±63
4-Biphenylcarboxaldehyde		86	2±0.1	833±45	416±2.6	2000±5.2

Overall PlamHNL structure and reaction mechanism

The X-ray crystal structure of recombinant PlamHNL-E was obtained after removing the His-tag, and the amino acid residues were numbered from the N-terminal Met residues with the signal peptide and beginning at amino acid 21 (Lys₂₁) of the mature protein. The ligand-free PlamHNL structure was determined by the molecular replacement method with the initial model ChuaHNL (PDB ID 6JHC)¹⁰⁷ and refined at 1.42 Å resolution. The PlamHNL crystal belonged to *P2*₁, and the fold is a homodimer composed of two α -helices and eight β -sheets in each subunit connected with three intramolecular and two intermolecular disulfide bonds (Figure 3-7A). The overall structure is almost identical to that of ChuaHNL¹⁰⁷ (Figure 3-7B), which showed high similarity to the folding pattern of the lipocalins. Lipocalins are secreted proteins that occur in animals, plants, and bacteria and play roles in different functions such as retinol transport,¹⁰⁸ lipids transport,^{109, 110} cryptic colorization,¹¹¹ facilitation of feeding for hemetophagus insects,¹¹² and the enzymatic synthesis of prostaglandins (lipocalin-type prostaglandin D synthase; LPGDS).¹¹³ The structural features of the lipocalin fold can also be seen in the tertiary structure of millipede HNLs, such as a large cup-shaped cavity, which is composed of structurally conserved regions (SCRs). Their tertiary structure consists of six- or eight-stranded β -barrels stabilized by one, two, or three sulfide bridges.^{107, 114, 115} Although PlamHNL showed low identity of overall SCR motif sequences with other lipocalins (Appendix XVI-A), its tertiary structure is closely related to human retinol binding protein 4 (RBP4), walleye sandercyanin fluorescent protein, and L-PGDS (Appendix XVI-B). L-PGDS is the only enzyme among the members of the lipocalin gene family, which form two large central cavities that separate the binding pocket (surrounded by aliphatic hydrophobic side chain) and a catalytic pocket connecting with the narrow tunnel between two pockets.¹¹⁶ In the case of PlamHNL structures, the ligand-binding cavity and catalytic residues were located in the same pocket and formed by many hydrophobic aromatic residues (such as F44, Y60, A74, F87, A96, L98, W109, F111, Y124, and A126) identical to those of ChuaHNL (Figure 3-7C). This deep hydrophobic cavity might facilitate the acceptance of various bulky cyanohydrins as observed previously.¹⁰⁷ There were two distinct corresponding amino acid residues between PlamHNL and ChuaHNL, one in the entrance region (I89 in PlamHNL and G90 in ChuaHNL) and another at the bottom of the ligand cavity (N85 in PlamHNL and Y86 in ChuaHNL) on the β_4 -barrel (Figure 3-7C). The longer aliphatic side chain I89 of PlamHNL made the binding cavity deeper than that of ChuaHNL (G90), and the lower part at N85 of PlamHNL created a small tunnel (Figure 3-7D). PlamHNL was different from ChuaHNL in that it formed a more compact pocket (Figure 3-7E).

This observation seems to be important to the effectiveness of the shape of the binding cavities and might be reflected in the substrate specificity of each millipede HNL.

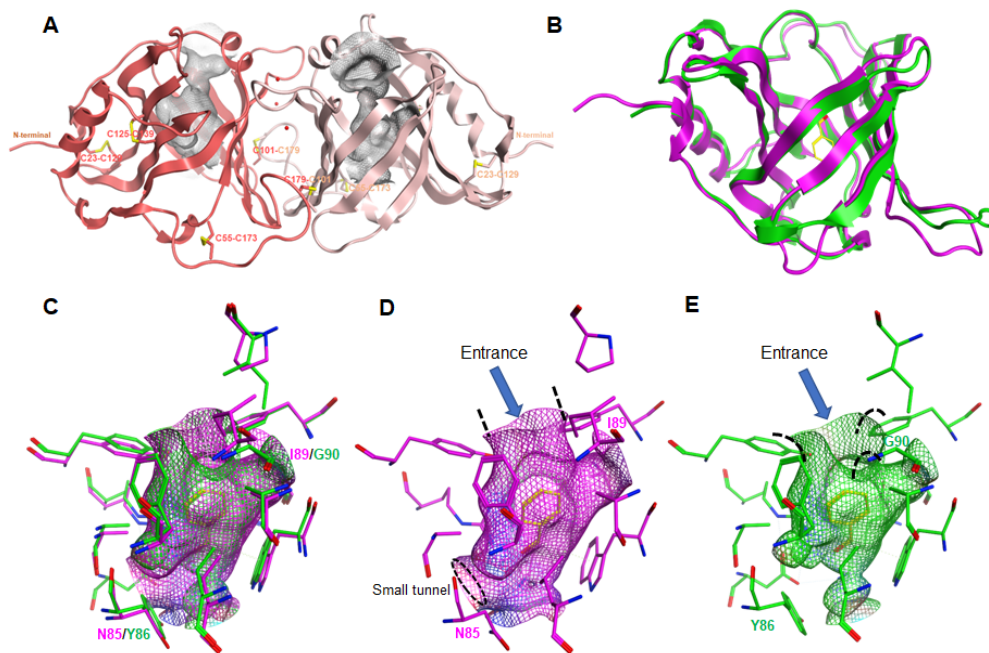


Figure 3-7. Crystal structure of PlamHNL. A) Overall structure of the PlamHNL dimer. B) Superimposed structure of PlamHNL (magenta) and ChuaHNL (green). C) Superimposed side chain and surface of the active pocket of PlamHNL (magenta) and ChuaHNL (green). D) and E) Stereo views of the side chain-forming cavity and entrance channel of PlamHNL and ChuaHNL, respectively.

The residues important for catalysis have been confirmed in our laboratory by mutagenesis and kinetic analyses.¹⁰⁷⁾ Furthermore, since we found that the residues (R58, Y124, Y60, D76, and K138) are conserved in all of the known HNLs from millipedes (Figure 3-8A), we propose that the mechanism of PlamHNL is identical. We successfully determined the structure of the complex of PlamHNL with benzaldehyde (Figure 3-8B) and benzaldehyde with thiocyanate (by co-crystallization with benzaldehyde in a buffer containing potassium thiocyanate (0.2 M), sodium acetate (0.1 M, pH 5.5), polyethylene glycol (PEG) 8000 (10 % w/v), and PEG 1000 (10 % w/v)) (Figure 3-8C). Although we performed many soaking experiments with various substrates, such as *rac*-mandelonitrile and *rac*-2-chloromandelonitrile, we could not obtain a structure of substrate-

bound PlamHNL. In the ligand-free form, two water molecules were observed in the active site. In the complex with benzaldehyde (Figure 3-8B), the polar part of the carbonyl group interacted with the epsilon-amino group of K138 (distance 2.71 Å), and one water formed a hydrogen bond with the OH-group of Y124 (2.79 Å) and the guanidine group of R58 (3.0 Å). In the complex with thiocyanate, which occupies the water position, the anion thiocyanate group interacted with the guanidinium group of R58 (Figure 3-8C). This observation of the negatively charged thiocyanate complex appears to be analogous to the cyanide source of mandelonitrile synthesis in PlamHNL, though it could not be utilized for the condensation of benzaldehyde to yield mandelonitrile. It is reasonable to speculate that the hydrogen bond length in the R58 and Y124 dyad in this complex might increase the basicity of R58 to deprotonate the proton of HCN. Subsequently, the negatively charged carbon of cyanide ion acts as a nucleophile and attacks the carbonyl carbon of benzaldehyde. The carbonyl oxygen receives an electron and a proton from cyanide and R58, respectively, to release (*R*)-mandelonitrile out of the active site of PlamHNL (Figure S6). Motojima *et al.* reported that the proposed catalytic mechanism of millipede HNLs, in which R58 and K138 are conserved and directly involved in the catalysis reaction, has been considered to be most likely by docking simulation with (*R*)-mandelonitrile and site-directed mutagenesis with results showing that the enzyme activities of mutations on these sites caused a complete loss of activity.¹⁰⁷⁾

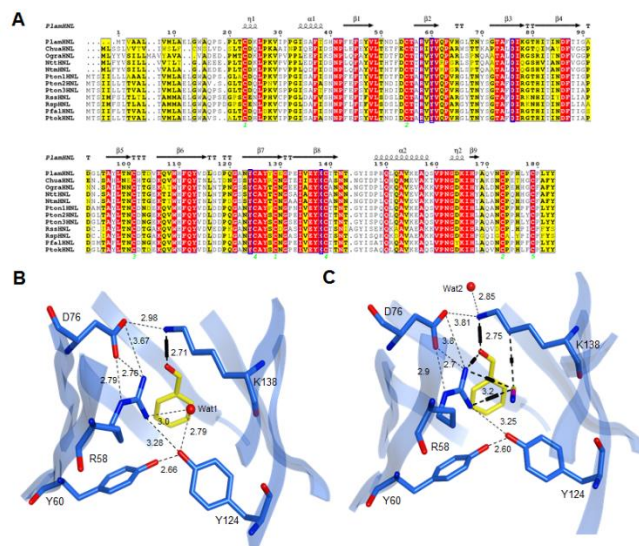


Figure 3-8. A) Amino acid sequence and secondary-based alignment of HNLs from millipedes. The secondary structure elements are shown as follows: α -helices, medium squiggles with α symbols; 3_{10} -helices, squiggles with η symbols; β -sheets, arrows with β symbols; and strict β -turns, TT letters. The pairs of cysteine residues forming disulfide-bonds are shown as green digits at the bottom of the alignment. The red background shows strictly conserved residues. Similar

groups are highlighted with a yellow background. Residues conserved between groups are boxed. The proposed catalytic residues are surrounded by blue boxes. Numbers refer to the PlamHNL sequence. The alignment was performed using PROMALS3D¹¹⁷) and illustrated using ESPript 3.0 (<http://espript.ibcp.fr>).¹¹⁸) B) Stereo view of the active site structure of PlamHNL complexed with benzaldehyde (yellow) and C) benzaldehyde and thiocyanate (magenta). The catalytic residues (R58, Y60, D76, Y124, and K138) are shown in blue. The hydrogen bonds are shown as black dotted lines, and their distances are labeled. Water molecules are depicted as red spheres (wat1, wat2).

Structure-guided engineering to improve the activity toward 2-(R)-chlorobenzaldehyde

To improve the catalytic activity and enantioselectivity of (*R*)-2-chloromandelonitrile synthesis of PlamHNL, computational approach was used for enzyme engineering. The (*R*)-2-chloromandelonitrile was docked into the substrate-binding site of PlamHNL crystal structure (PDB ID: 7BOW) using Molecular Operating Environment program (MOE). Based on the catalytic mechanism (Appendix XVII), the model of the wild type giving the most prominent affinity score ($\Delta S = -5.32$ kcal/mol) is shown in Figure 3-9A. The hydroxyl of (*R*)-2-chloromandelonitrile interact with R58 and K138 with bond lengths of 2.45 and 2.3 Å, respectively. The nitrile group form hydrogen bond with Y124 with bond lengths of 2.9 Å. This simulation of ligand was used to generate the mutants using the MOE program in the function of alanine and residue scanning around 14 amino acid residues at the binding pocket of PlamHNL in order to identify the critical residues for substrate affinity. The lowest scores of the affinity variants are shown in Table 3-4. From 13 preliminary selected hits, only 6 variants from three positions at N85, T95, and W109, including N85H, N85Y, W109H, N85E, N85Q, and T95A, could be detected for 2-chlorobenzaldehyde. In contrast, the mutant at positions R58, Y60, L98, and Y124 could not be detected due to these residues being involved in the activity mechanism of PlamHNL (Appendix XVII). The best variant obtained from alanine scanning functionality T95A exhibited an approximately 1.5 times higher specificity than the wild type, whereas N85H, N85Y, W109H, N85E, and N85Q showed lower specific activity than the wild type. These results indicate that the mutation at T95 might contribute to improving the enzyme activity toward 2-chlorobenzaldehyde. In order to demonstrate which variant can improve the enantiomeric excess of (*R*)-2-chloromandelonitrile, the synthesis reaction from 2-chlorobenzaldehyde and KCN was carried out with 1.2 U of all purified variants. Interestingly, N85Y increased to 92% enantiomeric excess of (*R*)-2-chloromandelonitrile compared to 76% of the wild type. Thus, the mutation at N85 to tyrosine contributes to increasing the enantiomeric excess in the production of (*R*)-2-chloromandelonitrile. The double mutation of N85H/T95A and N85Y/T95A was performed; both variants did not

improve the enzyme activity but slightly improved the enantiomeric excess of (*R*)-2-chloromandelonitrile compared with the wild type.

Table 3-4. HNL productivity, activity and enantiomeric excess of selected PlamHNL variants ^[a].

	dAffinity ^[b] (kcal/mol)	Productivity (U/L)	Specific activity ^[c] (U/mg)	% <i>e.e.</i> ^[d]
N85H	-0.9664	477	402 ± 36	85
L98W	-0.7173	N.D.	N.D.	N.D.
L98E	-0.6709	N.D.	N.D.	N.D.
N85Y	-0.6488	233	304 ± 18.2	92
R58W	-0.5245	N.D.	N.D.	N.D.
W109H	-0.4386	160	205 ± 14.2	22
A74W	-0.4163	N.D.	N.D.	N.D.
N85E	-0.3178	382	290 ± 7.6	53
R58K	-0.2335	N.D.	N.D.	N.D.
N85Q	-0.2192	180	165 ± 26.7	15
Y60A	-0.1977	N.D.	N.D.	N.D.
Y124A	-0.1061	N.D.	N.D.	N.D.
T95A	-0.0494	580	695 ± 42	86
N85H/T95A	-	450	420 ± 8.6	81
N85Y/T95A	-	457	401 ± 8.5	88
Wild-type	-	420	506 ± 7.8	76

[a] All of variants were selected based on a calculation of the Alanine or Residue Scanning functionality around 14 amino acid residues in binding pocket using MOE program to identify critical residues for affinity.

[b] dAffinity value refer to the relative binding affinity of the mutation to the wild type protein. dAffinity is equal to the Boltzmann average of the relative affinities of the ensemble. A more negative value indicates a mutation with better affinity.

[c] Specific activity was assayed after 5 min in a reaction mixture containing 2-chlorobenzaldehyde (50 mM) and KCN (100 mM) at 25°C with the same amount (0.64 µg) of purified PlamHNL variants.

[d] Enantiomeric excess was obtained for (*R*)-2-chloromandelonitrile synthesis by HPLC with the same concentration of purified PlamHNL variants (1.2 Unit).

N.D.= Not detectable.

Next, N85 was replaced by single amino acid substitutions with 19 amino acids, and expression, purification, and the HNL activity of all mutants was performed. For position N85, all variants exhibited a lower specificity toward 2-chlorobenzaldehyde than that of the wild type except N85C and N85S, which showed higher specificity. Some mutants, N85P and N85D, could not be expressed in the *E. coli* SHuffle T7 host (Appendix XVIII). At the same amount of enzyme (1.2 U),

most expressible mutants of amino acids with a hydrophobic side chain at position 85 showed similar % *e.e.* value compared to the N85 wild type in the range of 75%-92%. The maximum *e.e.* of 92% was obtained with the N85Y mutant, whereas the mutants of amino acids with polar side chains N85K, N85R, N85E, and N85Q showed a loss of activity and stereoselectivity for (*R*)-2-chloromandelonitrile synthesis (Appendix XVIII-B). This mutation at the N85 site is located at the bottom of the active pocket of PlamHNL that formed a large pocket and small opened tunnel, which is different from ChuaHNL (Figure 3-7D, 3-7E). Substitutions with a bulky amino acid side chain (especially tyrosine) at position N85 might reform the shape of the cavity active site to close a small tunnel and make the cavity more compact than the wild type, similar to the structure obtained at position Y86 in the ChuaHNL structure (Figure 3-7E). While the specific activity of mutant PlamHNL-N85Y was decreased, it showed good enantioselectivity for 2-chloromandelonitrile synthesis. This evidence was supported by docking simulation of (*R*)-2-chloromandelonitrile with active site of variant PlamHNL-N85Y, the distance of hydroxyl of (*R*)-2-chloromandelonitrile interact with R58 and K138 and the nitrile group form hydrogen bond with Y124 were longer bond lengths than that the wild type (3.3, 3.23, and 3.1 Å, respectively). Therefore, the estimated affinity of PlamHNL-N85Y to (*R*)-2-chloromandelonitrile was reduced (-4.40 kcal/mol), which showed the specific activity lower than that the wild type (Figure 3-9B). On the other hand, the improvement of enantioselectivity of PlamHNL-N85Y could be explained by the newly generated cavity form and the CH- π interaction between F87 and the (*R*)-2-chloromandelonitrile was abolished upon substitution of asparagine with tyrosine at position 85. The newly opposite CH- π interaction between (*R*)-2-chloromandelonitrile and Y60 of PlamHNL-N85Y was formed which might preferentially recognizes (*R*)-2-chloromandelonitrile than that the wild type (Figure 3-9A and B).

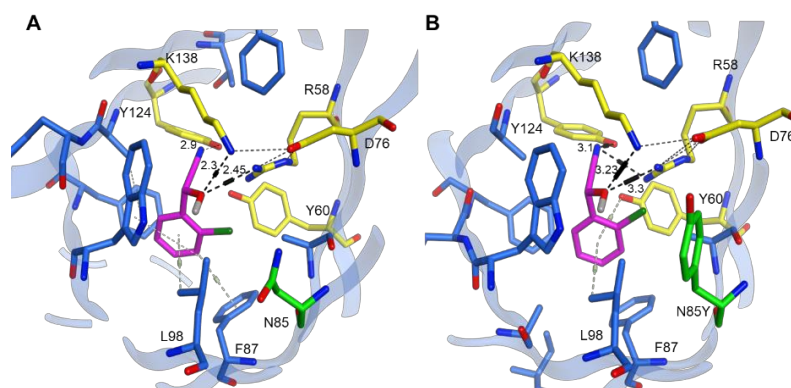


Figure 3-9. The docking model of the wild type (A) and PlamHNL-N85Y mutant (B) complexed with (*R*)-2-chloromandelonitrile (magenta). The crystal structure of wild type and mutant are colored in blue. The catalytic residues

(R58, Y60, D76, Y124, and K138) are shown in yellow. The hydrogen bonds are shown as black dotted lines, and their distances are labeled. The CH- π are shown as light-green dotted lines. N85 and N85Y are colored in green.

Enantioselectivity of PlamHNL mutants

The best mutant candidate for improving enantioselectivity, N85Y, was further studied to optimize the conditions for (*R*)-2-chloromandelonitrile synthesis. The maximum enantiomeric excess of (*R*)-2-chloromandelonitrile produced by N85Y increased to 96.33% when the enzyme quantity was increased (4 U), as compared with the maximum *e.e.* of 87% produced under the same conditions by the wild type enzyme (Figure 3-10A). Furthermore, the *e.e.* of 98.2% was obtained at low pH conditions (pH 3.5) by the N85Y mutant (Figure 3-10B). The production of (*R*)-2-chloromandelonitrile with purified enzyme N85Y showed higher conversion (91% conversion with 98.2% *e.e.*) than the wild type (76% conversion with 90% *e.e.*) at pH 3.5 and 25 °C for 30 min incubation (Figure 3-10C and D). These results showed that the enantioselectivity of PlamHNL for (*R*)-2-chloromandelonitrile synthesis was remarkably improved by the Asn85 mutation.

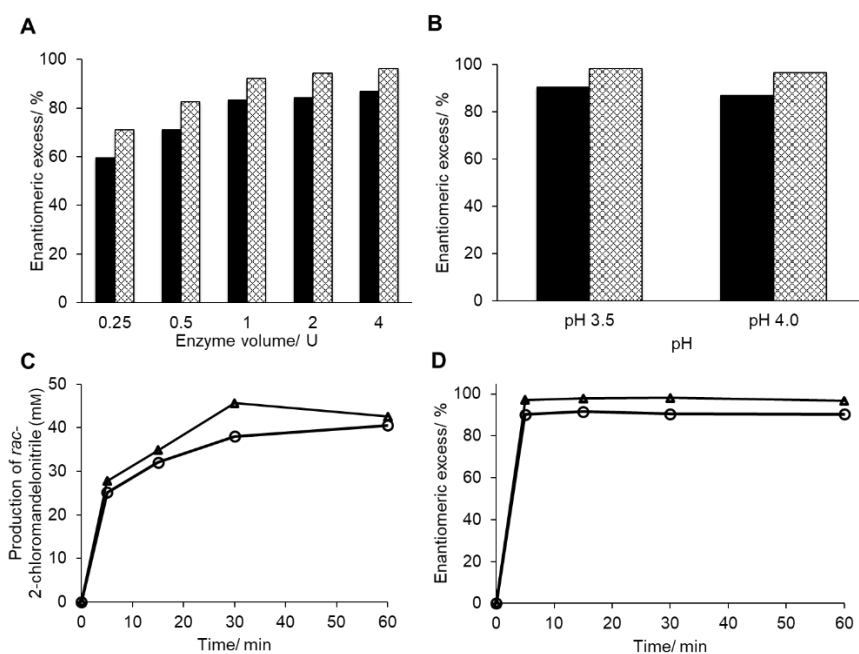


Figure 3-10. Enantioselectivity for (*R*)-2-chloromandelonitrile synthesis of purified enzymes from PlamHNL wild type (black bar or \circ) and mutant N85Y (cross-bar or \blacktriangle). (A) Effect of enzyme quantity. (B) Effect of low-pH conditions. (C) Time course of total (*R+S*)-2-chloromandelonitrile production and their enantiomeric excess (D) in the reaction of citrate buffer (300 mM, pH 3.5) containing 2-chlorobenzaldehyde (50 mM), KCN (100 mM), and enzyme (4 U) at 25 °C.

CHAPTER IV Identification of saturated and unsaturated 1-methoxyalkanes (alkyl methyl ethers) from the Thai millipede *Orthomorpha communis* [Paradoxosomatidae, Polydesmida] as “raincoat compounds,” other than mandelonitrile-related defense allomones

The Thai millipede *Orthomorpha communis* Likhitrakarn, Golovatch & Panha, 2011 (Paradoxosomatidae: Polydesmida) is distributed widely in the eastern part of Thailand close to the border with Cambodia and has newly been found in Hat Yai District, Songkhla Province, Thailand. The live coloration of blackish-brown body rings, with paraterga and epiproct showing a distinctive creamy yellow during contracting (Figure 4-1).



Figure 4-1. Pictures of the Thai millipede *O. communis*.

Using gas chromatography coupled with mass spectrometry (GC/MS) analysis of the species, we happened to detect a series of saturated and unsaturated wax-like components, other than the conventional mixtures derived from mandelonitrile. These wax-like components have not been reported in other millipedes, and our hypothesis is that they have a function similar to the “raincoat” found in the oribatid mite, *Liacarus subterraneus* (Acari: Oribatida). This mite secretes di-glycerides via an esterification reaction between fatty acids and glycerin¹¹⁹). As a result, the chemical structures of these saturated and unsaturated wax-like components was investigated.

MATERIALS AND METHODS

Millipedes: The adult species *Orthomorpha communis* Likhitrakarn, Golovatch & Panha, 2011 [Paradoxosomatidae: Polydesmida], often found individually embedded in soils after heavy rains, were manually collected in Banrai (6°54'38.3"N, 100°28'08.0"E), Hat Yai District, Songkhla Province (Southern Thailand), Thailand in December, 2015, and subsequently reared in the laboratory on leaf litters collected from the site under a natural photoperiod with high humidity (ca. 100 % RH) at 20 °C. Sizes of adult were as follows; body length 15.0–27.5 mm, a width of midbody pro- and meta-zona 1.1–2.3 and 1.5–3.1 mm, respectively.

Hexane extraction of millipedes: Five adult millipedes as a group and also one adult millipede (two times) were dipped into *n*-hexane (1 ml, separately) for three minutes, using appropriate glass vials (5 ml volume). Then, the resultant hexane extracts after separation from residual millipede bodies by decantation were each subjected to gas chromatography coupled with mass spectrometry (GC/MS) analyses.

Analytical methods: GC/MS spectra were obtained as reported¹²⁰⁾ using Hewlett Packard HP 5975C Inert XL EI/CI MSD with triple-Axis Detector at 75 eV coupled with a 7890A GC-system equipped with an HP-5 column (30 m × \varnothing 0.25 mm; 0.25 μ m in film thickness) operated in the split-less mode at 60 °C for 2 min, then programmed to increase at 10 °C /min to 290 °C, and finally held at this temperature for 5 min. Helium was used as the carrier gas at a flow rate of 1.00 ml/min. GC and GC/MS data were processed using ChemStation (Hewlett Packard Co.), with reference to an MS database (Wiley 9th/NIST 2011 MS Library; Hewlett Packard Co.). In the case of the non-resolved peaks (**5** and **6**), relative amounts were calculated by selected ion chromatography (SIC) using each base ion (*m/z* 105 and 109, respectively). Retention indices¹²¹⁾ were calculated under the same GC conditions mentioned above, as described in Bodner and Raspotnig¹²²⁾, and were used for structure elucidation. ¹H-NMR spectra (400 MHz, TMS at δ 00.00 as internal standard) were recorded on a Bruker Biospin AC400M spectrometer.

Chemicals: The following chemicals and solvents were obtained commercially and used as described in the text; two chemicals (**1** and **2**) from Wako Pure Chemical Industries, Japan, two chemicals (**5** and **7**) from Nacalai Tesque Inc., Japan, and five chemicals (**3**, **4**, **6**, **8**, and **10**) from Tokyo Chemical Industry Co., Japan. Compound **26** was prepared as reported previously¹²³⁾. Wako

Gel C-200, hexane, diethylether (Wako Pure Chemical Industries, Ltd.), lithium aluminum hydride, methyl iodide, and metallic sodium (Junsei Chemical, Japan) were used without purification.

Preparation of standard 1-methoxyalkane mixtures from palm oil: A mixture of alcohols derived from the palm oil by reduction with lithium aluminum hydride were lead to the corresponding mixture of 1-methoxyalkanes using Williamson's ether synthesis (reaction between sodium alcoholate and methyl iodide). After addition of water to dissolve solids, the reaction products were extracted by hexane. The hexane extracts, without concentration, was poured on a silica gel column (Wako Gel C-200, 300 mg, 0.5 × 2.1 cm in length) prepared without solvent, using a disposable glass pipette (1.5 ml, Iwaki). 1-Methoxyalkanes were obtained as an eluted fraction by 1 % ether in hexane.

Separation of millipede components by a silica gel column: To the silica gel column (Wako C-200, 300 mg) similarly prepared as mentioned above, the hexane extracts (1 ml) from seven millipedes was subjected, and the column was successively eluted with hexane (3 ml), 1 % Et₂O/hexane (3 ml), 2 % Et₂O/hexane (3 ml), 5 % Et₂O/hexane (3 ml), 10 % Et₂O/hexane (3 ml), 50 % Et₂O/hexane (3 ml), and 100 % Et₂O (3 ml). All fractions were subjected to GC/MS analysis.

RESULTS

GC/MS analyses of the Thai millipede: As shown in Figure 4-2, the GC-profile showed a total of 31 peaks (=compounds) including non-resolved components [2-methoxyphenol (**5**) and benzoyl cyanide (**6**), as mentioned later]. Mass spectra of all peaks are summarized in Table 4-1, including identified results by co-chromatography with authentic compounds and (or) elucidated structures, as mentioned later. A total of 20 peaks (11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 27, 28, 29, 30, and 31, 45.4 %) were newly detected in the present species and represented the first discovery of the hereafter named "waxy compounds" among Polydesmida. Their structures were later elucidated as 1-methoxyalkanes by 1) column chromatography (SiO₂) behavior, 2) NMR analysis, and 3) GC-mass spectra (the presence of M⁺ or M⁺-32, and m/z 45), using 1-methoxyalkanes prepared from palm oil as an authentic compound.

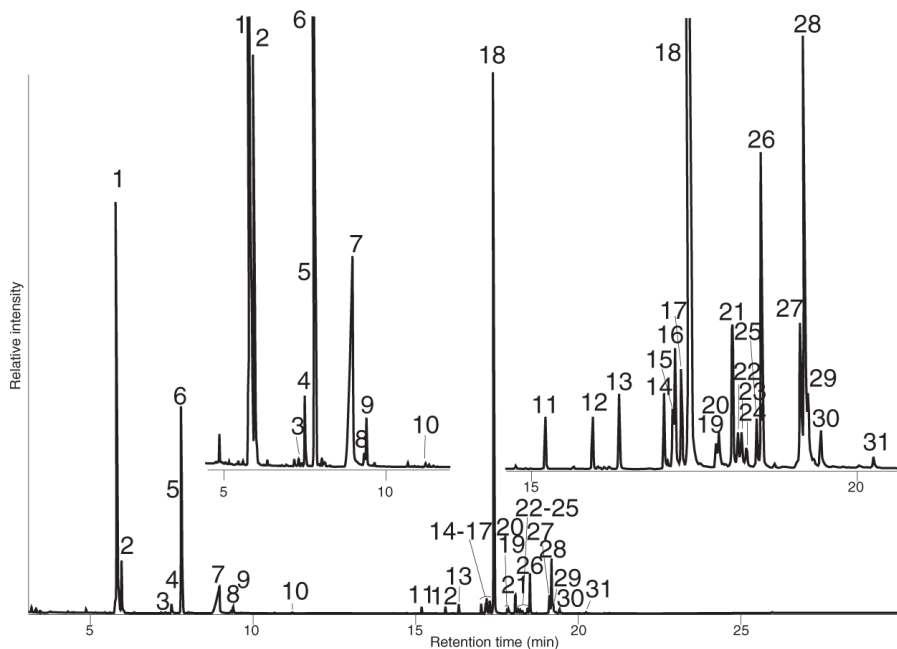


Figure 4-2. GC profile of the hexane extract from the Thai millipede *O. communis*.

1–10 and 26; components known as defensive allomone of polydesmid millipedes, 11–25 and 27–31; waxy compounds. Structure and names see Figure 4-4 and Table 4-1.

Chromatographic behavior of waxy compounds on SiO₂ column: When the hexane extracts from seven adult millipedes were separated by silica gel column (300 mg, 0.5 ϕ \times 2.1 cm) chromatography, most of the waxy compounds were recovered in the fraction (3 ml) eluted with the 1 % Et₂O in hexane, indicating their non-polar nature (less polar than esters). Methyl ether mixtures prepared from palm oil were also recovered in the same 1 % Et₂O in hexane fractions, as mentioned above. Peak 1 was eluted by 2 % Et₂O in hexane (3 ml) with the remaining waxy compounds. As shown in Figure 4-3, peaks 11, 18, 27, 28, and 30 derived from the palm oil were identical to those derived from millipedes using GC t_{RS} and mass spectra (mentioned below).

¹H-NMR spectrum of the waxy compounds from millipedes and 1-methoxyalkanes from palm oil:

The NMR spectrum of millipede compounds after being purified by the SiO₂ column proved to be almost the same as the spectra for the 1-methoxyalkanes found in palm oil, except for a difference in the integral on long chain methylenes. The following chemical shifts were observed for millipedes; ω -methyl (3H, *t*, J= 6.6 Hz at δ 0.88 ppm), long chain methylenes (26H, *m* at δ 1.20–1.38 ppm), -CH₂-CH₂-O- (2H, *m*, at δ 1.52–1.60 ppm), CH₃-O- (3H, *s*, at δ 3.33 ppm), and -CH₂-CH₂-O (2H, *t*, J= 6.6 Hz at δ 3.36 ppm), thus, identifying the structure as 1-methoxyhexadecane. As a result, waxy components from millipede were shown to be a mixture of saturated and unsaturated 1-methoxyalkanes.

Table 4-1. Gas chromatographic and mass spectral data of compounds from extracts of the Thai millipede, *O. communis*

Peak no.	Retention		Content (%)	Mass spectrometric fragmentation (m/z)	Compound identified	
	time	index*			as	Standard obtained (Library ID %)
1	5.838	959	26.8	106(M ⁺ ,100), 105(96), 77(90), 51(32)	benzaldehyde	com. (97%)
2	5.972	970	3.1	94 (M ⁺ , 100), 66 (30.4)	phenol	com. (91%)
3	7.185	1068	0.1	108(M ⁺ , 100), 107(87), 90 (21), 89 (13), 80 (11), 79 (28), 77 (27)	<i>o</i> -cresol	com.
4	7.512	1094	0.6	108(M ⁺ , 81), 107(100), 90 (7), 79 (20), 77(22)	<i>p</i> -cresol	com. (96%)
5	7.819	1118	0.2	124(M ⁺ , 94), 109(100), 81(44), 65(5), 53(15)	**2-methoxyphenol	com. (96%)
6	7.822	1119	14.9	131(M ⁺ , 72), 105(100), 77(55), 74(8), 51(23)	benzoyl cyanide	com. (93%)
7	8.982	1213	6.0	122(M ⁺ , 85), 105(100), 77(72), 51(35)	benzoic acid	com. (96%)
8	9.340	1242	0.1	138(M ⁺ , 77), 123(100), 95(25), 77(12), 67(21), 55(11), 51(9)	2-methoxy-5-methylphenol	com. (94%)
9	9.405	1247	0.5	138(M ⁺ , 96), 123(100), 95(26), 77(8), 67(13), 55(10), 51(8)	3-methoxy-4-methylphenol	94%
10	11.225	1394	trace	133(M ⁺ , 70), 115(39), 105(100), 77(95), 51(45)	mandelonitrile	com. (90%)
11	15.203	1716	0.3	196(M ⁺ -32, 26), 168(15), 140(8), 125(20),111(39), 97(72), 83(92), 69(71), 55(60), 45(100)	1-methoxy- <i>n</i> -tetradecane	syn. (90%)
12	15.930	1775	0.3	210(M ⁺ -32, 20), 195(4),182(18),167(3), 154(19), 139(6), 125(15), 111(40), 97(70), 83(92), 69(90), 56(96), 45(100)	1-methoxy-13-methyl-tetradecane	
13	16.341	1808	0.5	210(M ⁺ -32, 25),182(13),154(6), 140(8), 125(21), 111(46), 97(81), 83(100), 69(73), 55(61), 45(98)	1-methoxy- <i>n</i> -pentadecane	
14	17.033	1864	0.5	224(M ⁺ -32, 11), 209(2), 196(9), 181(3), 168(8), 153(5), 140(12), 125(18), 111(49), 97(72), 83(100), 69(95), 55(94), 45(97)	1-methoxy-14-methylpentadecane	
15	17.159	1874	0.5	254(M ⁺ , 2), 222(10), 194(2), 180(1), 166(2), 152(2), 137(6), 123(12), 109(31), 96(62), 82(100), 67(73), 55(47), 45(39)	1-methoxy- <i>n</i> -hexadecene	
16	17.203	1878	1.0	254(M ⁺ , 2), 222(14), 194(3), 179(1), 166(3), 152(3), 137(8), 123(16), 109(30), 96(71), 82(100), 67(59), 55(61), 45(45)	1-methoxy- <i>Z</i> -9-hexadecene	syn

Peak no.	Retention		Content	Mass spectrometric fragmentation (m/z)	Compound identified	
					as	Standard obtained (Library ID %)
17	17.300	1885	0.7	254(M ⁺ , 2), 222(11), 194(2), 180(1), 166(2), 152(3), 137(7), 123(14), 109(27), 96(74), 82(100), 67(54), 55(78), 45(49)	1-methoxy- <i>E</i> -9-hexadecene	
18	17.466	1899	32.9	224(M ⁺ -32, 31), 213(1), 196(15), 182(2), 168(5), 154(7), 139(10), 125(27), 111(53), 97(88), 83(100), 69(69), 55(59), 45(86)	1-methoxy- <i>n</i> -hexadecane	syn. (91%)
19	17.832	1928	0.2	268(M ⁺ , 2), 236(16), 208(3), 194(1), 180(2), 166(1), 152(3), 137(9), 123(19), 109(34), 96(66), 82(100), 67(54), 55(55), 45(42)	1-methoxy- <i>Z</i> -9-15-methyl-hexadecene	
20	17.872	1932	0.4	238(M ⁺ -32, 2), 223(1), 210(2), 196(1), 182(1), 169(9), 153(22), 140(11), 125(44), 111(66), 97(94), 83(93), 71(89), 69(65), 57(100), 45(68)	1-methoxy-methyl-hexadecane	
21	18.082	1948	1.1	238(M ⁺ -32, 13), 223(2), 210(11), 195(3), 182(10), 168(2), 154(6), 139(5), 125(14), 111(37), 97(57), 83(100), 69(74), 57(78), 45(80)	1-methoxy-15-methylhexadecane	
22	18.166	1955	0.3	238(M ⁺ -32, 5), 227(1), 209(15), 181(7), 168(4), 153(6), 139(12), 125(28), 111(44), 97(80), 83(83), 70(100), 57(62), 55(56), 45(66)	1-methoxy-14-methylhexadecane	
23	18.220	1960	0.3	268(M ⁺ , 2), 236(13), 208(2), 194(1), 180(2), 166(2), 152(3), 137(9), 123(16), 109(31), 96(72), 82(100), 67(56), 55(56), 45(37)	1-methoxy- <i>Z</i> -9-heptadecene	syn.
24	18.295	1966	0.2	268(M ⁺ , 2), 236(7), 208(3), 193(2), 180(3), 166(3), 151(5), 137(9), 123(17), 109(32), 96(72), 82(100), 67(59), 55(68), 45(48)	1-methoxy- <i>E</i> -9-heptadecene	
25	18.454	1979	0.4	238(M ⁺ -32, 20), 227(1), 210(10), 196(1), 182(4), 168(4), 154(5), 139(8), 125(24), 111(48), 97(85), 83(100), 69(68), 57(56), 45(82)	1-methoxy- <i>n</i> -heptadecane	
26	18.530	1985	2.3	237(M ⁺ , 17), 116(39), 105(100), 89(11), 77(25), 51(11)	mandelonitrile benzoate	prepd. (80%)
27	19.126	2033	1.2	280(M ⁺ , 2), 248(3), 219(1), 205(1), 191(1), 184(3), 177(2), 163(4), 152(9), 149(9), 135(25), 121(30), 109(30), 95(69), 81(89), 67(100), 55(51), 45(47)	1-methoxy- <i>Z</i> , <i>Z</i> -9,12-octadecadiene	syn.
28	19.195	2038	3.2	282(M ⁺ , 2), 250(12), 222(2), 208(1), 194(2), 180(1), 166(2), 152(3), 137(9), 123(16), 109(29), 96(70), 82(100), 67(54), 55(54), 45(37)	1-methoxy- <i>Z</i> -9-octadecene	syn. (99%)
29	19.241	2042	0.8	282(M ⁺ , 3), 250(14), 239(1), 222(2), 208(1), 194(2), 180(2), 165(2), 152(4), 137(10), 123(18), 109(31), 96(80), 82(100), 69(49), 55(58), 45(38)	1-methoxy- <i>E</i> -9-octadecene	

Peak no.	Retention		Content	Mass spectrometric fragmentation (m/z)	Compound identified	
					as	Standard obtained (Library ID %)
30	19.438	2058	0.4	252(M ⁺ -32, 19), 241(1), 224(8), 210(1), 196(3), 182(3), 168(4), 154(5), 139(9), 125(26), 111(49), 97(86), 83(100), 69(68), 57(58), 45(78)	1-methoxy- <i>n</i> -octadecane	syn.
31	20.251	2124	0.1	296(M ⁺ , 1), 264(5), 236(2), 222(1), 208(2), 194(1), 180(2), 165(2), 151(4), 137(9), 123(16), 109(28), 96(71), 82(100), 69(55), 55(70), 45(45)	1-methoxy- <i>Z</i> -9-nonadecene	

*: Retention index; calculated by HP-5 column under conditions described in text, as reported (Kovat. 1958),

** : Detectable by selected ion chromatography

com.; commercially available, syn.; prepared from palm oil, prepd; by synthesis

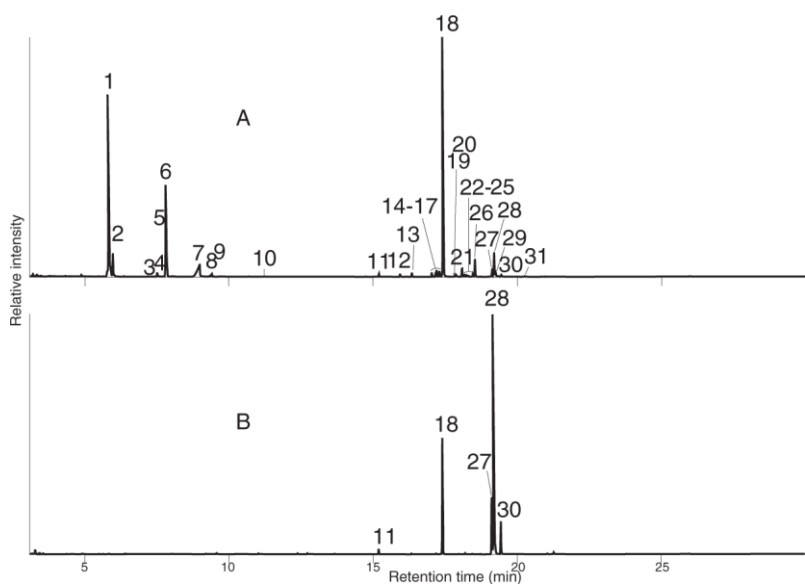


Figure 4-3. Comparison of GC profiles. (A) from Thai millipedes, and (B) 1-methoxyalkanes prepared from palm oil. Structure and names, see Figure 4-4 and Table 4-1.

GC/MS analysis of standard 1-methoxyalkanes prepared from palm oil:

Peaks 11, 18, 27, 28, and 30 in the preparation from the palm oil showed the same GC-retention times and mass spectra to those from millipedes (this corresponded to 84.4 % of the waxy composition in the millipede). Among them, significant results (more than 90 % identities) of library search results were available for peaks 11 (90 %), 18 (91 %) and 28 (99 %). As a result, structures of peaks 11, 18, and 28 were identified as 1-methoxy-*n*-tetradecane (**11**), 1-methoxy-*n*-hexadecane (**18**) and 1-methoxy-*Z*-9-octadecene (**28**), respectively (Figure 4-4).

Based on the above results, the other two peaks 27 and 30 were thought to be similar to the 1-methoxyalkanes originated from the fatty acid compositions of the palm oil, and described as 1-methoxy-*Z*, *Z*-9,12-octadecadiene (**27**) and 1-methoxy-*n*-octadecane (**30**) by mass spectra analysis (Table 4-1, mentioned later).

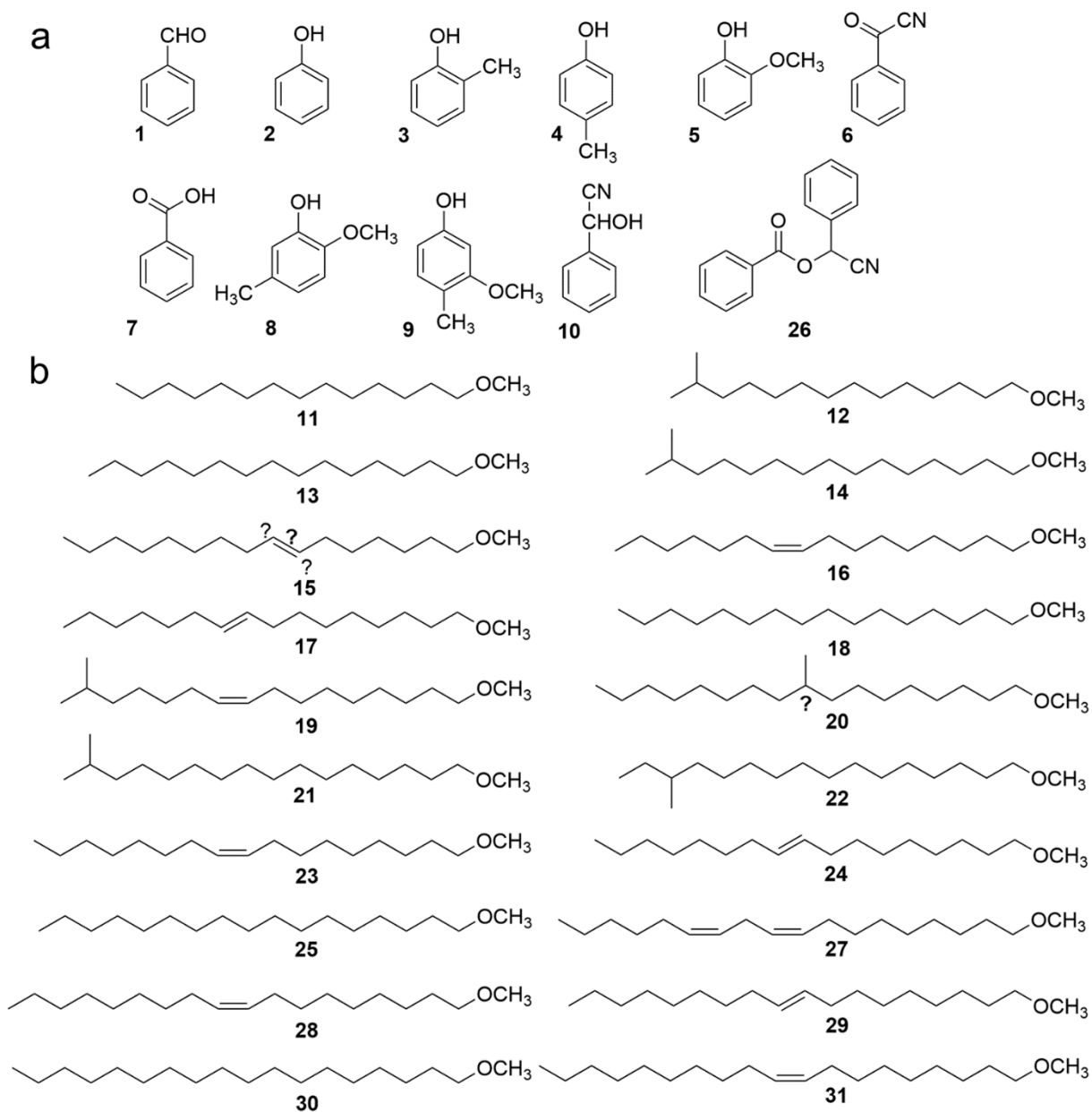


Figure 4-4. Chemical structures identified or elucidated in the hexane extract of the Thai millipede *O. communis*. (a) Compounds commonly detectable among polydesmid millipedes as defense allomone, (b) 1-methoxyalkanes, newly detected in the species.

Structure elucidation of waxy compounds by GC/MS: Mass spectra of waxy compounds mostly indicated alkene series of fragment ions with sigmoid intensity from m/z 41 to m/z 139 or m/z 153. The fragment ion at m/z 45 ($-\text{CH}_2^+-\text{OMe}$) was always observable in all compounds, indicating the

presence of oxygen in the molecule. Saturated compounds (**11**, **12**, **13**, **14**, **18**, **20**, **21**, **22**, **25**, and **30**, except **20** and **22**) displayed each reliable intensity (more than 11 %) of fragment at m/z $M^+ - 32(\text{MeOH})$ with trace intensity fragment at $M^+ - 1$, but no M^+ ion, and m/z 83 was constantly observable as the base ion peak, respectively. On the other hand, in the case of monoene compounds (**15**, **16**, **17**, **19**, **23**, **24**, **28**, **29**, and **31**), M^+ ions were detected with $M^+ - 32(\text{MeOH})$ and the base ion at m/z 82. In the case of dienoic compound (**27**), the base ion was observed at m/z 67. As a whole, these facts implied that elimination of $\text{CH}_3\text{-OH}$ or H^+ could occur in all those components, and each compound was elucidated as 1-methoxyalkane (**11**, **12**, **13**, **14**, **18**, **20**, **21**, **22**, **25**, and **30**), 1-methoxyalkene (**15**, **16**, **17**, **19**, **23**, **24**, **28**, **29**, and **31**), or 1-methoxyalkadiene (**27**) (Table 4-1).

Structure elucidation of 1-methoxyalkanes in the millipede: Mass spectra of **13** and **25** were identical to those of **11** [$M^+ - 32(\text{MeOH})$ ions at m/z 196], **18** [$M^+ - 32(\text{MeOH})$ ions at m/z 224] and **30** [$M^+ - 32(\text{MeOH})$ ions at m/z 252] derived from palm oil reduction, except $M^+ - 32(\text{MeOH})$ ions observed at m/z 210 and 238. Therefore, peak 13 was described as 1-methoxy-*n*-pentadecane (**13**) and peak 25 as 1-methoxy-*n*-heptadecane (**25**) (Table 4-1, Figure 4-4).

Peaks 12, 14, 20, 21, and 22 indicated $M^+ - 32(\text{MeOH})$ ions at m/z 210, m/z 224, m/z 238, m/z 238, and m/z 238 as the largest fragments, respectively, and these were thought to be 1-methoxy-methyl-branched (substituted) alkanes. Differences on Kovat's retention indexes (ΔKI between 1-methoxy-*n*-alkane and 1-methoxy-methyl substituted alkane = increment of additional methyl residue) were calculated as follows; KI of peak 12 (1775) was that of peak 11 (1716) + ΔKI 59, likewise ΔKI 56 for peak 14, ΔKI 33 for peak 20, ΔKI 49 for peak 21 and ΔKI 56 for peak 22. If we assume peak 12 and peak 14 are 1-methoxy-13-methyltetradecane (**12**, *iso*- C_{15} ether, ΔKI 59) and 1-methoxy-14-methylpentadecane (**14**, *iso*- C_{16} ether, ΔKI 56), respectively, peak 21 (ΔKI 49) or 22 (ΔKI 56) should be 1-methoxy-15-methylhexadecane (**21**, *iso*- C_{17} ether), while peak 20 (ΔKI 33) remained obscure as a branched $\text{C}_{17}\text{-OMe}$. If $\Delta\text{KI}'$ is compared between *n*- and *iso*-alkyl series, $\Delta\text{KI}'$ between peaks 12 and 13 was observed as 33, likewise between peaks 14 and 18 as 35, and between peaks 21 and 25 as 31, then 21 was more likely a 1-methoxy-15-methylhexadecane (**21**, *iso*- C_{17} ether), because $\Delta\text{KI}'$ between peaks 22 and 25 was 24. On the other hand, mass fragment $M^+ - 32(\text{MeOH}) - 15(\text{CH}_3)$ ion was observed in peaks 12 (4 %), 14 (2 %), 20 (1 %), 21 (2 %) and 22 (1 %) suggested the presence of methyl residue in the molecule, while peak 22 gave strong $M^+ - 32(\text{MeOH}) - 29(\text{CH}_3\text{CH}_2)$ ion (18 %), indicative of a removable ethyl residue (*ante-iso* carbon chains) in the structure. As a whole, peaks 12, 14, and 21 should be proposed as 1-methoxy-*iso*-alkanes; 1-methoxy-13-methyltetradecane (**12**, *iso*-

pentadecyl methyl ether), 1-methoxy-14-methylpentadecane (**14**, *iso*-hexadecyl methyl ether) and 1-methoxy-15-methylhexadecane (**21**, *iso*-heptadecyl methyl ether), respectively. Peak 22 was subsequently characterized as 1-methoxy-14-methylhexadecane (**22**, *ante-iso*-heptadecyl methyl ether). Peak 20 was supposed to be 1-methoxy-methyl-branched hexadecane (**20**, methyl-branched hexadecyl methyl ether), whose methyl position remained obscure. Stereo chemistry of each optically active carbon in **20** and **22** remained obscure.

Structure elucidation of 1-methoxy-alkenes (unsaturated methyl ethers) in the millipede: Based on calculated C=C double bond increment ($\Delta\text{KI} = -20$) for *Z*-9-ene (peak 28) to saturated (peak 30) of standards derived from palm oil, and likewise ($\Delta\text{KI} = -25$) for *Z,Z*-9, 12-diene (peak 27), peak 16 ($\Delta\text{KI} = -21$) was characterized as 1-methoxy-*Z*-9-hexadecene (**16**, *Z*-9-hexadecenyl methyl ether), and peak 23 ($\Delta\text{KI} = -19$) was characterized as 1-methoxy-*Z*-9-heptadecene (**23**, *Z*-9-heptadecenyl methyl ether). Calculated KI for putative C₁₉-OMe was 2143, then peak 31 (KI=2124) corresponded to the difference of $\Delta\text{KI} = -19$ and the structure of peak 31 was elucidated as 1-methoxy-*Z*-9-nonadecene (**31**, *Z*-9-nonadecenyl methyl ether). KI value of peak 19 (1928) indicated -20 to that of peak 21 (1946), then the resulting $\Delta\text{KI} = -20$ suggested the structure of peak 19 as 1-methoxy-15-methyl-*Z*-9-hexadecene (**19**, *Z*-9-*iso*-heptadecenyl methyl ether). Although peak 15 looked like 1-methoxy-hexadecene (**15**) by GC/MS, its $\Delta\text{KI} = -25$ were larger than *Z*-9-mono-ene ($\Delta\text{KI} = -19 \sim -20$), and its C=C bond position remained obscure. Peaks 17, 24, and 29 had almost the same mass spectra as peaks 16, 23, and 28 with $\Delta\text{KI} = -14, -13, \text{ and } -16$ corresponding saturated compounds. Those were proposed as being *E*-isomers, and tentatively elucidated as 1-methoxy-*E*-9-hexadecene (**17**, *E*-9-hexadecenyl methyl ether), 1-methoxy-*E*-9-heptadecene (**24**, *E*-9-heptadecenyl methyl ether), and 1-methoxy-*E*-9-octadecene (**29**, *E*-9-octadecenyl methyl ether). All structures of compounds from *O. communis*, identified, elucidated, or characterized, are presented in Figure 4-4.

Structure elucidation of conventional millipede components: Another five peaks (1, 6, 7, 10, and 26, a total of 50.0 %) were conventional polydesmid compounds and identified each as benzaldehyde (**1**), benzoyl cyanide (**6**), benzoic acid (**7**), mandelonitrile (**10**) and mandelonitrile benzoate (**26**), using authentic compounds. The other five peaks (2, 3, 4, 5, and 8, in total 4.1 %) were also found among Polydesmida and identified as phenol (**2**), *o*-cresol (**3**), *p*-cresol (**4**), 2-methoxyphenol (**5**) and 2-methoxy-5-methylphenol (**8**), using authentic compounds. Peak 9 (0.5 %) was elucidated as 3-methoxy-4-methylphenol (**9**) by comparison with MS database (Table 4-1).

DISCUSSION

The present polydesmid millipede has recently revealed the possession of a large amount of waxy compounds (mentioned later) in addition to the conventional chemical defense systems represented by five polydesmid compounds {**1**, **6**, **7**, **10**, and **26**, produced by two enzymes; hydroxynitrile lyase (HNL)³ and mandelonitrile oxidase (MOX)¹²⁴ from **10** as the common substrate with subsequent chemical and biochemical reactions} and of six phenolic compounds (**2**, **3**, **5**, **8**, and **9**). Then the species corresponds to one of HCN and H₂O₂ emitter⁶.

The structures of waxy compounds were elucidated and identified as saturated, unsaturated or methyl branched 1-methoxyalkanes (more than 45.4 % of total hexane extracts) by NMR and GC/MS spectra. Most of components (86.6 % of total ethers) were identified by co-chromatography with authentic compounds prepared from the palm oil, as follows; **11** (0.7 %), **16** (2.2 %), **18** (72.5 %), **23** (0.7 %), **27** (2.6 %), **28** (7.0 %), and **30** (0.9 %). Structures of the other ethers (11.6 % of total methyl ethers; **12**, **13**, **14**, **17**, **19**, **21**, **22**, **24**, **25**, **29**, and **31**) were elucidated by calculation using Kovat's retention index as summarized in Table 4-1. The position of a double bond in **15** (1.1 %) and that of methyl residue in **20** (0.9 %) remained obscure. Alkyl moieties of methyl ethers are mostly identical to those of long chain fatty acids, and presumably have those fatty acids as a precursor. At present, there are no other polydesmid millipedes known to possess those waxy compounds, such as hydrocarbons and fatty acid esters, nor other related presumably functioning as "raincoat compounds"

As far as our data searches are concerned, no other millipedes including Polydesmida or other animals have been reported to produce 1-methoxyalkanes (methyl alkyl ethers)^{125, 126}, except for the spider *Nephila clavipes* [Arachnida: Araneae]¹²⁷. The major components of the spider silk extracted by pentane or methylene chloride consist of a complex mixture of methyl-branched 1-methoxyalkanes [total 51 compounds, all of those contained up to four methyl groups in each carbon chain (chain length between C₂₈ and C₃₄)], together with small amounts of hydrocarbons and alcohols¹²⁷.

As a group of waxy components, fatty acid esters have been distributed in several species of millipedes belonging to Julida, other than methyl- and methoxy-substituted benzoquinones^{125, 126}. A mixture of hexadecyl acetate, Δ 9-hexadecenyl acetate and Δ 9-octadecenyl acetate has been identified in *Blaniulus guttulatus* [Julida: Blaniulidae]¹²⁸, likewise hexyl oleate and octyl oleate have been identified in *Cibiulus phlepsii* and *Nopoiulus kochii* [Julida: Blaniulidae]¹²⁹. Hexyl esters of alkanolic

acids have also been found in *Enantiulus nanus* and *Julus scandinavicus* [Julida: Julidae]¹³⁰). A total of 15 esters, composed of fatty acids (C₉–C₁₄, including *iso*- and *ante-iso*-) and *n*-alcohols (C₄–C₈) have been identified in *Anaulaciulus* sp. [Julida: Julidae]¹³¹). However, distribution of hydrocarbons, as waxy components or body surface compounds, has not been reported in millipedes. Two types of unwettability (raincoat effect) due to chemical compounds have been summarized in Liacaridae (Acari, Oribatida), other than physical and morphological “Lotus effect”; 1) a mixture of C₈–C_{10:2} acid, and their 1,2- and 1-3-di-glycerides, and 2) esters composed of C₁₅ and C₁₆ acids with C₁₄–C₁₇ alcohols¹³²)

CONCLUSION

In this study, the author investigated the molecular cloning, heterologous expression and effect of glycosylation on the biocatalytic properties of hydroxynitrile lyase from plant, *Passiflora edulis* and millipedes, *C. hualienensis* and *P. laminata*. Moreover, this dissertation also newly reported a series of saturated and unsaturated wax-like components, other than conventional mixtures derived from mandelonitrile in the Thai millipede *Orthomorpha communis*. The finding chapters could be summarized as follows.

CHAPTER I

The present data warrant consideration of *PeHNL* in the synthesis of important industrial intermediates. *PeHNL* cDNA was successfully cloned and expressed in *E. coli* BL 21 (DE3) and *P. pastoris* GS115 cells. The wild type enzyme and the glycosylated enzyme that was expressed in *P. pastoris* showed much better thermostability, pH stability, and organic solvent tolerance than non-glycosylated *PeHNL*s that were expressed in *E. coli* and the N105Q mutant that lacked the glycosylation site. These data indicate that the glycosylation system in *P. pastoris* is desirable when used as an expression host for *PeHNL*. Finally, synthesis of (*R*)-mandelonitrile and reusability in biphasic systems with *PeHNL*s were significantly affected by glycosylation and the type of organic solvent. This study confirms the key roles of *N*-glycosylation in enzyme activity, kinetic parameters, and the stability of *PeHNL*.

CHAPTER II

The natural and glycosylated form of *PeHNL* from the yellow passion fruit showed much higher thermostability, pH stability, and the tolerance to organic solvents than *PeHNL* from the purple passion fruit, indicating the important role of the glycosylation. The glycan present at one glycosylation site located on a high flexibility C-terminal region might stabilize the structure and thus reduce the structural flexibility of *PeHNL*. This hypothesis has been proved by generation of a C-terminal truncated form of *PeHNL* lacking 15 amino acid residues, and then expressing it in *E. coli* resulting in an enzyme lacking in glycosylation. The non-glycosylated and C-terminal truncated

mutant *PeHNL* Δ_{107} could improve the stability and reusability, as effectively as the glycosylated natural *PeHNL* and expressed in *P. pastoris*.

CHAPTER III

ChuaHNL was constructed a *P. pastoris* expression system for high-level production of His-ChuaHNL using the co-expression of codon-optimized *ChuaHNL* and *PpPDI*. Using this expression system, His-ChuaHNL productivity increased to $3,170 \pm 144.7$ U/L from 22.6 ± 3.8 U/L, about 140 times. We also revealed that the sugar chains are essential to the high-level production of His-ChuaHNL in the *Pichia* expression system. Thus, the codon-optimization of *ChuaHNL*, co-expression of *PpPDI*, and the site of glycosylation all play a very important role for efficient and correct folding of ChuaHNL in the *P. pastoris* expression system. Since His-tagged enzymes have a merit of easy purification, our expression system of His-ChuaHNL in *P. pastoris* will contribute to facilitating functional studies and industrial applications of ChuaHNL.

PlamHNL is a member of the Xystodesmidea family, which shows catalytic properties different from known millipede HNLs. The cDNA of PlamHNL was cloned and successfully expressed in the *E. coli* SHuffleT7 strain, which can correctly fold proteins with multiple disulfide bonds in the cytoplasm and is expressed in the eukaryotic *P. pastoris* system that has post-translational glycosylation. The glycosylated PlamHNL expressed in *P. pastoris* exhibited much better thermostability and pH stability than that of *E. coli*. This HNL also showed broad substrate specificity for aromatic aldehydes with high enantioselectivity. Although the PlamHNL structure showed high similarity to the folding pattern of the lipocalins akin to the tertiary structure of ChuaHNL and catalytic mechanism residues were identical to another HNL, the shape of the binding cavity was different and consequently might affect the substrate specificity of each millipede HNL. This structural characteristic was also used to guide improvements in stereoselectivity for (*R*)-2-chloromandelonitrile synthesis using a computational-directed evolution approach. We propose that the PlamHNL mutant can serve as a biocatalyst in industrial biotechnology.

CHAPTER IV

1-methoxyalkanes (a total of 20 compounds of methyl alkyl ethers, a class of compounds) were discovered as components of hexane extracts in the present species. It is only the second example of its presence in the animal kingdom, and it is presumably present on the body surface as a raincoat to keep off water penetration and also to prevent desiccation. The present discovery of unusual 1-methoxyalkanes in this polydesmid millipede might represent evidence of a chemical evolution of the species that adapts and thrives in heavy tropical rainfalls. Further examples of similar species should be expected.

ACKNOWLEDGMENTS

The author wishes to express his sincere thanks to Professor Yasuhisa Asano, Toyama Prefectural University, for his continuous guidance and encouragement during course of this work. He has made available his support in a many ways, starting from the application for the MEXT scholarship in 2013, accommodation, supporting my life after the scholarship was finished along with the supervision of this dissertation.

It is a pleasure to acknowledge the member of my thesis review committee including Professors Yasuhisa Asano, Yasuhiro Igarashi, Hidenobu Komeda, Makoto Hibi of Toyama Prefectural University and Professor Takao Hibi of Fukui Prefectural University.

I would like to express my gratitude to Associate Professor Dr. Makoto Hibi and Assistant Professor Dr. Daisuke Matsui of Laboratory of Enzyme Chemistry and Engineering, for their contribution in my experiment, guidance, comments and advises during these years.

The author is gratefully appreciates kindness shown by the member of Asano Active Enzyme Molecule Project (2014-2017), the Exploratory Research for Advanced Technology (ERATO) including Professor Dr. Yasumasa Kuwahara, Professor Dr. Kimiyasu Isobe, Dr. Mohammad Dadashpour, Dr. Yuko Ishida, Dr. Fumihito Motojima, Dr. Takuya Yamaguchi, Dr. Zhenyu Zhai, Dr. Suguru Shinoda and the JSPS Postdoctoral Fellowships members, Dr. Siriporn Chaikhaew and Dr. Makoto Nakabayashi, for their contribution in my experiment, teaching, guidance, comments and advises during these years.

I am thankful to Professor Aran H-Kittikun and Associate Professor Tipparat Hongpattarakere, of Faculty of Agro-Industry, Prince of Songkla University, Associate Professor Chartchai Khanongnuch, of Faculty of Agro-Industry, Chiang Mai University, for their kind discussion and for their help in the collaboration between Thailand and Japan.

The author appreciate the MEXT (Ministry of Education, Culture, Sports, Science and Technology) of Japan for financial support of this Ph. D. course (2013-2016).

I offer my regards and blessing to all my lab members and friends who helped in many ways. It is a great honor for me to thank all the people who taught and trained me until today.

Last, but not least, the author thanks his family for their encouragements, supports and tolerances.

REFERENCES

- 1) Zagrobelny, M., Bak, S. and Moller, B. L.: Cyanogenesis in plants and arthropods, *Phytochemistry*, **69**, 1457-1468 (2008).
- 2) Yamaguchi, T., Kuwahara, Y. and Asano, Y.: A novel cytochrome P450, CYP3201B1, is involved in (*R*)-mandelonitrile biosynthesis in a cyanogenic millipede, *FEBS Open Bio*, **7**, 335-347 (2017).
- 3) Dadashpour, M., Ishida, Y., Yamamoto, K. and Asano, Y.: Discovery and molecular and biocatalytic properties of hydroxynitrile lyase from an invasive millipede, *Chamberlinius hualienensis*, *Proc. Natl. Acad. Sci.*, **112**, 10605-10610 (2015).
- 4) Yamaguchi, T., Nuylert, A., Ina, A., Tanabe, T. and Asano, Y.: Hydroxynitrile lyases from cyanogenic millipedes: molecular cloning, heterologous expression, and whole-cell biocatalysis for the production of (*R*)-mandelonitrile, *Sci. Rep.*, **8**, 3051 (2018).
- 5) Ishida, Y., Kuwahara, Y., Dadashpour, M., Ina, A., Yamaguchi, T., Morita, M., Ichiki, Y. and Asano, Y.: A sacrificial millipede altruistically protects its swarm using a drone blood enzyme, mandelonitrile oxidase, *Sci. Rep.*, **6**, 26998 (2016).
- 6) Kuwahara, Y., Yamaguchi, T., Ichiki, Y., Tanabe, T. and Asano, Y.: Hydrogen peroxide as a new defensive compound in “benzoyl cyanide” producing polydesmid millipedes, *Sci Nat*, **104**, 19 (2017).
- 7) Kuwahara, Y., Shimizu, N. and Tanabe, T.: Release of hydrogen cyanide via a post-secretion Schotten-Baumann reaction in defensive fluids of polydesmoid millipedes, *J. Chem. Ecol.*, **37**, 232-238 (2011).
- 8) Dadashpour, M. and Asano, Y.: Hydroxynitrile lyases: insights into biochemistry, discovery, and engineering, *ACS Catal.*, **1**, 1121-1149 (2011).
- 9) Jansen, I., Woker, R. and Kula, M.-R.: Purification and protein characterization of hydroxynitrile lyases from sorghum and almond, *Biotechnol. Appl. Biochem.*, **15**, 90-99 (1992).
- 10) Yemm, R. S. and Poulton, J. E.: Isolation and characterization of multiple forms of mandelonitrile lyase from mature black cherry (*Prunus serotina* Ehrh.) seeds, *Arch. Biochem. Biophys.*, **247**, 440-445 (1986).
- 11) Nanda, S., Kato, Y. and Asano, Y.: A new (*R*)-hydroxynitrile lyase from *Prunus mume*: asymmetric synthesis of cyanohydrins, *Tetrahedron*, **61**, 10908-10916 (2005).
- 12) Asif, M. and Bhalla, T. C.: Hydroxynitrile lyase of wild apricot (*Prunus armeniaca* L.): Purification, characterization and application in synthesis of enantiopure mandelonitrile, *Catalysis Letters*, **146**, 1118-1127 (2016).
- 13) Zheng, Y. C., Xu, J. H., Wang, H., Lin, G. Q., Hong, R. and Yu, H. L.: Hydroxynitrile lyase isozymes from *Prunus communis*: identification, characterization and synthetic applications, *Adv. Synth. Catal.*, **359**, 1185-1193 (2017).
- 14) Albrecht, J., Jansen, I. and Kula, M. R.: Improved purification of an (*R*)-oxynitrilase from *Linum usitatissimum* (flax) and investigation of the substrate range, *Biotechnol. Appl. Biochem.*, **17**, 191-203 (1993).
- 15) Wajant, H., Forster, S., Selmar, D., Effenberger, F. and Pfizenmaier, K.: Purification and characterization of a novel (*R*)-mandelonitrile lyase from the fern *Phlebodium aureum*, *Plant physiol.*, **109**, 1231-1238 (1995).
- 16) Andexer, J., von Langermann, J., Mell, A., Bocola, M., Kragl, U., Eggert, T. and Pohl, M.: An *R*-selective hydroxynitrile lyase from *Arabidopsis thaliana* with an α/β -hydrolase fold, *Angew. Chem. Int. Ed.*, **46**, 8679-8681 (2007).

- 17) Ueatrongchit, T., Kayo, A., Komeda, H., Asano, Y. and H-Kittikun, A.: Purification and Characterization of a novel (*R*)-Hydroxynitrile lyase from *Eriobotrya japonica* (Loquat), *Biosci. Biotechnol. Biochem.*, **72**, 1513-1522 (2008).
- 18) Asano, Y., Tamura, K. i., Doi, N., Ueatrongchit, T., H-Kittikun, A. and Ohmiya, T.: Screening for new hydroxynitrilases from plants, *Biosci. Biotechnol. Biochem.*, **69**, 2349-2357 (2005).
- 19) Ueatrongchit, T., Tamura, K. i., Ohmiya, T., H-Kittikun, A. and Asano, Y.: Hydroxynitrile lyase from *Passiflora edulis*: Purification, characteristics and application in asymmetric synthesis of (*R*)-mandelonitrile, *Enzyme Microb. Technol.*, **46**, 456-465 (2010).
- 20) Nuylert, A., Ishida, Y. and Asano, Y.: Effect of glycosylation on the biocatalytic properties of hydroxynitrile lyase from the passion fruit, *Passiflora edulis*: A comparison of natural and recombinant enzymes, *ChemBioChem*, **18**, 257-265 (2017).
- 21) Lanfranchi, E., Pavkov-Keller, T., Koehler, E.-M., Diepold, M., Steiner, K., Darnhofer, B., Hartler, J., Van Den Bergh, T., Joosten, H.-J. and Gruber-Khadjawi, M.: Enzyme discovery beyond homology: a unique hydroxynitrile lyase in the Bet v1 superfamily, *Sci. Rep.*, **7**, 46738 (2017).
- 22) Isobe, K., Kitagawa, A., Kanamori, K., Kashiwagi, N., Matsui, D., Yamaguchi, T., Fuhshuku, K.-i., Semba, H. and Asano, Y.: Characterization of a novel hydroxynitrile lyase from *Nandina domestica* Thunb, *Biosci. Biotechnol. Biochem.*, **82**, 1760-1769 (2018).
- 23) Hanefeld, U., Straathof, A. J. and Heijnen, J. J.: Study of the (*S*)-hydroxynitrile lyase from *Hevea brasiliensis*: Mechanistic implications, *Biochim. Biophys. Acta, Protein Struct. Mol. Enzymol.*, **1432**, 185-193 (1999).
- 24) Hughes, J., Decarvalho, J. and Hughes, M. A.: Purification, characterization, and cloning of α -hydroxynitrile lyase from cassava (*Manihot esculenta* Crantz), *Arch. Biochem. Biophys.*, **311**, 496-502 (1994).
- 25) Wajant, H. and Mundry, K.-W.: Hydroxynitrile lyase from *Sorghum bicolor*: a glycoprotein heterotetramer, *Plant Sci.*, **89**, 127-133 (1993).
- 26) Dadashipour, M., Yamazaki, M., Momonoi, K., Tamura, K., Fuhshuku, K., Kanase, Y., Uchimura, E., Kaiyun, G. and Asano, Y.: *S*-selective hydroxynitrile lyase from a plant *Baliospermum montanum*: molecular characterization of recombinant enzyme, *J. Biotechnol.*, **153**, 100-110 (2011).
- 27) Dreveny, I., Gruber, K., Glieder, A., Thompson, A. and Kratky, C.: The hydroxynitrile lyase from almond: A lyase that looks like an oxidoreductase, *Structure*, **9**, 803-815 (2001).
- 28) Fukuta, Y., Nanda, S., Kato, Y., Yurimoto, H., Sakai, Y., Komeda, H. and Asano, Y.: Characterization of a new (*R*)-hydroxynitrile lyase from the Japanese apricot *Prunus mume* and cDNA cloning and secretory expression of one of the isozymes in *Pichia pastoris*, *Biosci Biotechnol Biochem*, **75**, 214-220 (2011).
- 29) Hajnal, I., Lyskowski, A., Hanefeld, U., Gruber, K., Schwab, H. and Steiner, K.: Biochemical and structural characterization of a novel bacterial manganese-dependent hydroxynitrile lyase, *FEBS J.*, **280**, 5815-5828 (2013).
- 30) Wiedner, R., Gruber-Khadjawi, M., Schwab, H. and Steiner, K.: Discovery of a novel (*R*)-selective bacterial hydroxynitrile lyase from *Acidobacterium capsulatum*, *Comput. Struct. Biotechnol. J.*, **10**, 58-62 (2014).
- 31) Laemmli, U. K.: Cleavage of structural proteins during the assembly of the head of bacteriophage T4, *Nature*, **227**, 680-685 (1970).
- 32) Ueatrongchit, T., Kayo, A., Komeda, H., Asano, Y. and H-Kittikun, A.: Purification and characterization of a novel (*R*)-hydroxynitrile lyase from *Eriobotrya japonica* (Loquat), *Bioscience, Biotechnology, and Biochemistry*, **72**, 1513-1522 (2008).

- 33) Ueatrongchit, T., Komeda, H., Asano, Y. and H-Kittikun, A.: Parameters influencing asymmetric synthesis of (*R*)-mandelonitrile by a novel (*R*)-hydroxynitrile lyase from *Eriobotrya japonica*, *J. Mol. Catal. B Enzym.*, **56**, 208–214 (2009).
- 34) Jansen, I., Woker, R. and Kula, M. R.: Purification and protein characterization of hydroxynitrile lyases from Sorghum and Almond, *Biotechnol. Appl. Biochem.*, **15**, 90-99 (1992).
- 35) Albrecht, J., Jansen, I. and Kula, M. R.: Improved purification of and (*R*)-oxynitrilase from *Linum usitatissimum* (flax) and investigation of the substrate range, *Biotechnol. Appl. Biochem.*, **17**, 191-203 (1993).
- 36) Wajant, H., Forster, S., Selmar, D., Effenberger, F. and Pfizenmaier, K.: Purification and characterization of a novel (*R*)-mandelonitrile lyase from the fern *Phlebodium aureum*, *Plant Physiology*, **109**, 1231-1238 (1995).
- 37) Dgany, O., Gonzalez, A., Sofer, O., Wang, W., Zolotnitsky, G., Wolf, A., Shoham, Y., Altman, A., Wolf, S. G., Shoseyov, O. and Almog, O.: The structural basis of the thermostability of SP1, a novel plant (*Populus tremula*) boiling stable protein, *J Biol Chem*, **279**, 51516-51523 (2004).
- 38) Bingman, C. A., Johnson, K. A., Peterson, F. C., Frederick, R. O., Zhao, Q., Thao, S., Fox, B. G., Volkman, B. F., Jeon, W. B. and Smith, D. W.: Crystal structure of the protein from gene At3g17210 of *Arabidopsis thaliana*, *Proteins*, **57**, 218-220 (2004).
- 39) Yang, X., Matsui, T., Kodama, T., Mori, T., Zhou, X., Taura, F., Noguchi, H., Abe, I. and Morita, H.: Structural basis for olivetolic acid formation by a polyketide cyclase from *Cannabis sativa*, *The FEBS journal* (2016).
- 40) Wildt, S. and Gerngross, T. U.: The humanization of *N*-glycosylation pathways in yeast, *Nat Rev Micro*, **3**, 119-128 (2005).
- 41) Yáñez, E., Carmona, T. A., Tiemblo, M., Jiménez, A. and Fernández-Lobato, M.: Expression of the *Schwanniomyces occidentalis* SWA2 amylase in *Saccharomyces cerevisiae*: role of *N*-glycosylation on activity, stability and secretion, *Biochem. J.*, **329**, 65-71 (1998).
- 42) Zou, S., Guo, S., Kaleem, I. and Li, C.: Purification, characterization and comparison of *Penicillium purpurogenum* β -glucuronidases expressed in *Escherichia coli* and *Pichia pastoris*, *J. Chem. Technol. Biotechnol.*, **88**, 1913-1919 (2013).
- 43) Okrob, D., Metzner, J., Wiechert, W., Gruber, K. and Pohl, M.: Tailoring a stabilized variant of hydroxynitrile lyase from *Arabidopsis thaliana*, *ChemBioChem*, **13**, 797-802 (2012).
- 44) Mopai, K. L.: Screening, purification and characterisation of an active hydroxynitrile lyase (nitrilase) from indigenous South African plants, University of Limpopo, (2013).
- 45) Kawahara, N. and Asano, Y.: Mutagenesis of an Asn156 residue in a surface region of *S*-selective hydroxynitrile lyase from *Baliospermum montanum* enhances catalytic efficiency and enantioselectivity, *ChemBioChem*, **16**, 1891-1895 (2015).
- 46) Anbarasan, S., Janis, J., Paloheimo, M., Laitaoja, M., Vuolanto, M., Karimaki, J., Vainiotalo, P., Leisola, M. and Turunen, O.: Effect of glycosylation and additional domains on the thermostability of a family 10 xylanase produced by *Thermopolyspora flexuosa*, *Appl Environ Microbiol*, **76**, 356-360 (2010).
- 47) Skropeta, D.: The effect of individual *N*-glycans on enzyme activity, *Bioorg Med Chem*, **17**, 2645-2653 (2009).
- 48) Sola, R. J. and Griebenow, K.: Effects of glycosylation on the stability of protein pharmaceuticals, *J Pharm Sci*, **98**, 1223-1245 (2009).
- 49) Zou, S., Huang, S., Kaleem, I. and Li, C.: *N*-Glycosylation enhances functional and structural stability of recombinant beta-glucuronidase expressed in *Pichia pastoris*, *J Biotechnol*, **164**, 75-81 (2013).

- 50) Weis, R., Gaisberger, R., Gruber, K. and Glieder, A.: Serine scanning: a tool to prove the consequences of *N*-glycosylation of proteins, *J. Biotechnol.*, **129**, 50-61 (2007).
- 51) Meldgaard, M. and Svendsen, I.: Different effects of *N*-glycosylation on the thermostability of highly homologous bacterial (1,3-1,4)- β -glucanases secreted from yeast, *Microbiology*, **140**, 159-166 (1994).
- 52) Fonseca-Maldonado, R., Vieira, D. S., Alponi, J. S., Bonneil, E., Thibault, P. and Ward, R. J.: Engineering the pattern of protein glycosylation modulates the thermostability of a GH11 xylanase, *J. Biol. Chem.*, **288**, 25522-25534 (2013).
- 53) Ahern, T. J. and Klibanov, A. M.: Analysis of processes causing thermal inactivation of enzymes. In: *Methods Biochem. Anal.*, p.91-128, John Wiley & Sons, Inc., (2006).
- 54) Klibanov, A. M.: Improving enzymes by using them in organic solvents, *Nature*, **409**, 241-246 (2001).
- 55) Ognyanov, V. I., Datcheva, V. K. and Kyler, K. S.: Preparation of chiral cyanohydrins by an oxynitrilase-mediated transcyanation, *J. Am. Chem. Soc.*, **113**, 6992-6996 (1991).
- 56) Bauer, M., Griengl, H. and Steiner, W.: Parameters influencing stability and activity of a S-hydroxynitrile lyase from *Hevea brasiliensis* in two-phase systems, *Enzyme Microb. Technol.*, **24**, 514-522 (1999).
- 57) Persson, M., Costes, D., Wehtje, E. and Adlercreutz, P.: Effects of solvent, water activity and temperature on lipase and hydroxynitrile lyase enantioselectivity, *Enzyme Microb. Technol.*, **30**, 916-923 (2002).
- 58) Costes, D., Wehtje, E. and Adlercreutz, P.: Hydroxynitrile lyase-catalyzed synthesis of cyanohydrins in organic solvents: Parameters influencing activity and enantiospecificity, *Enzyme Microb. Technol.*, **25**, 384-391 (1999).
- 59) Hickel, A., Radke, C. J. and Blanch, H. W.: Role of organic solvents on Pa-hydroxynitrile lyase interfacial activity and stability, *Biotechnol. Bioeng.*, **74**, 18-28 (2001).
- 60) Ueatrongchit, T., Komeda, H., Asano, Y. and Aran, H.: Parameters influencing asymmetric synthesis of (*R*)-mandelonitrile by a novel (*R*)-hydroxynitrile lyase from *Eriobotrya japonica*, *J. Mol. Catal. B: Enzym.*, **56**, 208-214 (2009).
- 61) Pollard, D. J. and Woodley, J. M.: Biocatalysis for pharmaceutical intermediates: the future is now, *Trends in Biotechnology*, **25**, 66-73 (2007).
- 62) Zarschler, K., Janesch, B., Pabst, M., Altmann, F., Messner, P. and Schaffer, C.: Protein tyrosine *O*-glycosylation-a rather unexplored prokaryotic glycosylation system, *Glycobiology*, **20**, 787-798 (2010).
- 63) Reetz, M. T., Carballeira, J. D. and Vogel, A.: Iterative saturation mutagenesis on the basis of B factors as a strategy for increasing protein thermostability, *Angew. Chem. Int. Ed. Engl.*, **45**, 7745-7751 (2006).
- 64) Motojima, F., Nuylert, A. and Asano, Y.: The crystal structure and catalytic mechanism of hydroxynitrile lyase from passion fruit, *Passiflora edulis*, *FEBS J.*, **285**, 313-324 (2018).
- 65) Nagashima, Y., von Schaewen, A. and Koiwa, H.: Function of N-glycosylation in plants, *Plant Sci.*, **274**, 70-79 (2018).
- 66) Rayon, C., Lerouge, P. and Faye, L.: The protein *N*-glycosylation in plants, *J. Exp. Bot.*, **49**, 1463-1472 (1998).
- 67) Faye, L. and Chrispeels, M. J.: Apparent inhibition of β -fructosidase secretion by tunicamycin may be explained by breakdown of the unglycosylated protein during secretion, *Plant physiol.*, **89**, 845-851 (1989).
- 68) Faye, L. and Chrispeels, M. J.: Transport and processing of the glycosylated precursor of concanavalin A in jack-bean, *Planta*, **170**, 217-224 (1987).

- 69) Rodriguez-Amaya, D. B.: PASSION FRUITS. In: Encyclopedia of Food Sciences and Nutrition (Second Edition) (Caballero, B. ed., p.4368-4373, Academic Press, Oxford, (2003).
- 70) Liu, Z., Pscheidt, B., Avi, M., Gaisberger, R., Hartner, F. S., Schuster, C., Skranc, W., Gruber, K. and Glieder, A.: Laboratory evolved biocatalysts for stereoselective syntheses of substituted benzaldehyde cyanohydrins, *ChemBioChem*, **9**, 58-61 (2008).
- 71) Pscheidt, B., Liu, Z., Gaisberger, R., Avi, M., Skranc, W., Gruber, K., Griengl, H. and Glieder, A.: Efficient biocatalytic synthesis of (*R*)-pantolactone, *Adv. Synth. Catal.*, **350**, 1943-1948 (2008).
- 72) Wiedner, R., Kothbauer, B., Pavkov-Keller, T., Gruber-Khadjawi, M., Gruber, K., Schwab, H. and Steiner, K.: Improving the properties of bacterial *R*-selective hydroxynitrile lyases for industrial applications, *ChemCatChem*, **7**, 325-332 (2015).
- 73) Pscheidt, B., Avi, M., Gaisberger, R., Hartner, F. S., Skranc, W. and Glieder, A.: Screening hydroxynitrile lyases for (*R*)-pantolactone synthesis, *J. Mol. Catal., B Enzym.*, **52-53**, 183-188 (2008).
- 74) Rej, R. K., Das, T., Hazra, S. and Nanda, S.: Chemoenzymatic asymmetric synthesis of fluoxetine, atomoxetine, nisoxetine, and duloxetine, *Tetrahedron Asymmetry*, **24**, 913-918 (2013).
- 75) Terreni, M., Pagani, G., Ubiali, D., Fernandez-Lafuente, R., Mateo, C. and Guisán, J. M.: Modulation of penicillin acylase properties via immobilization techniques: one-pot chemoenzymatic synthesis of cephamandole from cephalosporin C, *Bioorg. Med. Chem. Lett.*, **11**, 2429-2432 (2001).
- 76) Ema, T., Ide, S., Okita, N. and Sakai, T.: Highly efficient chemoenzymatic synthesis of methyl (*R*)-*o*-chloromandelate, a key intermediate for clopidogrel, via asymmetric reduction with recombinant *Escherichia coli*, *Adv. Synth. Catal.*, **350**, 2039-2044 (2008).
- 77) Wajant, H. and Effenberger, F.: Hydroxynitrile lyases of higher plants, *Biol Chem*, **377**, 611-617 (1996).
- 78) Caruso, C. S., de Fatima Travençolo, R., de Campus Bicudo, R., de Macedo Lemos, E. G., Ulian de Araujo, A. P. and Carrilho, E.: alpha-Hydroxynitrile lyase protein from *Xylella fastidiosa*: Cloning, expression, and characterization, *Microb. Pathog.*, **47**, 118-127 (2009).
- 79) Hussain, Z., Wiedner, R., Steiner, K., Hajek, T., Avi, M., Hecher, B., Sessitsch, A. and Schwab, H.: Characterization of two bacterial hydroxynitrile lyases with high similarity to cupin superfamily proteins, *Appl. Environ. Microbiol.*, **78**, 2053-2055 (2012).
- 80) Bradford, M. M.: A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding, *Anal. Biochem.*, **72**, 248-254 (1976).
- 81) Sinclair, G. and Choy, F. Y.: Synonymous codon usage bias and the expression of human glucocerebrosidase in the methylotrophic yeast, *Pichia pastoris*, *Protein Expr. Purif.*, **26**, 96-105 (2002).
- 82) Ishida, Y., Kuwahara, Y., Dadashipour, M., Ina, A., Yamaguchi, T., Morita, M., Ichiki, Y. and Asano, Y.: A sacrificial millipede altruistically protects its swarm using a drone blood enzyme, mandelonitrile oxidase, *Sci. Rep.*, **6**, 26998 (2016).
- 83) Tian, G., Xiang, S., Noiva, R., Lennarz, W. J. and Schindelin, H.: The crystal structure of yeast protein disulfide isomerase suggests cooperativity between its active sites, *Cell*, **124**, 61-73 (2006).
- 84) Zhao, G. J., Yang, Z. Q. and Guo, Y. H.: Cloning and expression of hydroxynitrile lyase gene from *Eriobotrya japonica* in *Pichia pastoris*, *J. Biosci. Bioeng.*, **112**, 321-325 (2011).
- 85) Chang, S.-W., Lee, G.-C. and Shaw, J.-F.: Codon Optimization of *Candida rugosa* lip1 gene for improving expression in *Pichia pastoris* and biochemical characterization of the purified recombinant LIP1 lipase, *J. Agric. Food Chem.*, **54**, 815-822 (2006).

- 86) Gao, Z., Li, Z., Zhang, Y., Huang, H., Li, M., Zhou, L., Tang, Y., Yao, B. and Zhang, W.: High-level expression of the *Penicillium notatum* glucose oxidase gene in *Pichia pastoris* using codon optimization, *Biotechnol. Lett.*, **34**, 507-514 (2012).
- 87) Hu, S., Li, L., Qiao, J., Guo, Y., Cheng, L. and Liu, J.: Codon optimization, expression, and characterization of an internalizing anti-ErbB2 single-chain antibody in *Pichia pastoris*, *Protein Expr. Purif.*, **47**, 249-257 (2006).
- 88) Kuse, M., Yanagi, M., Tanaka, E., Tani, N. and Nishikawa, T.: Identification of a fluorescent compound in the cuticle of the train millipede *Parafontaria laminata* armigera, *Biosci., Biotechnol., Biochem.*, **74**, 2307-2309 (2010).
- 89) Hashimoto, M., Kaneko, N., Ito, M. T. and Toyota, A.: Exploitation of litter and soil by the train millipede *Parafontaria laminata* (Diplopoda: Xystodesmidae) in larch plantation forests in Japan, *Pedobiologia*, **48**, 71-81 (2004).
- 90) Kabsch, W.: Xds, *Acta Crystallogr., Sect. D: Biol. Crystallogr.*, **66**, 125-132 (2010).
- 91) Evans, P.: Scaling and assessment of data quality, *Acta Crystallogr., Sect. D: Biol. Crystallogr.*, **62**, 72-82 (2006).
- 92) Vagin, A. and Teplyakov, A.: MOLREP: an automated program for molecular replacement, *J. Appl. Crystallogr.*, **30**, 1022-1025 (1997).
- 93) Emsley, P., Lohkamp, B., Scott, W. G. and Cowtan, K.: Features and development of Coot, *Acta Crystallogr., Sect. D: Biol. Crystallogr.*, **66**, 486-501 (2010).
- 94) Murshudov, G. N., Vagin, A. A. and Dodson, E. J.: Refinement of macromolecular structures by the maximum-likelihood method, *Acta Crystallogr., Sect. D: Biol. Crystallogr.*, **53**, 240-255 (1997).
- 95) Yamaguchi, T., Kuwahara, Y. and Asano, Y.: A novel cytochrome P450, CYP3201B1, is involved in (*R*)-mandelonitrile biosynthesis in a cyanogenic millipede, *FEBS Open Bio*, **7** (3), 335-347 (2017).
- 96) Eisner, T. and Meinwald, J.: Defensive secretions of arthropods, *Science*, **153**, 1341-1350 (1966).
- 97) Dadashipour, M., Ishida, Y., Yamamoto, K. and Asano, Y.: Discovery and molecular and biocatalytic properties of hydroxynitrile lyase from an invasive millipede, *Chamberlinius hualienensis*, *Proceedings of the National Academy of Sciences*, **112**, 10605-10610 (2015).
- 98) Glieder, A., Weis, R., Skranc, W., Poechlauer, P., Dreveny, I., Majer, S., Wubbolts, M., Schwab, H. and Gruber, K.: Comprehensive step-by-step engineering of an (*R*)-hydroxynitrile lyase for large-scale asymmetric synthesis, *Angew. Chem. Int. Ed. Engl.*, **42**, 4815-4818 (2003).
- 99) Bauer, M., Griengl, H. and Steiner, W.: Parameters influencing stability and activity of a *S*-hydroxynitrile lyase from *Hevea brasiliensis* in two-phase systems, *Enzyme Microb. Technol.*, **24**, 514-522 (1999).
- 100) Jansen, I., Woker, R. and Kula, M.-R.: Purification and protein characterization of hydroxynitrile lyases from sorghum and almond, *Biotechnol. Appl. Biochem.*, **15**, 90-99 (1992).
- 101) Kiernan, J. A.: Histological and histochemical methods: theory and practice, *Shock*, **12**, 479 (1999).
- 102) Dombkowski, A. A., Sultana, K. Z. and Craig, D. B.: Protein disulfide engineering, *FEBS Lett.*, **588**, 206-212 (2014).
- 103) Ueatrongchit, T., Kayo, A., Komeda, H., Asano, Y. and A, H. K.: Purification and characterization of a novel (*R*)-hydroxynitrile lyase from *Eriobotrya japonica* (Loquat), *Biosci Biotechnol Biochem*, **72**, 1513-1522 (2008).

- 104) Wajant, H., Forster, S., Selmar, D., Effenberger, F. and Pfizenmaier, K.: Purification and characterization of a novel (*R*)-mandelonitrile lyase from the fern *Phlebodium aureum*, *Plant Physiol.*, **109**, 1231-1238 (1995).
- 105) Kuroki, G. W. and Conn, E. E.: Mandelonitrile lyase from *Ximenia americana* L.: stereospecificity and lack of flavin prosthetic group, *Proc. Natl. Acad. Sci. U. S. A.*, **86**, 6978-6981 (1989).
- 106) van Langen, L. M., van Rantwijk, F. and Sheldon, R. A.: Enzymatic hydrocyanation of a sterically hindered aldehyde. Optimization of a chemoenzymatic procedure for (*R*)-2-chloromandelic acid, *Org. Process Res. Dev.*, **7**, 828-831 (2003).
- 107) Motojima, F., Izumi, A., Nuylert, A., Zhai, Z., Y. Dadashpour, M., Shichida, S., Yamaguchi, T., Nakano, S. and Asano, Y.: *R*-hydroxynitrile lyase from the cyanogenic millipede, *Chamberlinius hualienensis* -A new entry to the carrier protein family Lipocalines -, *FEBS J.*, published online on 17 July, 2020.
- 108) Rask, L., Anundi, H. and Peterson, P. A.: The primary structure of the human retinol-binding protein, *FEBS letters*, **104**, 55-58 (1979).
- 109) Drayna, D., Fielding, C., McLean, J., Baer, B., Castro, G., Chen, E., Comstock, L., Henzel, W., Kohr, W. and Rhee, L.: Cloning and expression of human apolipoprotein D cDNA, *J. Biol. Chem.*, **261**, 16535-16539 (1986).
- 110) Xu, N. and Dahlbäck, B.: A novel human apolipoprotein (apoM), *J. Biol. Chem.*, **274**, 31286-31290 (1999).
- 111) Quarmby, R., Nordens, D., Zagalsky, P., Ceccaldi, H. and Daumas, R.: Studies on the quaternary structure of the lobster exoskeleton carotenoprotein, crustacyanin, *Comp. Biochem. Physiol. B, Biochem. Mol. Biol.*, **56**, 55-61 (1977).
- 112) Paesen, G., Adams, P., Harlos, K., Nuttall, P. and Stuart, D.: Tick histamine-binding proteins: isolation, cloning, and three-dimensional structure, *Mol. cell*, **3**, 661-671 (1999).
- 113) Urade, Y., Fujimoto, N. and Hayaishi, O.: Purification and characterization of rat brain prostaglandin D synthetase, *J. Biol. Chem.*, **260**, 12410-12415 (1985).
- 114) Grzyb, J., Latowski, D. and Strzałka, K.: Lipocalins—a family portrait, *J. Plant Physiol.*, **163**, 895-915 (2006).
- 115) Flower, D. R.: The lipocalin protein family: structure and function, *Biochem. J.*, **318**, 1-14 (1996).
- 116) Kumasaka, T., Aritake, K., Ago, H., Irikura, D., Tsurumura, T., Yamamoto, M., Miyano, M., Urade, Y. and Hayaishi, O.: Structural basis of the catalytic mechanism operating in open-closed conformers of lipocalin type prostaglandin D synthase, *J. Biol. Chem.*, **284**, 22344-22352 (2009).
- 117) Pei, J. and Grishin, N. V.: PROMALS3D: multiple protein sequence alignment enhanced with evolutionary and three-dimensional structural information, *Methods Mol. Biol.*, **1079**, 263-271 (2014).
- 118) Robert, X. and Gouet, P.: Deciphering key features in protein structures with the new ENDscript server, *Nucleic acids Res.*, **42**, W320-W324 (2014).
- 119) Rasputnig, G. and Leis, H.-J.: Wearing a raincoat: exocrine secretions contain anti-wetting agents in the oribatid mite, *Liacarus subterraneus* (Acari: Oribatida), *Exp. App. Acarol.*, **47**, 179-190 (2009).
- 120) Kuwahara, Y., Ichiki, Y., Morita, M., Tanabe, T. and Asano, Y.: Chemical polymorphism in defense secretions during ontogenetic development of the millipede *Niponia nodulosa*, *J. Chem. Ecol.*, **41**, 15-21 (2015).

- 121) Kovats, v. E.: Gas-chromatographische charakterisierung organischer verbindungen. Teil 1: retentionsindices aliphatischer halogenide, alkohole, aldehyde und ketone, *Helv. Chim. Acta*, **41**, 1915-1932 (1958).
- 122) Bodner, M. and Raspotnig, G.: Millipedes that smell like bugs:(E)-alkenals in the defensive secretion of the julid diplopod *Allajulus dicentrus*, *J. Chem. Ecol.*, **38**, 547-556 (2012).
- 123) Kuwahara Y, Mori N, Sakuma M and Tanabe T: (1Z)-and (1E)-2-nitroethenylbenzenes, and 2-nitroethylbenzene as natural products in defense secretions of a millipede *Thecodesmus armatus* Miyosi (Polydesmida: Pyrgodesmidae), *Jpn. J. Environ. Entomol. Zool.*, **14**, 149-155 (2003).
- 124) Ishida, Y., Kuwahara, Y., Dadashipour, M., Ina, A., Yamaguchi, T., Morita, M., Ichiki, Y. and Asano, Y.: A sacrificial millipede altruistically protects its swarm using a drone blood enzyme, mandelonitrile oxidase, *Sci. Rep.*, **6**, 26998 (2016).
- 125) Shear, W. A.: The chemical defenses of millipedes (Diplopoda): biochemistry, physiology and ecology, *Biochem. Syst. Ecol.*, **61**, 78-117 (2015).
- 126) Makarov, S. E.: Diplopoda-Integument. In: Minelli A (ed) Treatise on Zoology-Anatomy, Taxonomy, Biology, *The Myriapoda*, **2** (Brill, 2015).
- 127) Schulz, S.: Composition of the silk lipids of the spider *Nephila clavipes*, *Lipids*, **36**, 637-647 (2001).
- 128) Weatherston, J., Tyrrell, D. and Percy, J.: Long chain alcohol acetates in the defensive secretion of the millipede *Blaniulus guttulatus*, *Chem. Phys. Lipids*, **7**, 98-100 (1971).
- 129) Vujisić, L. V., Antić, D. Ž., Vučković, I. M., Sekulić, T. L., Tomić, V. T., Mandić, B. M., Tešević, V. V., Ćurčić, B. P., Vajs, V. E. and Makarov, S. E.: Chemical defense in millipedes (Myriapoda, Diplopoda): do representatives of the family Blaniulidae belong to the 'quinone' clade?, *Chem. Biodiv.*, **11**, 483-490 (2014).
- 130) Huth, A.: Defensive secretions of millipedes: more than just a product of melting point decrease?, *Frag. Faun. Warszawa*, **43**, 191-200 (2000).
- 131) Shimizu, N., Kuwahara, Y., Yakumaru, R. and Tanabe, T.: *n*-Hexyl laurate and fourteen related fatty acid esters: new secretory compounds from the julid millipede, *Anaulaciulus* sp, *J. Chem. Ecol.*, **38**, 23-28 (2012).
- 132) Brückner, A., Stabentheiner, E., Leis, H.-J. and Raspotnig, G.: Chemical basis of unwettability in Liacaridae (Acari, Oribatida): specific variations of a cuticular acid/ester-based system, *Exp. App. Acarol.*, **66**, 313-335 (2015).

APPENDICES

Appendix I. Primers used for PCR in chapter I

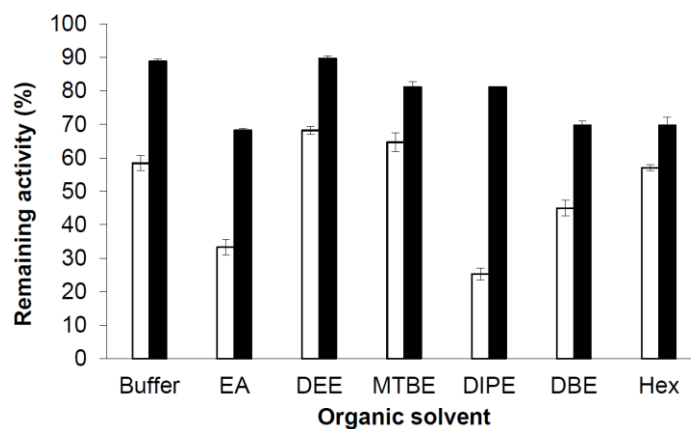
Primer	Sequence (5'-3')	Usage
D	AA(C/T)CC(A/C/G/T)CC(A/C/G/T)GA(A/G)AT(A/C/T)G	Degenerated primer from N-
1	CCCCGAGATGAAGGAATGGAAGTGGGGTAC	Amplified for the 5' and 3' end and the full-length coding <i>PeHNL</i> region.
2	CGATGTCTAGATTGGATCTT GGACATACCC	
3	GAAACCATCAGCTCTGTCTCCACCGCCGG	
4	CACTGTGAAATGTCGATTCATATGCATGTG	
5	GCAATGT TTTGCAGGTGCCATAGTCTGCG	
6	P-5'-CCATCAGCTCTGTC-3' (P indicates 5' phosphorylated end.)	
7	CTCAAGAGAGAGTGACTCTGTGAGAGTACGTACG	
8F	ATATGAGCTCAACCTCCGGAGATAGT	Amplified for ligation with pColdI
8R	CTGCTCTAGATCAGGTGACTACGTATTTG	
9F	ATATGAATTCATGAGAAAACCCAGGAGAGACATT	Amplified for ligation with pPICZαA
9R	CGCGGGCCCTCAGGTGACTACGTATTTGTTTGG	
10F	CTACATAATCCAGGAGACTTTTCCATAC	Removed the <i>N</i> -glycosylation site N105Q
10R	GTATGGAAAAGTCTCCTGGATTATGTAG	

Appendix II. Primers used for PCR in chapter II

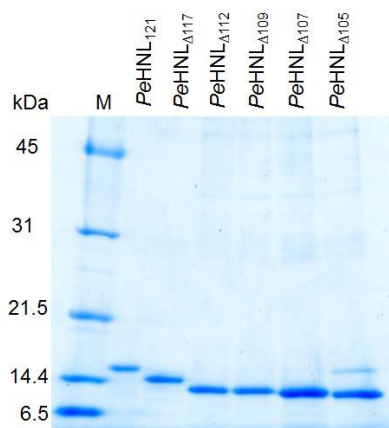
Primer	Sequence (5'-3')	Usage
Fw- <i>PeHNL</i> -Np	CATAATCAACGAGGCTTTTCCATACACATG	<i>PeHNL</i> from purple passion fruit
Rv- <i>PeHNL</i> -Np	CATGTGTATGAAAAAGCCTCGTTGATTATG	
Fw- <i>PeHNL</i> -K117	GCACTTCCAAACTGATACGTAGTCACCTGA	<i>PeHNL</i> Δ ₁₁₇
Rv- <i>PeHNL</i> -K117	TCAGGTGACTACGTATCAGTTTGAAGTGC	
Fw- <i>PeHNL</i> -W112	TTTCCATACACATGAGCACTTCCAAACAAA	<i>PeHNL</i> Δ ₁₁₂
Rv- <i>PeHNL</i> -W112	TTTGTGTTGAAGTGCTCATGTGTATGGAAA	
Fw- <i>PeHNL</i> -P109	TCAACGAGACTTTTTAGTACACATGGGCAC	<i>PeHNL</i> Δ ₁₀₉
Rv- <i>PeHNL</i> -P109	GTGCCCATGTGTACTAAAAAGTCTCGTTGA	
Fw- <i>PeHNL</i> -T107	CATAATCAACGAGTGATTTCCATACACATG	<i>PeHNL</i> Δ ₁₀₇
Rv- <i>PeHNL</i> -T107	CATGTGTATGAAAATCACTCGTTGATTATG	
Fw- <i>PeHNL</i> -N105	TCAACTACATAATCTAGGAGACTTTTCCAT	<i>PeHNL</i> Δ ₁₀₅
Rv- <i>PeHNL</i> -N105	ATGAAAAAGTCTCCTAGATTATGTAGTTGA	
Fw- <i>PeHNL</i> -Y102	CGTCGTGCTCAACTAGATAATCAACGAGAC	<i>PeHNL</i> Δ ₁₀₂
Rv- <i>PeHNL</i> -Y102	GTCTCGTTGATTATCTAGTTGAGCACGACG	

PeHNL-Ny	1	NPPEIVRHIVFNRYKSQLSQKIDQIIADYGNLQNIAPEMKEWKWGTDLG	50
PeHNL-Np	1	NPPEIVRHIVFNRYKSQLSQKIDQIIADYGNLQNIAPEMKEWKWGTDLG	50
PeHNL-Ny	51	PAVEDRADGFTHAYESTFHSVADFLNFFYSPPALEFAKEFFPACEKIVVL	100
PeHNL-Np	51	PAVEDRADGFTHAYESTFHSVADFLNFFYSPPALEFAKEFFPACEKIVVL	100
PeHNL-Ny	101	NYIINE <u>ET</u> FPYTWALPNKYVVT	121
PeHNL-Np	101	NYIINE <u>EA</u> FPYTWALPNKYVVT	121

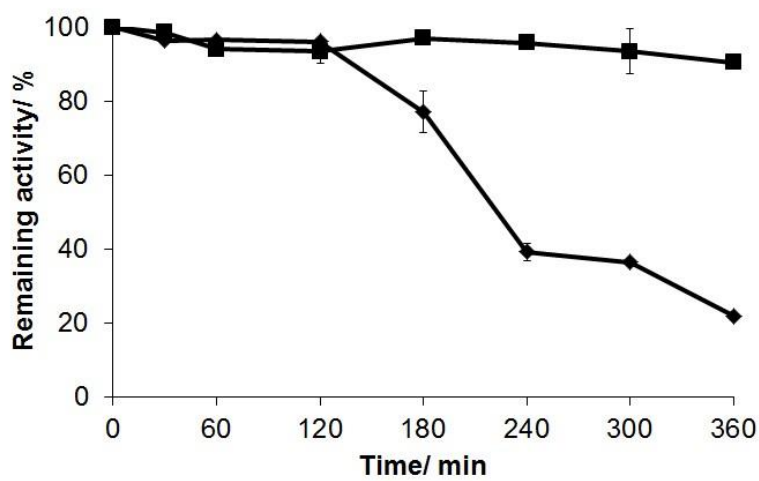
Appendix III. Sequence alignment of native yellow (*PeHNL-Ny*) and purple *PeHNL* (*PeHNL-Np*). Underlined amino acid residues is the consensus sequences for the *N*-glycosylation site on N105 (●).



Appendix IV. Effect of organic solvents on the stability of purified *PeHNL-Ny* (solid bar) and *PeHNL-Np* (open bar). After 12 h incubation at 10°C in each biphasic system at a citrate buffer/organic solvent ratio of 50:50, remaining activities of each *PeHNL* were measured by HPLC with a chiral column.



Appendix V. SDS-PAGE analysis of purified of C-truncated *PeHNL* various.



Appendix VI. Time course of remaining activity of wild type *PeHNL*₁₂₁ (◆) and *PeHNL* Δ ₁₀₇ (■) in biphasic systems at a citrate buffer/organic solvent ratio of 50:50.

Appendix VII. Oligonucleotide primers used in chapter III.

Primer name	5'–3' sequence
His–ChuaHNL-Fw	GCGCTCGAGAAAAGAGCACATCATCATCATCATCATGAAAATTTA TACTTCCAAGGGTCACTGACTTGTGATCAACTTCCC
His–ChuaHNL-Rv	CGCTCTAGATTAGTAAAAAGCAAAGCAACCGTGGGTTTC
ChuaPDI1-InFu-Fw	TCGAAACGAGGAATTCACCATGAAAGTTATCCACTAAATTATTAGC
ChuaPDI1-InFu-Rv	TGTCTAAGGCGAATTCCTATAGTTCATCTTTGCTCGG
ChuaPDI2-InFu-Fw	TCGAAACGAGGAATTCACCATGAAAGTTACAACAATGAAGTCG
ChuaPDI2-InFu-Rv	TGTCTAAGGCGAATTCCTTACAATCCGTCTTGTCGCTC
PpPDI-InFu-Fw	TCGAAACGAGGAATTCACCATGCAATTCAACTGGGATATT
PpPDI-InFu-Rv	TGTCTAAGGCGAATTCCTTAAAGCTCGTCGTGAGCGTC
N99Q-Syn-Fw	TTGTTGACTGACCAGGGTTCCTCCTACTG
N99Q-Syn-Rv	CAGTAGGAGGAACCCTGGTCAGTCAACAA
N109Q-Syn-Fw	TGTGCTTACAGATGTCAGGGAAGTCTGAGATC
N109Q-Syn-Rv	GATCTCAGTTCCTGACATCTGTAAGCACA
N123Q-Syn-Fw	TCCAACAACCAGGGTACTGACCCATTG
N123Q-Syn-Rv	CAATGGGTCAGTACCCTGGTTGTTGGA

```

OriChuaHNL   L   T   C   D   Q   L   P   K   A   A   I   N   P   I   Q   E   F   I   D   S
SynChuaHNL   T-- --- --- --C --G T-G --A --G --- --- --C --C --A --C --A --G --C --C --C 60

OriChuaHNL   N   P   L   E   F   E   Y   V   L   T   E   T   F   E   C   T   T   R   I   Y
SynChuaHNL   AAT CCT TTG GAA TTC GAG TAC GTT CTG ACT GAA ACC TTC GAA TGC ACC ACT CGA ATT TAT 120
SynChuaHNL   --C --A --G --G --- --- --- --- T-- --- --G --T --- --G --T --T --- A-- --C --C 120

OriChuaHNL   V   Q   P   A   R   W   S   T   T   K   A   P   T   A   L   D   I   K   G   T
SynChuaHNL   GTG CAA CCT GCT CGC TGG TCC ACT ACC AAA GCC CCA ACT GCA TTG GAC ATT AAA GGA ACT 180
SynChuaHNL   --T --G --A --- A-A --- --- --- --T --G --T --- --- --- --T --- --C --G --T --- 180

OriChuaHNL   Q   I   M   A   Y   D   F   V   G   G   P   E   N   S   A   H   L   N   E   C
SynChuaHNL   CAA ATT ATG GCT TAC GAT TTC GTC GGT GGT CCT GAA AAC TCA GCT CAC CTC AAC GAA TGC 240
SynChuaHNL   --G --C --- --- --- --C --- --- --- --A --- --- --- --T --- --- T-G --- --- --T 240

OriChuaHNL   H   T   G   D   K   Q   V   W   Y   F   Q   Y   T   N   L   L   T   D   N   G
SynChuaHNL   CAT ACA GGA GAT AAA CAA GTT TGG TAC TTT CAA TAT ACC AAT CTG TTA ACA GAC AAT GGA 300
SynChuaHNL   --C --T --T --C --G --G --- --- --- --G --C --T --C T-- --G --T --- --C --T 300

OriChuaHNL   S   S   Y   C   A   Y   R   C   N   G   T   E   I   I   E   Y   K   C   A   S
SynChuaHNL   AGT TCC TAT TGC GCT TAC AGA TGC AAC GGC ACC GAA ATA ATT GAG TAC AAA TGC GCT TCA 360
SynChuaHNL   TCC --- --C --T --- --- --- --- --T --- --A --T --G --C --- --- --G --T --- --C 360

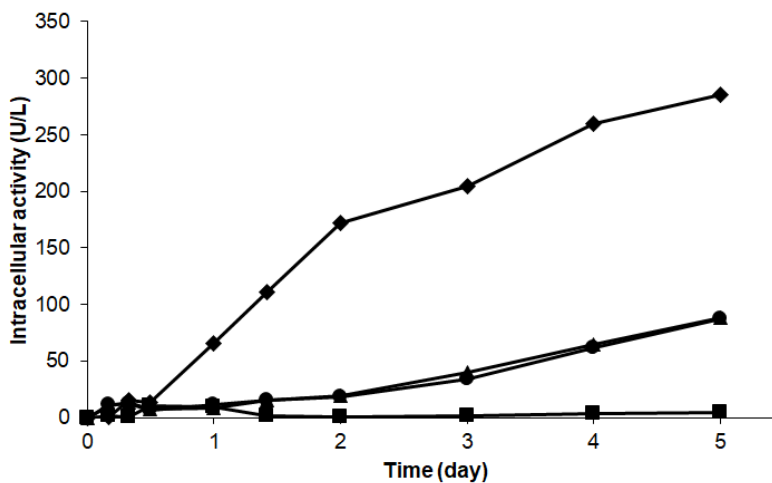
OriChuaHNL   N   N   N   G   T   D   P   L   Q   H   Q   A   M   E   V   A   K   T   V   P
SynChuaHNL   AAC AAT AAC GGA ACT GAT CCA CTC CAA CAC CAA GCG ATG GAA GTA GCA AAA ACA GTT CCA 420
SynChuaHNL   --- --C --- --T --- --C --- --- --- T-G --- --- --- --G --T --- --- --T --G --- --- 420

OriChuaHNL   N   G   D   K   I   H   Y   A   K   S   N   C   P   E   T   H   G   C   F   A
SynChuaHNL   AAC GGC GAC AAG ATT CAT TAT GCC AAA TCA AAT TGC CCC GAA ACC CAC GGT TGC TTT GCT 480
SynChuaHNL   --- --T --- --- --C --C --C --T --G --T --C --T --A --G --T --- --- --T --C --- 480

OriChuaHNL   F   Y   Stop
SynChuaHNL   TTT TAC TAA 489
SynChuaHNL   --C --- --- 489

```

Appendix VIII. Sequence alignment of original ChuaHNL cDNA and synthetic ChuaHNL DNA. The changed codons of SynChuaHNL were labeled out.



Appendix IX. Effects of glycosylation on intracellular His-ChuaHNL production in BMM medium. His-ChuaHNL-WT (diamonds), His-ChuaHNL-N99Q (triangles), His-ChuaHNL-N109Q (squares), and His-ChuaHNL-N123Q (circles).

Appendix X. Purification summary of PlamHNLs.

Step	Total protein	Total activity	Specific activity	Yield	Purification
PlamHNL-N					
Crude (800 millipedes)	28	1120	40	100	1
Butyl -Toyopearl	2.67	574	215	51.3	5.38
Mono Q	0.33	303	897	27	22.4
Resource PHE	0.04	52.6	1320	4.7	33
PlamHNL-E					
Cell-free extract (1L)	361	6890	19.1	100	1
Ni Sepharose	1.75	3840	2190	55.7	115
PlamHNL-P					
Supernatant (1L)	27.5	351	12.7	100	1
Butyl-Toyapearl	0.16	336	2100	95	165

Appendix XI. Effect of inhibitors and metal ions on synthesis activity of PlamHNLs

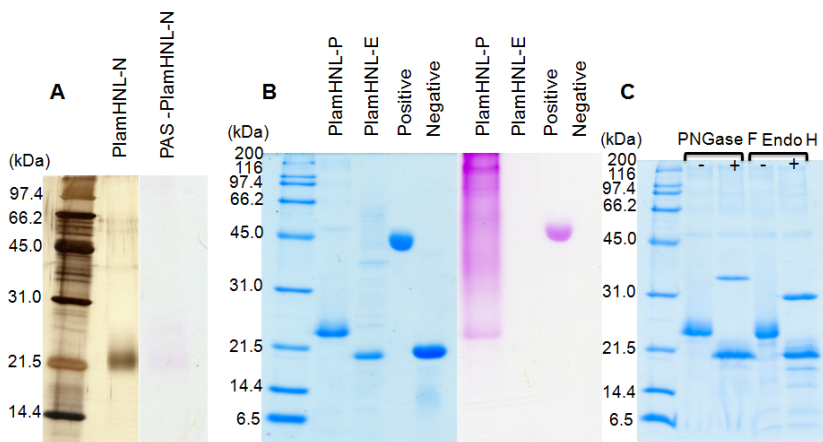
Metals	Remaining activity (%)	
	PlamHNL-E	PlamHNL-P
CaCl ₂	100±2.3	96.6±1.0
MgSO ₄	73.1±0.5	84.1±2.6
MnSO ₄	73.0±0.7	94.0±0.8
ZnSO ₄	88.8±1.2	79.8±1.1
CdCl ₂	79.1±0.3	80.2±0.6
CuSO ₄	79.4±2.1	80.5±2.5
KCl	81.1±0.6	82.2±2.4
FeCl ₃	79.7±4.2	80.8±0.5
LiCl	65.9±0.3	66.8±0.4
CoCl ₂	88.3±1.5	89.4±0.1
NaF	84.2±0.9	85.3±1.5
Sulphydryl reagents		
Iodoacetamide (10 mM)	26.5±0.6	30.2±1.5
Dithionitrobenzoic acid (10 mM)	69.8±1.7	66.8±0.8
Iodoacetic acid (10 mM)	18.3±0.6	21.3±1.3
His modifiers		
Diethylpyrocarbonate	68.4±0.5	62.5±0.8
Serine protease inhibitors		
Trypsin inhibitor (T-9378)	79.2±1.8	86.9±0.6
Trypsin inhibitor (T-9003)	93.0±2.6	80.0±2.8
Trypsin-chymotrypsin inhibitor (T-9777)	85.8±0.8	84.6±0.8
PMSF	81.0±2.2	89.3±1.5
Aspartic protease inhibitor		
Pepstatin A (100 µg/mL)	79±0.8	90.8±2.6
Reducing agents		
2-Mercaptoethanol (10 mM)	83.5±0.8	90.8±0.5
DTT	100±1.2	100±0.9
Biotin reagents		
Biotin	87.3±0.0	95.8±0.9
Avidin	90.0±1.2	90.0±1.6
Chelating reagent		
EDTA	100±0.0	100±0.0

Enzyme (5 U) was used in the reaction. The concentration of applied metal ion was 1.0 mM except where indicated.

Appendix XII. Statistics for data collection and refinement

Data set	Ligand-free	Benzaldehyde (Soaking)	Benzaldehyde&thiocyanate
PDB entry	7BOW	7BPO	7BR1
Date for data collection	5/29/2018	5/29/2018	7/4/2018
X-ray source	KEK-PF BL-5A	KEK-PF BL-5A	KEK-PF BL-5A
Wavelength (Å)	1	1	1
Space group	$P2_1$	$P2_1$	$P2_1$
Unit cell dimensions (Å)	$a = 68.24$ $b = 33.15$ $c = 72.07$	$a = 68.30$ $b = 33.13$ $c = 72.02$	$a = 56.93$ $b = 32.76$ $c = 84.63$
(°)	$\alpha = \gamma = 90.00,$ $\beta = 106.23$	$\alpha = \gamma = 90.00,$ $\beta = 106.17$	$\alpha = \gamma = 90.00,$ $\beta = 102.84$
Resolution range (Å) *	42.07 – 1.42 (1.50 – 1.42)	42.10 – 1.37 (1.44 – 1.37)	41.94 – 1.25 (1.32 – 1.25)
Total No. of observed reflections *	390,845 (55,634)	420,603 (48,071)	556,553 (77,278)
No. of unique reflections *	59,066 (8,551)	63,260 (8,085)	84,891 (12,324)
Mean $I/\sigma(I)$ *	12.0 (2.2)	11.7 (2.3)	18.1 (2.1)
R_{merge} *	0.081 (0.790)	0.082 (0.731)	0.053 (0.839)
Mn (I) half-set correlation CC (1/2)	0.998 (0.797)	0.997 (0.834)	1.000 (0.770)
Completeness (%) *	100.0 (100.0)	96.3 (85.5)	100.0 (100.0)
Redundancy *	6.6 (6.5)	6.6 (5.9)	6.6 (6.3)
Refinement statistics			
$R_{\text{work}}/R_{\text{free}}$	0.152/ 0.178	0.159/ 0.176	0.140/ 0.148
Number of atoms			
- Protein	2,569	2,579	2,594
- Ligands	12	28	52
- Waters	430	434	459
R.m.s. deviations			
Bond lengths (Å)	0.006	0.006	0.006
Bond angles (°)	1.404	1.376	1.357

* Values in parentheses are for the highest-resolution shells.



Appendix XIII. SDS-PAGE analysis of purified HNL from millipedes (PlamHNL-N) (A) and the recombinant PlamHNLS, the one expressed in *E. coli* (PlamHNL-E) and *P. pastoris* (PlamHNL-P) stained by Coomassie blue and by periodic acid Schiff (PAS) (B), treated with *N*-glycosidase F from *Flavobacterium meningosepticum* and endoglycosidase H from *Streptomyces plicatus* (C)

```

      10      20      30      40      50      60
M T V A A L I V M L A E L G W A Q P S P
ATGACTGTTG CTGCACTGAT CGTTATGTTG GCCGAATTGG GCTGGGCTCA ACCTTCACCT

      70      80      90      100     110     120
L T C D K L P K V I P P G I S A F T S H
CTCACTTGGG ACAAGTCCC AAAAGTCATC CCACCTGGCA TTAGTGCTTT CACTTCCCAC

      130     140     150     160     170     180
N P F E F S Y V L T N D L D C T A R V Y
AATCCITTTG AATTCTGTA TGTGTTGACT AACGATCTCG ACTGTACCGC ACGAGTCTAC

      190     200     210     220     230     240
V Q P V H G L T N Y S G T A F D I K G T
GTACAGCCTG TACATGGACT GACCAATTAC AGTGGAACTG CATTGACAT CAAAGGAACT

      250     260     270     280     290     300
H I T I N D F T I G A D G L T A Y L T N
CACATAACAA TAAATGACTT CACCATTGGT GCCGATGGTC TGACAGCCTA ITTGACTAAT

      310     320     330     340     350     360
C D T D V K Q V W H F Q Y V D L G D P Q
TGTGATACTG ACGTAAAACA GGTTTGGCAT TTTCRAATAG TCGACCTAGG TGATCCCCAA

      370     380     390     400     410     420
G A N Y C A Y Y C E G P E I V E Y K C T
GGTGCCAACT AATTGCGATA CTATTGCGAA GGTCGCCGAAA TAGTGAATA CAAATGCACT

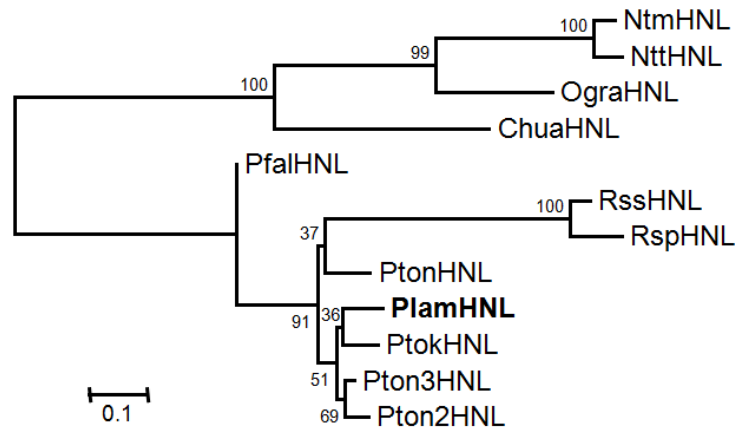
      430     440     450     460     470     480
T N T G Y I S P R Q L Q A V K E A Q S V
ACGAATACTG GATACATATC GCCTCGACAA CTCAGGCTG TAAAAGAGGC ACAATCAGTC

      490     500     510     520     530     540
P N G D K I H P A Q V N C P P H L Y C P
CCAAATGGTG ACAAGATTCA TCCAGCCGAG GTCAATTGCC CTCCTCACCT TTACTGTCCC

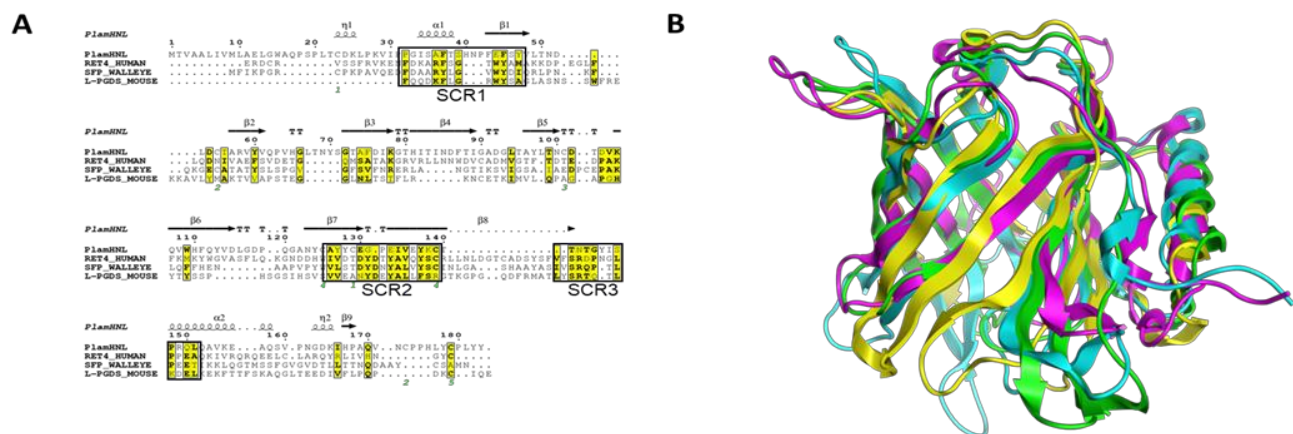
      550     560     570     580     590     600
L Y Y *
CTCTATTACT AA

```

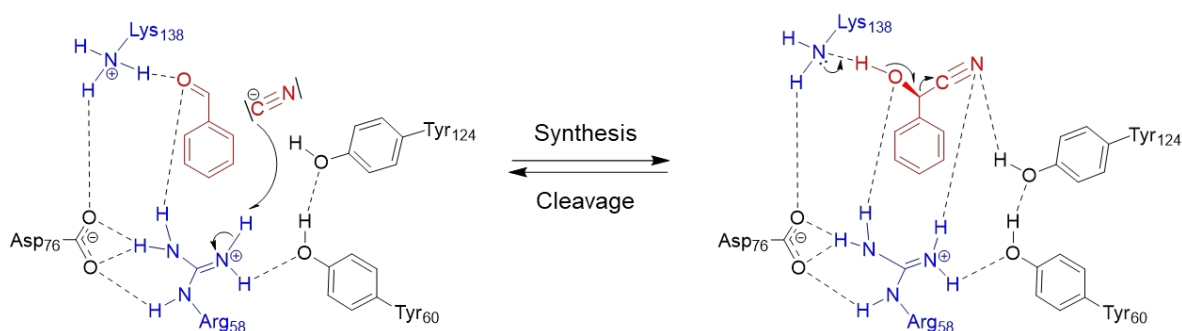
Appendix XIV. cDNA and deduced amino acid sequences of hydroxynitrile lyase from the millipedes, *Parafontaria laminata*. Underline indicates a 30 amino acid-long signal peptide. A box indicates predicted *N*-glycosylation sites.



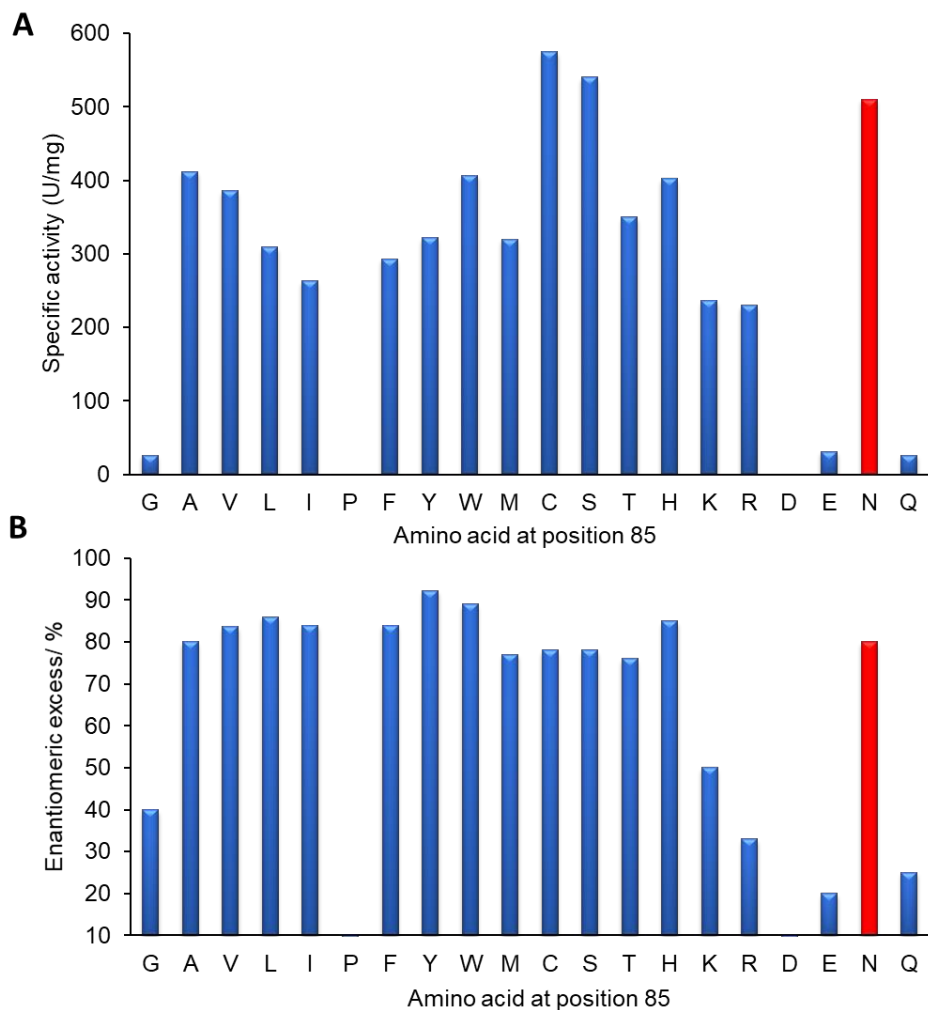
Appendix XV. Phylogenetic tree of HNLs from millipedes. The phylogenetic tree was performed using the neighbor-joining method in MEGA 6.0 with 1,000 bootstrap trials. The number at nodes indicate bootstrap value percentages. A bar indicated a 10% divergence.



Appendix XVI. A) The secondary structure-based multiple sequence alignment among PlamHNL and the similarity Dali Z-score¹ protein structure comparison of typical lipocalins, such as walleye sandercyanin fluorescent protein (SFP_WALLEYE PDB ID: 5E2Z, Dali Z-score: 14.2), human retinol binding protein 4 (RET4_HUMAN, PDB ID: 1JYD, Dali Z-score: 13.9), and mouse lipocalin-type prostaglandin D (L-PGDS_MOUSE, PDB ID: 2CZU, Dali Z-score: 11.6). The three motifs structurally conserved region (SCRs 1-3) are indicated by black boxes. The secondary structure elements are shown as follows; α -helices, medium squiggles with α symbols; 3_{10} -helices, squiggles with η symbols; β -sheets, arrows with β symbols; strict β -turns, TT letters. Similar groups are highlighted with yellow boxes. B) The superposed structures of monomer PlamHNL (cyan), RET4_HUMAN (blue), SFP_WALLEYE (green) and L-PGDS_MOUSE (yellow) are shown. Superposition with secondary structure matching was evaluated using MOE program.



Appendix XVII. Proposed catalytic mechanism for PlamHNL. Cyanohydrin synthesis of (R)-mandelonitrile based on complex-crystal structure and our previous report (Motojima et. al., 2020).



Appendix XVIII. HNL specific activities (A) and enantiomeric excess (B) for (*R*)-2-chloromandelonitrile synthesis of mutants generated by saturation mutagenesis at Asn85. Specific activity was assayed after 5 min in a reaction mixture containing 2-chlorobenzaldehyde (50 mM) and KCN (100 mM) at 25°C with the same amount (0.64 μg) of purified PlamHNL variants. Enantiomeric excess was obtained for (*R*)-2-chloromandelonitrile synthesis by HPLC with the same concentration of purified PlamHNL variants (1.2 Unit).

Appendix XIX. Primers used for PCR in Chapter 3 section 2.

Primer	Sequence (5'-3')
1F	TGACWTGYGATMAACTTCC
1R	TARTARAAAGCAAARCAACCG
2	ATGAATCTTGTACCGTTTGGAACTGATCG
3	GAGTTGTTTAGGCGATATGTATCCAGTATTC
4	ATTCAAGGAACTCACATAACAATAAATGACTTC
5	GTGCCGATGGTCTGACAGCCTATTTG
6	ATGACTTCGATCATTCTCCTCATGACTG
7	GCTTAATTCAATTGCACTTTAATTTTTATATC
8F	ATATGGATCCCTCACTTGCGACAAGCTCCC
8R	CTGCAAGCTTTTAGTAATAGAGGGGACAGTA
9F	gtctcgatcgggtaccCTCACTTGCGACCAGCTC
9R	gagttttgtctagaTTAGTAATAGAGGGGACAGT
10	catcaccatcaccatcacCTCACTTGCGACCAGCTC
11	GGTACCGATCCGAGACGGCCGGCTG
PlamHNL-R58K-Fw	CTCGACTGTACCGCAAAGTCTACGTACAGCCT
PlamHNL-R58K-Rv	AGGCTGTACGTAGACTTTTGCGGTACAGTCGAG
PlamHNL-R58W-Fw	CTCGACTGTACCGCATGGGTCTACGTACAGCCT
PlamHNL-R58W-Rv	AGGCTGTACGTAGACCCATGCGGTACAGTCGAG
PlamHNL-Y60A-Fw	ACCGCACGAGTCGCTGTACAGCCTGTACAT
PlamHNL-Y60A-Rv	ATGTACAGGCTGTACAGCGACTCGTGCGGT
PlamHNL-A74W-Fw	AATTACAGTGGAACCTGGTTTGACATCAAAGGA
PlamHNL-A74W-Rv	TCCTTTGATGTCAAACCAAGTTCCTACTGTAATT
PlamHNL-N85H-Fw	ACTCACATAACAATACACGACTTCACCATTGGT
PlamHNL-N85H-Rv	ACCAATGGTGAAGTCGTGATTGTTATGTGAGT
PlamHNL-N85Y-Fw	ACTCACATAACAATATACGACTTCACCATTGGT
PlamHNL-N85Y-Rv	ACCAATGGTGAAGTCGTATATTGTTATGTGAGT
PlamHNL-N85E-Fw	ACTCACATAACAATAGAAGACTTCACCATTGGT
PlamHNL-N85E-Rv	ACCAATGGTGAAGTCTTCTATTGTTATGTGAGT
PlamHNL-N85Q-Fw	ACTCACATAACAATACAAGACTTCACCATTGGT
PlamHNL-N85Q-Rv	ACCAATGGTGAAGTCTTGTATTGTTATGTGAGT
PlamHNL-N85I-Fw	ACTCACATAACAATAATCGACTTCACCATTGGT
PlamHNL-N85I-Rv	ACCAATGGTGAAGTCGATTATTGTTATGTGAGT
PlamHNL-N85L-Fw	ACTCACATAACAATACTCGACTTCACCATTGGT
PlamHNL-N85L-Rv	ACCAATGGTGAAGTCGAGTATTGTTATGTGAGT
PlamHNL-N85M-Fw	ACTCACATAACAATAATGGACTTCACCATTGGT
PlamHNL-N85M-Rv	ACCAATGGTGAAGTCCATTATTGTTATGTGAGT
PlamHNL-N85F-Fw	ACTCACATAACAATATTCGACTTCACCATTGGT
PlamHNL-N85F-Rv	ACCAATGGTGAAGTCGAATTATTGTTATGTGAGT
PlamHNL-N85W-Fw	ACTCACATAACAATATGGGACTTCACCATTGGT
PlamHNL-N85W-Rv	ACCAATGGTGAAGTCCCATTATTGTTATGTGAGT
PlamHNL-N85V-Fw	ACTCACATAACAATAGTTGACTTCACCATTGGT
PlamHNL-N85V-Rv	ACCAATGGTGAAGTCAACTATTGTTATGTGAGT
PlamHNL-N85R-Fw	ACTCACATAACAATACGTGACTTCACCATTGGT
PlamHNL-N85R-Rv	ACCAATGGTGAAGTCACGTATTGTTATGTGAGT
PlamHNL-N85K-Fw	ACTCACATAACAATAAAGGACTTCACCATTGGT
PlamHNL-N85K-Rv	ACCAATGGTGAAGTCCTTTATTGTTATGTGAGT
PlamHNL-N85S-Fw	ACTCACATAACAATATCTGACTTCACCATTGGT
PlamHNL-N85S-Rv	ACCAATGGTGAAGTCAGATTATTGTTATGTGAGT
PlamHNL-N85T-Fw	ACTCACATAACAATAACTGACTTCACCATTGGT
PlamHNL-N85T-Rv	ACCAATGGTGAAGTCAGTTATTGTTATGTGAGT
PlamHNL-N85C-Fw	ACTCACATAACAATATGTGACTTCACCATTGGT
PlamHNL-N85C-Rv	ACCAATGGTGAAGTCACATATTGTTATGTGAGT
PlamHNL-N85G-Fw	ACTCACATAACAATAGGTGACTTCACCATTGGT

PlamHNL-N85G-Rv	ACCAATGGTGAAGTCACCTATTGTTATGTGAGT
PlamHNL-N85P-Fw	ACTCACATAACAATACCTGACTTCACCATTGGT
PlamHNL-N85P-Rv	ACCAATGGTGAAGTCAGGTATTGTTATGTGAGT
PlamHNL-N85D-Fw	ACTCACATAACAATAGATGACTTCACCATTGGT
PlamHNL-N85D-Rv	ACCAATGGTGAAGTCATCTATTGTTATGTGAGT
PlamHNL-L98K-Fw	GGTCTGACAGCCTATAAGACTAATTGTGATACT
PlamHNL-L98K-Rv	AGTATCACAATTAGTCTTATAGGCTGTCAGACC
PlamHNL-L98W-Fw	GGTCTGACAGCCTATTGGACTAATTGTGATACT
PlamHNL-L98W-Rv	AGTATCACAATTAGTCCAATAGGCTGTCAGACC
PlamHNL-L98E-Fw	GGTCTGACAGCCTATGAACTAATTGTGATACT
PlamHNL-L98E-Rv	AGTATCACAATTAGTTTCATAGGCTGTCAGACC
PlamHNL-W109H-Fw	GGCGAAAAACAGGTTTCATCATTTTCAATATGTC
PlamHNL-W109H-Rv	GACATATTGAAAATGATGAACCTGTTTTTCGCC

PUBLICATIONS

1. A. Nuylert, Y. Ishida, and Y. Asano. Effect of glycosylation on the biocatalytic properties of hydroxynitrile lyase from the passion fruit, *Passiflora edulis*: A comparison of natural and recombinant enzymes. *ChemBioChem* **18** (3), 257-265 (2017).
2. A. Nuylert, F. Motojima, C. Khanongnuch, T. Hongpattarakere and Y. Asano. Stabilization of hydroxynitrile lyases from two variants of passion fruit, *Passiflora edulis* Sims and *Passiflora edulis* Forma *flavicarpa*, by C-terminal truncation. *ChemBioChem*, **21**(1-2), 181-189 (2020)
3. Z. Zhai*, A. Nuylert*, K. Isobe, and Y. Asano. Effects of codon optimization and glycosylation on the high-level production of hydroxynitrile lyase from *Chamberlinius hualienensis* in *Pichia pastoris*, *J. Ind. Microbiol. Biotechnol.*, **46** (7), 887-898 (2019).
*The authors have contributed equally to this work.
4. A. Nuylert, Y. Kuwahara, T. Hongpattarakere, and Y. Asano. Identification of saturated and unsaturated 1-methoxyalkanes from the Thai millipede *Orthomorpha communis* as potential “Raincoat Compounds”. *Sci. Rep.* 8: 11730. (2018)
5. A. Nuylert, M. Nakabayashi, T. Yamaguchi, and Y. Asano. Discovery and structural basis to improve the enzyme activity and enantioselectivity of hydroxynitrile lyase from *Parafontaria laminata* millipedes for (*R*)-2-chloromandelonitrile synthesis. *ACS Omega* **43** (5), 27896-27908. (2020)

The articles are not included in the dissertation:

1. F. Motojima, A. Nuylert, and Y. Asano. The crystal structure and catalytic mechanism of hydroxynitrile lyase from passion fruit, *Passiflora edulis*, *FEBS J.*, 285 (2), 313-324 (2017)
2. T. Yamaguchi, A. Nuylert, A. Ina, T Tanabe, and Y. Asano. Hydroxynitrile lyases from cyanogenic millipedes: molecular cloning, heterologous expression, and whole-cell biocatalysis for the production of (*R*)-mandelonitrile, *Sci. Rep.* 8: 3051. (2018)
3. F. Motojima, A. Izumi, A. Nuylert, Z. Zhai, M. Dadashipour, S. Shichida, T. Yamaguchi, S. Nakano, and Y. Asano. *R*-hydroxynitrile lyase from the cyanogenic millipede, *Chamberlinius hualienensis* -A new entry to the carrier protein family Lipocalines, *FEBS J.*, published online on 17 July, 2020.