Matthias Diepold, BSc

Structural investigations
on the hydroxynitrile lyase
from Davallia tyermanii

MASTER'S THESIS

to achieve the university degree of

Master of Science

Master's degree programme: Molecular Microbiology

submitted to

Graz University of Technology

Supervisor

Univ.-Prof. Dr. Karl Gruber

Institute of Molecular Biosciences
University of Graz

Dr. Tea Pavkov-Keller
ACIB GmbH, Graz

Graz, January 2016

AFFIDAVIT

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

19.01.2016
Date

Signature

Acknowledgements

Foremost I would like to offer my sincerest gratitude to my two extraordinary supervisors, Dr. Karl Gruber and Dr. Tea Pavkov-Keller who supported me with knowledge, ideas, enthusiasm and most of all patience whilst allowing me the room to work in my own way.

I attribute the level of my Master's degree to their encouragement and effort and without them this thesis, too, would not have been completed or written. I could not wish for better or friendlier supervisors and their intellectual and professional influence will follow me wherever I may go.

I would also like to thank all the members of the Gruber workgroup and in general the all members of the structural biology department for their guidance, moral support and the one or other words of cheer.

Furthermore, I would like to thank the Anton Glieder workgroup for the endless support of protein and the spadework which made this project possible.

Table of contents

Table of figures

Abbreviations

Å	Angstrom
ADH	Alcohol dehydrogenase
BEZ	Benzoic acid
C	Cysteine
CN	Cyanide
CPD	Carboxypeptidase
D	Aspartic acid
DtHNL	Hydroxynitrile lyase from *Davallia tyermanii*
E	Glutamic acid
F	Phenylalanine
FAD	Flavin adenine dinucleotide
GMC oxidoreductase	Glucose-methanol-choline oxidoreductase
GtHNL	Hydroxynitrile lyase from *Granulicella tundricola*
HbHNL	Hydroxynitrile lyase from *Hevea brasiliensis*
HCN	Hydrogencyanide
HNL	Hydroxynitrile lyase
I	Isoleucine
L	Leucine
LuHNL	Hydroxynitrile lyase from *Linum usitatissimum*
M	Methionine
MeHNL	Hydroxynitrile lyase from *Manihot esculenta*
MSE	Selenomethionine
MXN	(R)-mandelonitrile
NAD	Nicotinamide adenine dinucleotid
P	Phenylalanine
PaHNL	Hydroxynitrile lyase from *Prunus amygdalus*
R	Arginine
S	Serine
SAD	Single anomalous diffraction
SbHNL	Hydroxynitrile lyase from *Sorghum bicolor*

T	Threonine
V	Valine
W	Tryptophane
Y	Tyrosine

Abstract of the thesis

Hydroxynitrile lyases are involved in a widely spread defence mechanism in the flora, the cyanogenic pathway. During maceration of plant cells, cyanogenic glycosides, which are amongst others stored in vacuoles, are catabolized by the enzymes hydroxylnitrile lyases and β-glucosidase. The result is the release of hydrogen cyanide (HCN). [1] In industrial biocatalysis, hydroxynitrile lyases are used to catalyse the stereospecific formation of α-cyanohydrins from hydrogen cyanide and ketones and aldehydes. [2]

A new hydroxynitrile lyase has recently been identified in the fern *Davallia tyermanii* (DtHNL). An efficient expression system for this enzyme was established and the biochemical characterization showed that DtHNL is a highly active, (R)-selective hydroxynitrile lyase. [3] Here I report the structural characterisation of this new HNL.

Five crystal structures were presented (native DtHNL, selenomethionine derivate, native DtHNL with substrate, inhibitor and product in active site) and DtHNL exhibits a Bet V1 fold. As DtHNL does not show any sequence or structural similarity to any other hydroxynitrile lyase and does not possess HNL typical conserved motifs, therefore DtHNL represent a new class of HNLs. Based on the obtained structures and the analysis of possible proton distributions in the active site of the enzyme three possible mechanisms were proposed.

Zusammenfassung

Hydroxynitrillyasen sind an der Cyanogenese einem, in der Pflanzenwelt weit verbreiteten Verteidigungsmechanismus beteiligt. In den Vakuolen der Pflanzen werden cyanogene Glykoside gespeichert welche von Hydroxynitrillyasen und β-Glucosidasen abgebaut werden und als Resultat dieses Abbaus wird Blausäure (HCN) freigesetzt. [1] In der industriellen Biokatalyse werden Hydroxynitrillyasen genutzt um stereospezifisch α-Cyanhydrine aus Blausäure und Ketonen oder Aldehyden herzustellen. [2]

Kürzlich wurde eine neue Hydroxynitrillyase im Farn *Davallia tyermanii* (DtHNL) entdeckt. Es wurde ein effizientes Expressionssystem etabliert und die biochemische Charakterisierung zeigte, dass es sich bei DtHNL um eine hoch aktive, (R)-selektive Hydroxynitrillyase handelt. [3] In dieser Masterarbeit wird die strukturelle Charakterisierung dieser neuen HNL beschrieben.

Es wurden fünf Kristallstrukturen gelöst (natives DtHNL, ein Selenomethionin-Derivat, natives DtHNL je mit Substrat, Inhibitor und Produkt im aktiven Zentrum). Dabei stellte sich heraus, dass DtHNL eine Bet V1 Faltung besitzt. Da die Hydroxynitrillyase von *Davallia tyermanii* keine Sequenz- und Strukturähnlichkeiten zu anderen HNLs zeigt und auch keine HNL typischen konservierten Motive besitzt, repräsentiert DtHNL eine neue Klasse von HNLs. Basierend auf den erhaltenen Strukturen und der möglichen Protonenverteilung im aktiven Zentrum konnten drei mögliche katalytische Mechanismen vorgeschlagen werden.

1   Introduction

Hydroxynitrile lyases (HNLs, EC 4.1.2.10, EC 4.1.2.11, EC 4.1.2.46, and EC 4.1.2.47) are enzymes which catalyse the enantioselective cleavage of cyanohydrins to hydrocyanic acid and the associated aldehyde or ketone. (Figure 1)
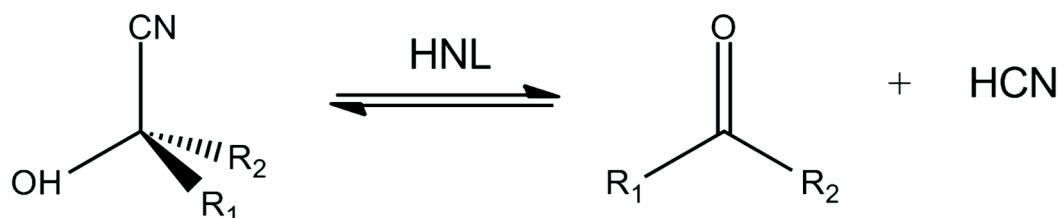
Figure 1. The chemical reaction catalysed by hydroxynitrile lyases.

The first hydroxylnitrile lyase activity was described by Wöhler and Liebig in 1837 as they investigated the formation of hydrocyanic acid also called hydrogen cyanide (HCN). [1]

Like the cleavage of cyanohydrins, the catalysed enzymatic reaction is reversible and the equilibrium is highly dependent on the pH value, solvent composition and accessible concentrations of cyanohydrin, carbonyl compound and HCN. In aqueous solutions, the cleavage of cyanohydrin occurs spontaneously above values of pH 5, when in fact the enzymatic reaction also occurs at lower pH values. [2]

The reverse reaction was observed by Rosenthaler for the first time in 1908. He used an emulsion derived from bitter almonds (*Prunus amygdalus*) called emulsin to prepare (R)-mandelonitrile. [3] Since this time cyanogenesis has been reported in many plant species like Rosaceae, Linaceae, Clusiceae, Gramineae, Olacaceae, Euphorbiaceae and Polypodiacae, but also in other organisms like fungi, arthripods, insects, lichen or bacteria. [4] [5] Interestingly HNLs were also identified in non-cyanogenic plants such as *Arabidopsis thaliana*. [6]

In the last years hydroxylnitrile lyases have come up as potential biocatalysts for the chemical industry. The C-C bond forming condensation emerged as a useful reaction in which a new chiral centre is formed, the carbon chain is prolonged by one carbon atom and the nitrile is added as an additional versatile functional group. Cyanohydrin products are biologically active compounds and can be converted into an assortment of useful starting materials for syntheses, like α-hydroxy carboxylic acids, α-hydroxy ketones or β-amino alcohols. [7] [8] [9] [10]

Biologically the release of hydrogen cyanide serves as a defence mechanism against attacks of herbivores or microorganisms for many plants [2] [11] [12] and could also serve as a nitrogen source for the biosynthesis of the amino acid asparagine. [13] [14]

1.1 HNL folds

Several unmistakably different types of hydroxynitrile lyases exist, they developed from different ancestral proteins which make HNLs an interesting example of convergent evolution [12]. At least five unrelated classes of proteins with hydroxynitrile lyase activity associated to different structure fold families like α/β-hydrolases [15] [16], FAD-dependent oxidoreductases [17] [18], carboxypeptidases [19] zinc-dependent alcohol dehydrogenases [20] and the cupine superfamily [21] have been known so far. In addition to the structural difference between the divergent HNLs, the construction of the active sites is also remarkable different. [22]

1.1.1 Hydroxynitrile lyases related to α/β-hydrolases

The higly similar homologous (S)-enantioselective hydroxynitrile lyases from *Hevea brasiliensis* (HbHNL) and *Manihot esculenta* (MeHNL) represent the best studied members of the α/β-hydrolase superfamily. [22]

The (S)-selective HbHNL is an in aquatic solutions homodimeric protein with 29.2kDa per monomer [23] [24] and shows a high similarity to the HNL of *Manihot esculenta* (76% sequence identity). [25] Acetone cyanohydrin is its natural substrate, but a wide range of

Figure 2. The crystal structure of the HNL from *Hevea brasiliensis* [29] drawn from the cordinates of pdb-entry 1yas.

aliphatic, aromatic and heterocyclic aldehydes or ketones are accepted in vitro. [26] [27] [28]

The crystallographic 3D structure analysis clearly confirmed that the HbHNL belongs to the α/β-hydrolase superfamily with an active site deeply buried in the interior of the protein (Figure 4). [29]

11
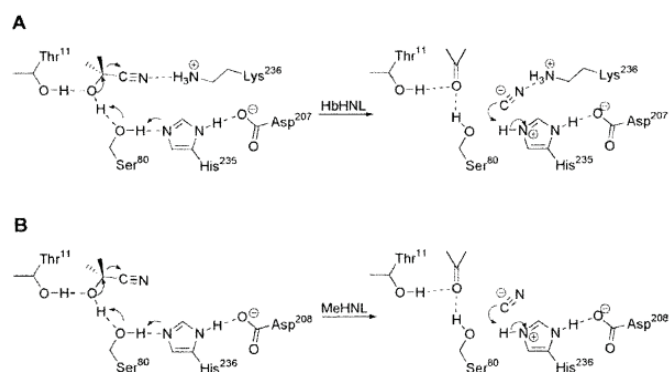
Figure 3. (A) Proposed reaction mechanism for the HNL from *Hevea brasiliensis.* (B) Proposed mechanism for the HNL from *Manihot esculenta*. [22]

Based on sequence homology, modelling studies and structural studies on complexes, a catalytic mechanism was proposed. (Figure 3 A) Therein the catalytic triad residues act as general acid/base with the central histidine residue as the active species. One active site lysine plays also an essential role in the mechanism. The lysine stabilizes the complementary negative charge, build up on the cyano group, with its positive charge. This stabilisation facilitates the cleavage of the cyanohydrin and respectively the deprotonation of hydrogen cyanide in the reverse reaction. [30] [31]

The structural analyses of MeHNL produced similar mechanistic proposals. (Figure 3 B) However the Lysine in the MeHNL active site is not suggested to take part in the catalytic mechanism. But the role of this Lysine is still a controversial issue. [22]

1.1.2 FAD dependent hydroxyl nitrile lyases

FAD dependent HNLs usually occur in Rosaceae. The single chain enzymes, with usually 50-80kDa, shows sequence homology with FAD-dependent oxidoreductases. The natural substrate of this FAD-dependent HNL is (R)-(+)-mandelonitrile. [17] Flavin dependent enzymes usually catalyse redox reactions but the FAD-HNLs do not appear to use the redox characteristics of its cofactor. [32] At least there is no evidence that the FAD participates directly in the catalytic mechanism. [33]

It can be supposed that the remaining FAD cofactor is an evolutionary remnant from an oxidoreductase precursor, in the non-redox reactions performing HNL. [22]

The first 3D structure analysis of a FAD-dependent hydroxynitrile lyase was made with the HNL from *Prunus amygdalus* (PaHNL) by multiple-wavelength anomalous dispersion of a mercury derivative. [34] (Figure 4)

Like the GMC oxidoreductase family PaHNL consists of a FAD-binding and a substrate binding domain. [35] While the FAD-binding domain shows homology to dinucleotide-binding proteins [36] the substrate binding domain is built by a six stranded β-sheet and four surrounding helices.

Based on structure analysis [34] and docking simulations [37] a mechanism for PaHNL was proposed. (Figure 5) In which a histidine acts as a general base and deprotonates the hydroxyl group of the substrate (R)-(+)-mandelonitrile.
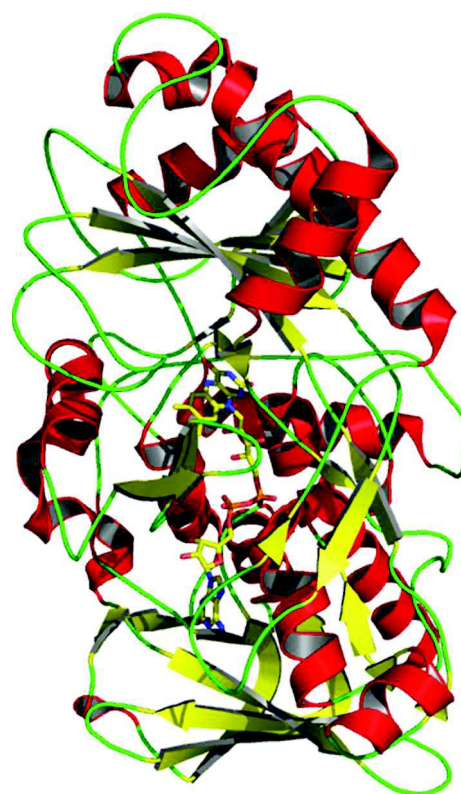
Figure 4. The crystall structure of the HNL from *Prunus amygdalus* [34] drawn from the pdb-entry 1ju2.

Opposing to the HbHNL positively charged residue in direct contact with the substrate was not found.

A positive electrostatic potential at the binding site was generated by a collective effort of charged residues more than 10Å away. [37]
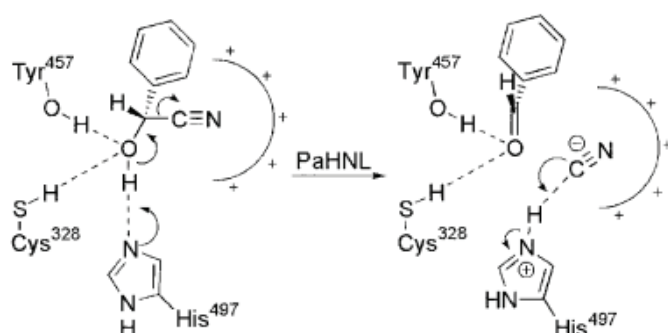
Figure 5. Proposed mechanism for the HNL from *Prunus amygdalus.* [22]

1.1.3 Hydroxynitrile lyases related to carboxypeptidases

The HNL isolated from *Sorghum bicolor* (SbHNL) is a 510 amino acids long polypeptide, which is post-translationally cleaved in one α and one β peptide. The active enzyme consists of two α, β heterodimers and forms a α2β2 heterotetramer. [38] The natural substrate of SbHNL is (S)-4-hydroxymandelonitrile. [39]

The SbHNL and the wheat serine carboxypeptidase (CPD) showed a high sequence identity (60%) but the SbHNL hold a two amino acid deletion directly neighbouring a

CPD catalytic serine residue. [19] This deletion suggests that CPD and SbHNL do not have the same active-site-geometry. [22]

The 3D structure was determined by molecular replacement with wheat CPD as search model. [40] The structure analysis revealed that sequence-identity accordant the heterodimeric structure of SbHNL was very similar to the CPD structure. SbHNL showed a α/β-hydrolase fold which consists of a central eleven stranded β-sheet surrounded by more than a dozen α-helices. [22] (Figure 6)
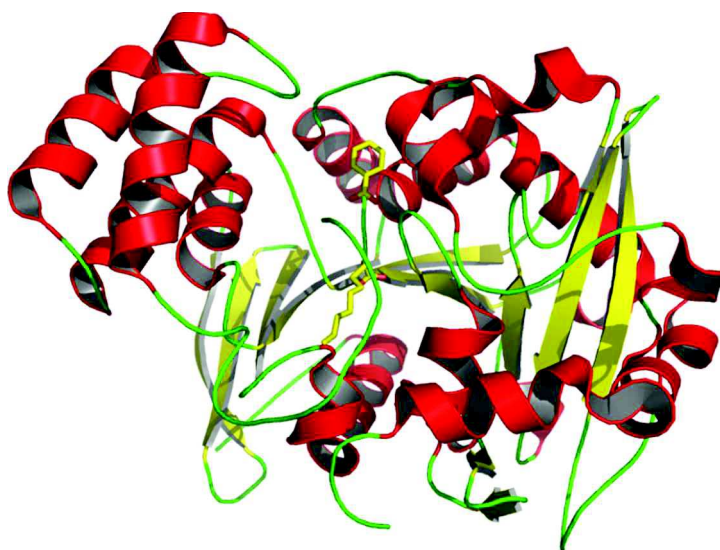
Figure 6. Crystal structure of the HNL from *Sorghum bicolor* [41] drawn from the pdb-entry 1gxs.

Based on modelling simulations of the SbHNL active site with the product and an active site water molecule, implications regarding a possible reaction mechanism were made. (Figure 7)

Figure 7. The proposed reaction mechanism for the hydroxynitrile lyase from *Sorghum bicolor.* [22]

The substrate was predicted to form a hydrogen bond between the hydroxyl group of the cyanohydrin with an active site serine, and with one oxygen atom of the C-terminal carboxylate group of a tryptophan. The water molecule was located between nitrile group of the substrate and the other carboxylate oxygen of the tryptophan residue. This indicated that the carboxylate group of C-terminal tryptophan acts as the catalytic base at first abstracting a proton from the hydroxyl group of the cyanohydrin. Meanwhile it is suggested that the water molecule transfer the abstracted proton to the cyanide following C-C bond cleavage. [41]

1.1.4 Hydroxynitrile lyases related to Zn^{2+} dependent alcohol dehydrogenases

The hydroxynitrile lyase from flax *Linum usitatissimum* (LuHNL) is a homodimer with 87kDa. [42] This HNL showed no activity against 4-hydroxymandelonitrile or mandelonitrile but was inhibited by the natural substrate acetone cyanohydrin in high concentrations. (Figure 8)

The enzyme also acted on numerous aliphatic aldehydes and ketones, but no aromatic carbonyl substrates. [43]
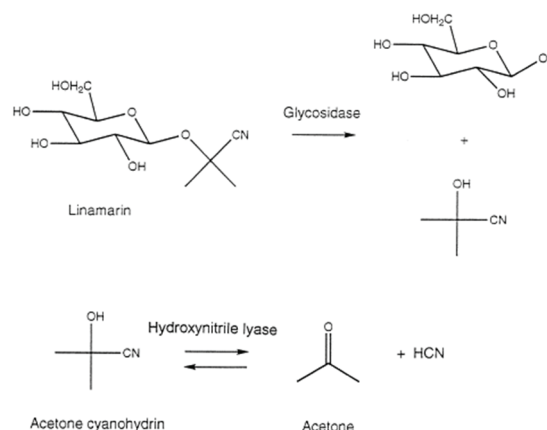
Figure 8. The role of the HNL from *Linum usitatissimum* during cyanogenesis. [47]

Database searches and analyses showed that the hydroxynitrile lyase from *Linum usitatissimum* featured homologies to Zn^{2+} dependent alcohol dehydrogenases. [20] The highest scores of the homology searches [44] gave medium chain alcohol dehydrogenases (ADH) of class I and class III. ADHs from class III are presumed to be the ancestral form, while class I ADHs appears to originate from gene duplication during early vertebrate evolution. [45] [46]

The conserved amino acids found in LuHNL relative to ADHs include several glycines and ligands of the Zn^{2+}, which are catalytically and structurally important. A NAD specifying aspartic acid and a serine and threonine which promote the proton removal from the substrate were also found. [47] To date no 3D structure of LuHNL was published.

1.1.5 Hydroxynitrile lyases related to the cupine superfamily

Most of the so far known hydroxylnitrile lyases have been discovered in plants. A lately discovered HNL was found in endophytic bacteria. [21] It turned out that this HNL was not related to any known HNL but showed sequence homology to cupine superfamily proteins. The many known cupin related proteins stand out for their small β-barrel folds with various functionalities. [48] [49] [50] This cupin related hydroxynitrile lyase presents a novel class of HNLs.

The manganese dependent hydroxynitrile lyase from acidobacterium *Granulicella tundricola* (GtHNL) is one representative of this new class of HNLs. [51]

Many cupin proteins are metal binding and bind Mn^{2+}, Fe^{2+}, Ni^{2+}, Cu^{2+}, Zn^{2+} in the active site. [52] Two conserved motifs are involved in the metal ion binding, G-(X)5-H-X-H-(X)3,4-E-(X)6-G (motif 1) and G-(X)5-P-X-G-(X)2-H-(X)3-N (motif 2). [50]

These motifs are conserved in the GtHNL sequence, indicating that the HNL is also metal dependent. Interestingly the HNL activity is increased in the presence of Mn^{2+} compared to colonies without supplementation. [51]

The 3D analysis revealed that the enzyme contained two homotetrameres in the asymmetric unit. Each monomer contains eleven β-strands, which form a jellyroll β-sandwich, a characteristic cupin-barrel fold (Figure 9).

Figure 9. Crystall strucutre of one homotetramer of the HNL from *Granulicella tundricola* [51] drawn from the pdb-entry 4uxa.

Until now no co-crystallization or soaking with the substrate in the active site was published.

1.2 A new hydroxynitrile lyase from the fern Davallia tyermanii

Very recently a new hydroxynitrile lyase was discovered in the fern *Davallia tyermanii* (DtHNL). This (R)-selective HNL shows due to its high activity and pH stability promising characteristics for possible applications in the chemical industry. An efficient expression system for this enzyme was established and it has been biochemicaly characterised. (Data not shown)

In this master thesis I describe the structural characterisation and mechanistic analysis of DtHNL which, like LuHNL [42] or PaHNL [53] has (R)-mandelonitrile as its natural substrate. However, in contrast to the previously described hydroxynitrile lyase folds, DtHNL was found to be structurally related to the birch pollen allergen Bet V1. To date no structure of this enzyme has been published and no HNL with this specific fold was reported.

2   Materials and methods

The cloning, expression and purification of native DtHNL and the DtHNL selenomethionine derivate was performed by our collaborators at the ACIB GmbH (austrian centre of industrial biotechnology).

2.1   DtHNL cloning

DtHNL isoenzyme 1 was cloned into the vector pEHISTEV (*NcoI/HindIII*) and verified via sequencing through LGC Genomics. *E. coli* BL21- (DE3) Star competent cells were transformed with standard methods. Resulting clones were selected on LB agar plates with kanamycin (50mg L^{-1}). [54]

2.2   Expression of native DtHNL

A single colony of *E. coli* BL21 Star (DE3), transformed with pEHISTEV_isoenzyme1, was incubated overnight in LB medium (also called LB-Lennox) (20g/L: 10g/L Tryptone, 5g/L Yeast extract, 5 g/L NaCl), Carl ROTH, 76185 Karlsruhe, Germany) with 50mg L^{-1} kanamycin at 37°C. After 16h, 500ml LB medium with kanamycin (50mg L^{-1}) were inoculated with an aliquot of the overnight culture to a final OD600 of 0.1. The culture was incubated at 37°C and 150rpm until an OD$_{600}$ of 0.7 was reached. 0.5mM IPTG was added and the culture was incubated at 25°C and 150rpm for 20h. [54]

2.3   Expression of the selenomethionine derivate of DtHNL

A single colony of *E. coli* BL21 Star (DE3), transformed with pEHISTEV_isoenzyme1, was incubated overnight in LB medium with 50mg L^{-1} kanamycin at 37°C. The cell pellet was collected after 16h, and was washed two times with M9 minimal media (M9 salts, 2% glucose, 2mM MgSO$_4$, 0.01mg/mL thiamine, 0.01mg mL^{-1} FeCl$_3$).

400mL minimal media with kanamycin (50mg L^{-1}) and selenomethionine (50mg L^{-1}) was inoculated with an aliquot of the preculture to a final OD$_{600}$ of 0.25. The culture was incubated at 37°C and 150rpm until an OD600 of 0.5 was reached. 0.5mM IPTG was added and the culture was incubated at 25°C and 150rpm for 38h. [54]

2.4 Preparation of cell free extract and protein purification of native DtHNL

Cell pellets were resuspended in 25mL of 50mM potassium phosphate buffer, pH 6 and disrupted by sonication (80% duty cycle, 7 output, 6 minutes). During cell disruption the samples were kept on ice.

The purification of DtHNL was performed by affinity chromatography with a His-Trap (5mL) column using an ÄKTA purifier system. *E. coli* lysate was prepared in 20mM sodium phosphate, 0.5M NaCl, 10mM imidazole pH 7.4 as described above.

Clear lysate (filtered through a 0.45µm syringe filter) was loaded on the, previously with 20mM sodium phosphate, 0.5M NaCl, 10mM imidazole at pH 7.4 (start buffer) equilibrated, column. 10 column volumes of start buffer were used to wash unbound material. A gradient from 0 to 100% elution buffer (20mM sodium phosphate, 0.5M NaCl, 500mM imidazole, pH 7.4) for 15 column volumes was used for elution. A last washing step was performed with 100% elution buffer for 5 column volumes. DtHNL containing fractions were combined and desalted HiPrep 26/10 Desalting (life technologies). The purified protein fractions were stored in 50mM sodium phosphate pH 6.5, at -80°C. [54]

2.5 Preparation of cell free extract and protein purification of SeMet-DtHNL

The cell free extract was prepared as described above. The selenomethionine derivate was purified by affinity chromatography (NiSepharose 6 Fast Flow resin). The elution was performed with 20mM sodium phosphate, 0.5 M NaCl, 300 mM imidazole, pH 7.4. SeMet-DtHNL containing fractions were combined and desalted (PD10 Desalting columns, GE Healthcare, LifeScience). The purified protein fractions were stored in 50mM potassium phosphate pH 6, at -20°C. [54]

2.6 Crystallisation

For crystallisation the buffer was changed to 10mM Tris-HCl pH8 by diluting and reconcentrating the enzyme in Ultra 2mL Centrifugal Filters (Amicon). The final protein concentration of native DtHNL was 4mg ml^{-1} and of SeMet DtHNL was 3mg ml^{-1}. Concentrations were determined by UV absorption measurements at 280nm. Screening for crystallization conditions was performed using Morpheus Screen MD 1-46, JCSG+ MD1-37 (Molecular Dimentions) and Index HT HR2-144 (Hampton

Research) by the sitting drop vapour-diffusion method in 96-well plates using an Oryx8 crystallization robot (Douglas Instruments Ltd.). The drops contained equal volumes of protein and reservoir solution and were equilibrated against 35µL of reservoir solution. The robot worked at 20°C and the crystal plates were stored in an incubator at 16°C.

The best native DtHNL crystals were grown by combining 0.5µL protein sample with 1µL reservoir solution (0.1 M Sodium/potassium phosphate, 25 % v/v 1,2-Propandiol) by the sitting drop vapour diffusion method in Crystal Clear Duo crystallization frames at 16°C. To improve crystal quality micro-seeding with 1:10000 diluted native DtHNL crystals, grown in 0.09M $NaNO_3$, Na_2HPO_4, $(NH_4)_2SO_4$ mix, 0.1M Tris-Bicine pH8 buffer system and 30% (w/v) polyethylene glycol monomethyl ether 550 & polyethylene glycol 20K (Morpheus Screen MD 1-46 condition C9), was applied. The orthorhombic crystals appeared within 2 days and grow to dimensions of about 0.18X0.04X0.006mm (Figure 10 A) in approximately 3 days.

The best diffracting SeMet-DtHNL crystals were grown by combining 0.5µL protein solution (3mg mL^{-1} in 10mM Tris HCl) und 0.5µL reservoir solution composed of 0.2M sodium thiocyanate, 20% (w/v) polyethylene glycol 3350 by the sitting drop vapour diffusion method in Crystal Clear Duo crystallization frames at 16°C. To improve crystal quality micro-seeding with 1:10000 diluted native DtHNL crystals, grown in 0.09M $NaNO_3$, Na_2HPO_4, $(NH_4)_2SO_4$ mix, 0.1M Tris-Bicine pH8 buffer system and 30% (w/v) polyethylene glycol monomethyl ether 550 & polyethylene glycol 20K (Morpheus Screen MD 1-46 condition C9), was applied. The orthorhombic crystals with typical dimensions of 0.06X0.01X0.006 mm (Figure 10 B) appeared within 2-3 days.

After supplementation of 30% glycerol, the crystals were flash cooled in liquid nitrogen.
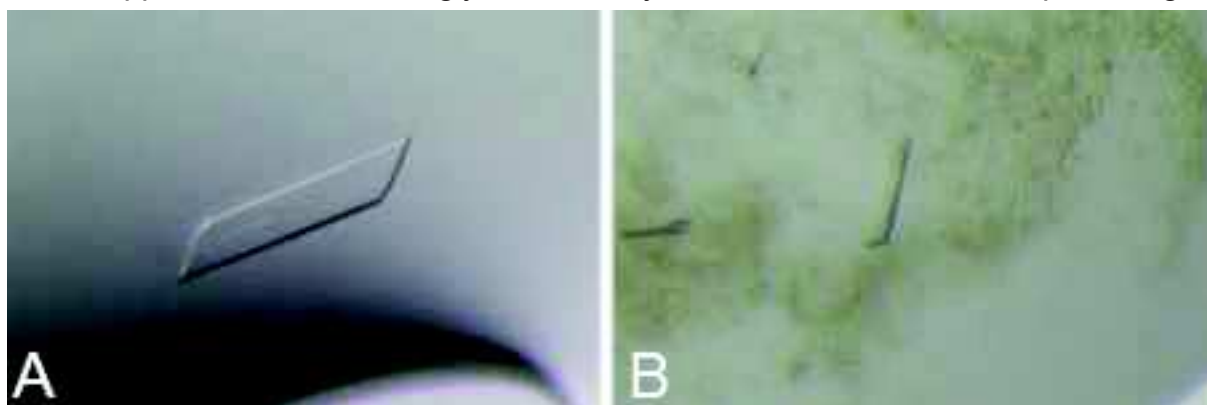
Figure 10. (A) Native DtHNL crystals with dimensions of about 0.18X0.04X0.006mm. (B) Seleno-methionine labelled DtHNL crystals with dimensions of 0.06X0.01X0.006 mm.

2.7 Soaking experiments

Soaking experiments were performed with native DtHNL crystals grown as described above. Crystal Clear Duo crystallisation frames containing mature crystals were opened and neat (R)-mandelonitrile (MXN) (Catalog # 469653, Sigma Aldrich), benzoic acid (BEZ) (catalog # 242381, Sigma Aldrich) was added with a small CryoLoop and thereby the crystallisation solution as well the crystals were saturated with the respective substrate, inhibitor or product.

After an incubation period of 30s, 1min, 5min, and 15min crystals were harvested, flash-cooled in liquid nitrogen and used for diffraction data collection.

2.8 Data collection, processing, model building, and refinement

Diffraction datasets were collected at cryogenic temperatures on synchrotron beamlines at the European Molecular Biology Laboratory (EMBL) in Grenoble, France and at the Elettra Syncrotrone in Trieste, Basovizza, Italy.

As no homologous structure was known at the beginning of this experimental work for this (R)-selective hydroxynitrile lyase, experimental phasing techniques had to be applied. Therefore, a selenomethionine derivate of DtHNL could be crystallized as described above. The selenium substructure of SeMet-DtHNL was suitable for subsequent phasing by single anomalous diffraction (SAD) with eight selenomethionine sites in the asymmetric unit.

The SeMet-DtHNL dataset were processed and scaled using the XDS program package [55]. The AutoSol program [56] and the AutoBuild program [57] from the Phenix software suit [58] were used to define the selenium heavy-metal atom sites using a SeMet-DtHNL SAD dataset, as well as to build an initial model. The resulting model was again completed manually in Coot [59] and refined with Phenix.

The native and soaked DtHNL datasets were processed and scaled using the program iMosflm [60] and SCALA [61]. Molecular replacement was performed with Phaser-MR [62]. The previously obtained SeMet-DtHNL structure was used as model for the molecular replacement with the native and soaked datasets. The resulting models were again completed manually in Coot and refined with Phenix.

Estimations of molecules within the asymmetric unit were done in CCP4 Matthews [63]. 5% randomly chosen reflections, which were not used during refinement cycles, were set aside for the calculation of the R_{free} values. [64]

The details of the data collection, the processing and structure refinement are summarised in Table 1.

Table 1. Data collection and refinement statistics. Values in parentheses are for the highest resolution shell.

	SeMet-DtHNL	Native DtHNL	DtHNL-MXN I	DtHNL-MXN II	DtHNL-BEZ
Data processing statistics					
X-ray source	ID29 (EMBL)	ID23 (EMBL)	XRD1 (Elettra)	BM14U (EMBL)	BM14U (EMBL)
Temperature	100K	100K	100K	100K	100K
Wavelenght [Å]	0.979	0.972	0.972	0.918	0.918
Resolution range	47.01-1.67 (1.73-1.67)	47.31-1.93 (2.00-1.93)	46.91-2.05 (2.12-2.05)	36.72-1.63 (1.69-1.63)	36.5-1.75 (1.81-1.75)
Space group	I222	I222	I222	I222	I222
Cell parameters					
a [Å]	73.64	73.68	73.50	73.44	73.61
b [Å]	94.02	94.63	93.82	94.18	93.86
c [Å]	117.05	117.72	116.48	116.39	116.16
$\alpha=\beta=\gamma$ [°]	90	90	90	90	90
R_{merge}	0.067	0.043	0.186	0.098	0.159
$I/\sigma(I)$	8.0 (1.10)	13.35 (2.21)	5.97 (1.97)	14.02 (1.42)	12.10 (1.23)
Unique reflections	47726 (4604)	28304 (2367)	25637 (2515)	49863 (4499)	40351 (3862)
Completeness [%]	99.74 (97.58)	90.50 (76.26)	99.91 (99.17)	98.71 (90.38)	98.57 (94.94)
Redundancy	2.0	2.0	2.0	5.6	6.6
Unique reflections	47726 (4604)	28304 (2367)	25637 (2515)	49863 (4499)	40351 (3862)
Refinement statistics					
Protein residues	350	349	350	349	349
No. of non H atoms	3072	3059	2934	3029	3076
r.m.s. deviations from ideality					
Bond lengths [Å]	0.011	0.008	0.008	0.007	0.012
Bond angles [°]	1.26	1.07	1.05	1.08	1.27
Ramachandran plot					
favoured [%]	97	96	97	97	97
outliers [%]	0	0	0	0	0
R_{work} [%]	16.45	19.10	18.05	17.89	18.36
R_{free} [%]	19.86	24.43	22.68	20.43	22.7

3 Results

3.1 SeMet-DtHNL

The crystal structure of the DtHNL selenomethionine derivate revealed a symmetric dimer with 8 antiparallel beta sheets and 5 alpha helices in each protomer. The refined structure consists of 2 chains of 3072 non H protein atoms containing 350 of 368 residues of both molecules in the asymmetric unit (Figure 11). Chain A is lacking residues 1 to 9 and 184 and in chain B the residues 1 to 8 are missing. It was not possible to place these residues within the existing electron density map due to missing density. 241 water molecules were placed and 8 selenomethionines at the positions M92, M100, M102 and M114 (each in chain A and B) are present. The refinement of the final model approached to R_{work} = 16.45%, R_{free} = 19.86% for 47726 reflections in the resolution range 47.01-1.67Å. In the Ramachandran plot are 97.5% of the residues in favoured regions, 2.6% are located in allowed regions (Table 1). Alternative conformations for the residues S54A (Chain A), MSE92A (selenomethionine 92 chain A), MSE100A, MSE102, MSE114, E61B, T65B, S81B, MSE92B, MSE102B and MSE114B were made to provide a structure with better electron density map fitting.
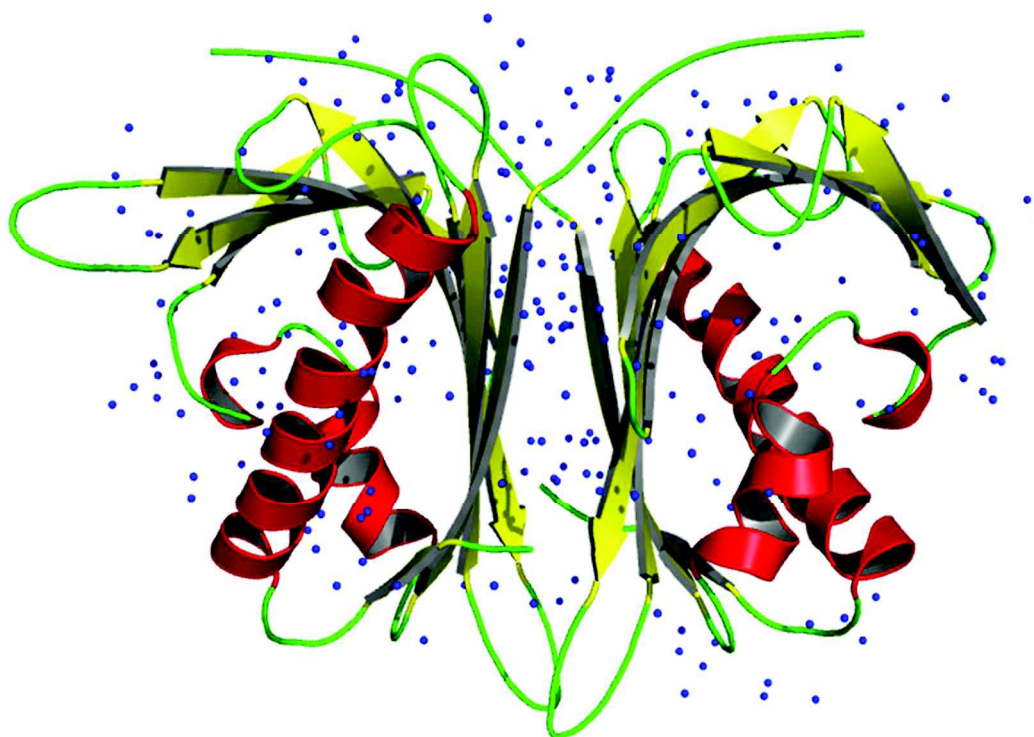
Figure 11. Cartoon illustration of the overall structure of the dimeric SeMet-DtHNL. α-helices are coloured in red, loops in green, β-sheets in yellow and water molecules in blue.

After the refinement the (F(o) - F(c)) density map of the SeMet DtHNL-dataset shows positive density (green) in each cavity of the two molecules. (Figure 12)
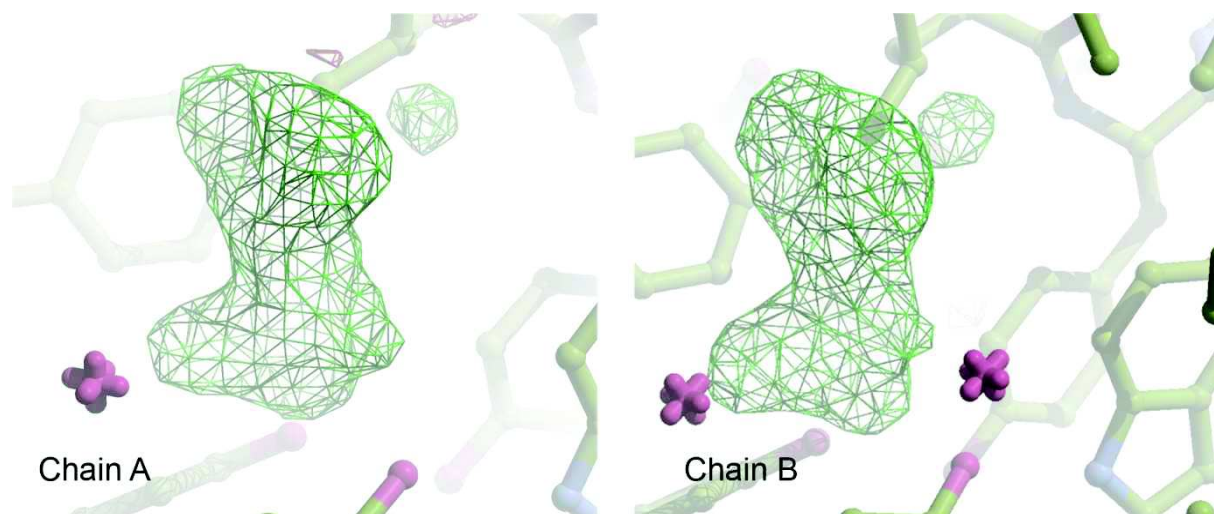
Figure 12. $F_o - F_c$ densities in the active site of the SeMet-DtHNL structure. (left) Chain A; (right) Chain B

The density shows a distinctive form and indicated that a ligand could be present in each cavity.

Several attempts have been made to identify the ligand. For example, the Ligand identification program of the Phenix software suit. But the density is not specific enough to make doubtless assumptions. The ligand could be a remnant from the expression or purification steps and is therefore difficult to identify.

The overall fold consists of an 8-stranded antiparallel β-sheet (β1-8) which bends around a long C-terminal located α-helix (α3). Two sequenced, shorter α-helices (α1-2) separate the C-terminal part of the long α-helix and the β-sheet. On the long helix follows C-terminal a short β-strand and the C-terminus (Figure 13 A).

A CaSoX accessibility analysis revealed a cavity between the β-sheet and the 3 α-helices that extend through the structure with two openings from the active site to the protein surface. The smaller exit tunnel is located between the helices α1 and α2 and the larger tunnel lies between α3 and β3-5. (Figure 13B)
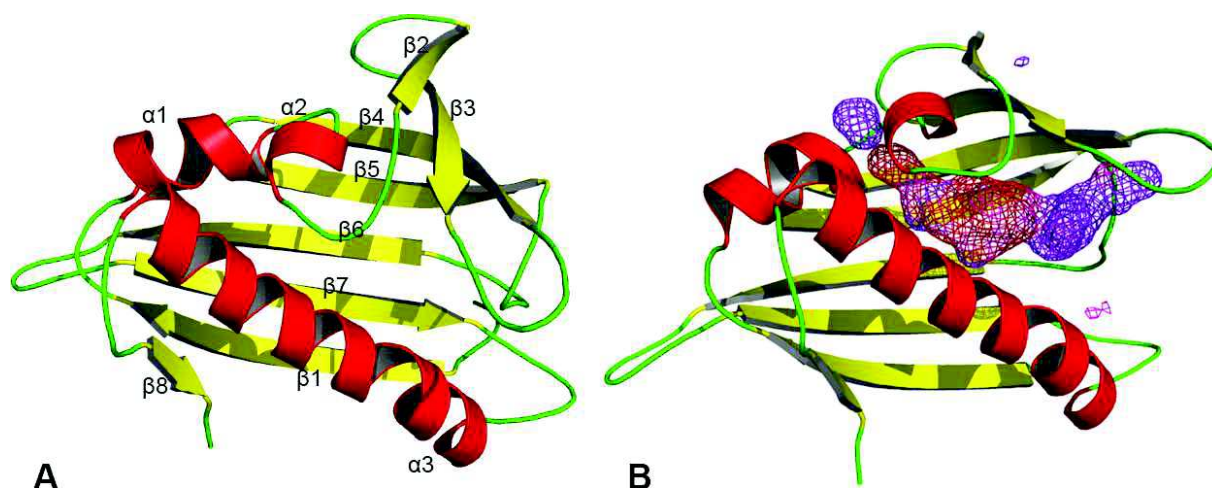
Figure 13. (A) Cartoon representation of a monomer of DtHNL. The major secondary-structural elements are numbered: α-helices (α1-α3) are coloured red, β-strands (β1-β8) are coloured yellow and loops are coloured green. (B) CaSox accessibility analysis of the binding pocket of a SeMet-DtHNL monomer, revealed two tunnels from the active site to the surface.

Structure homology searches exposed that the tertiary fold of the SeMet-DtHNL structure is related to the birch pollen allergen Bet V1 [65]. (Figure 14)
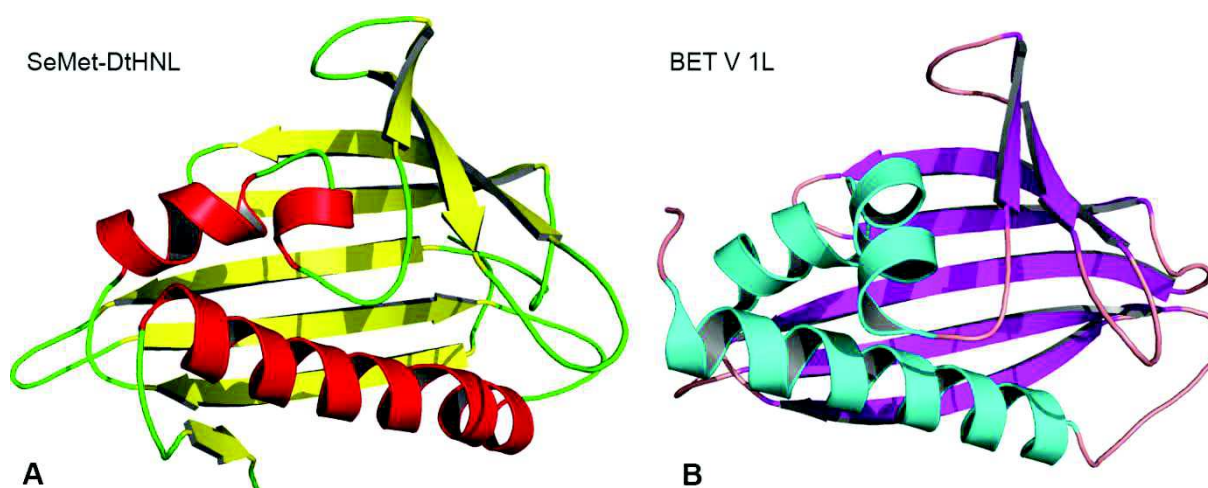

Figure 14. Cartoon illustration of a monomer from SeMet-DtHNL. α-helices are coloured in red, loops in green, β-sheets in yellow. (B) Cartoon illustration of the monomeric birch pollen allergen Bet V1L [65] drawn from the pdb-entry 1FM4. α-helices are coloured in cyan loops in salmon, β-sheets in magenta

3.2   Native DtHNL

The refined structure of the native DtHNL dataset consist of 3059 non H protein atoms containing 349 of 368 residues in both molecules in the asymmetric unit. The residues 1 to 9 and 184 are missing in chain A and chain B is lacking the residues 1 to 9. The existing electron density maps did not provide enough density to place these residues. 315 water molecules were placed. Alternative conformations for

the residues S54A, E61B and T65B were made to gain better electron density map fitting of the structure. Successively performed refinement cycles converged to R factors of R_{work} = 19.10%, R_{free} = 24.43% for 28304 reflections in the resolution range 47.31-1.93 Å. In the Ramachandran plot are 96% of the residues in favoured regions and 4% are located in allowed regions (Table 1). The SeMet-DtHNL structure (see 3.1 SeMet-DtHNL) was used as model for molecular replacement.

Equally to the SeMet-DtHNL structure the overall crystal structure of the native DtHNL consists of an 8-stranded antiparallel β-sheet which bends around a long C-terminal located α-helix and two consecutive shorter α-helices N-terminal located. (Figure 15)

A cavity is located between the β-sheet and the 3 α-helices that extend through the structure with two openings on the protein surface.
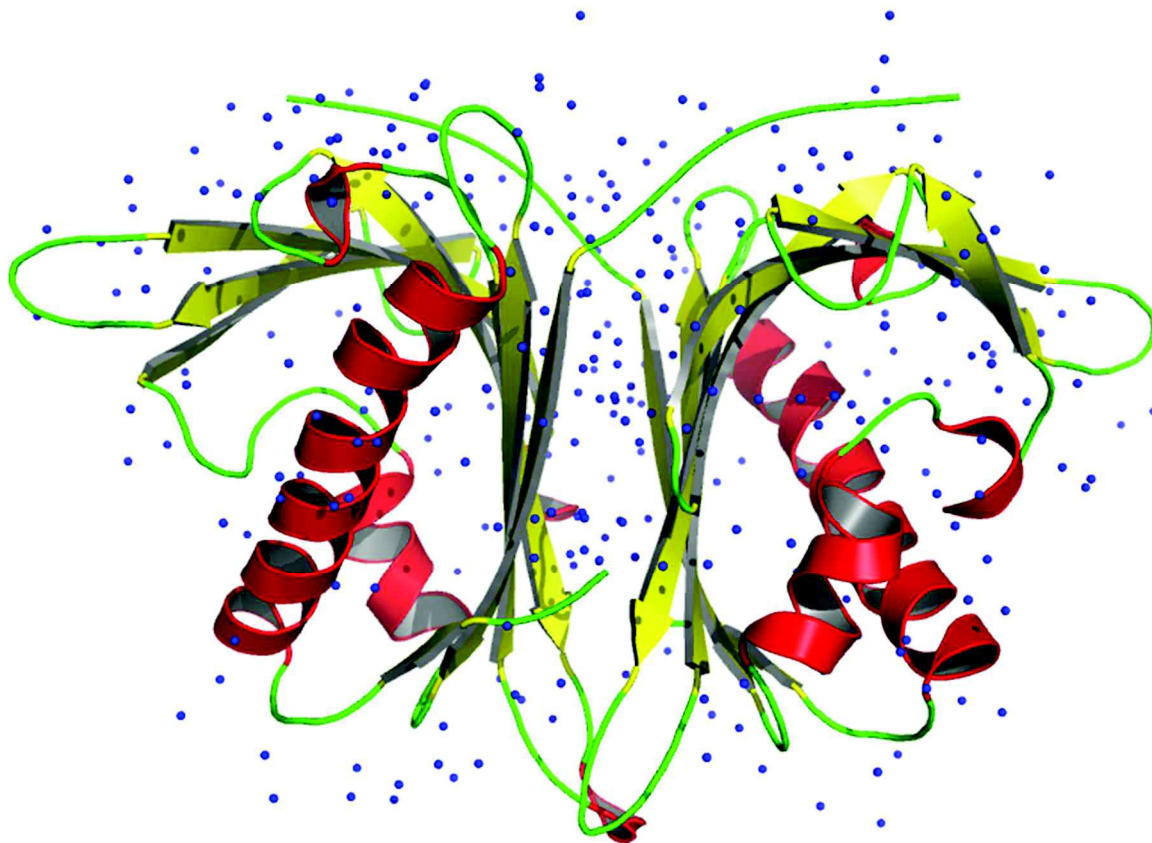
Figure 15. Cartoon illustration of the overall structure of the dimeric native DtHNL. α-helices are coloured in red, loops in green, β-sheets in yellow and water molecules in blue.

The (F(o) - F(c)) density map of the native DtHNL-dataset shows positive density in each cavity of the two molecules. (Figure 16) The positive density blobs are located in the same region as the discovered positive density in the SeMet-DtHNL structure but the density is not as distinctive. After the attempt to fill the density with water molecules,

atoms were still missing. This indicates that a larger ligand could be present in the cavity. The Phenix software suite tool Ligand identification was used to identify the ligand but no significant results were achieved. The ligand could be a remnant from the expression or purification steps and is therefore difficult to identify.
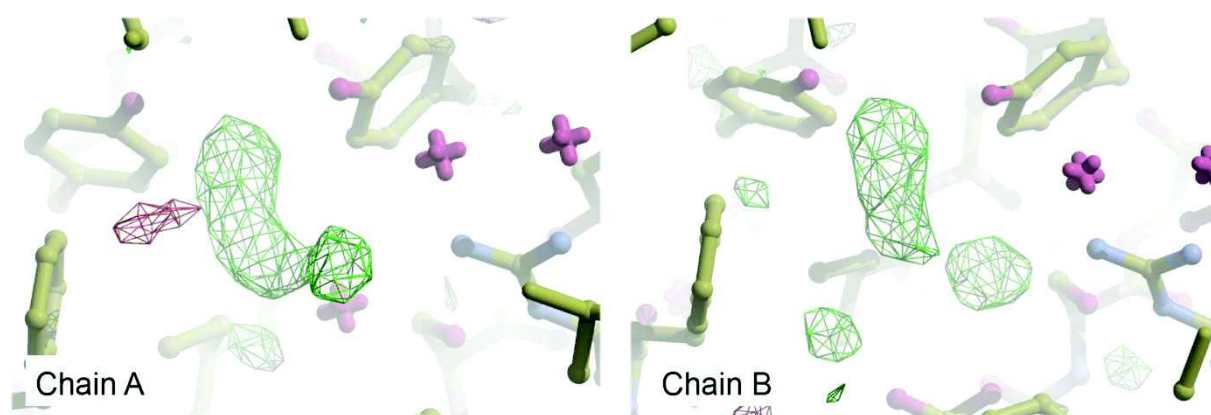
Figure 16. $F_o - F_c$ densities in the active site of the native DtHNL structure. (left) Chain A; (right) Chain B

3.3 DtHNL complex with (R)-mandelonitrile

The first refined structure of the DtHNL complex with soaked (R)-mandelonitrile (DtHNL-MXN I) consists of 2934 non H protein atoms containing 350 of 368 residues in both molecules in the asymmetric unit. The residues 1 to 9 are missing in both chains (A & B). The N-terminal ends of the monomers did not provide enough electron density to place the missing residues. 134 water molecules were placed. Alternative conformations for the residues M100A, E61B, M100B and S122B were made to gain better electron density fitting. Successively performed refinement cycles converged to R factors of R_{work} = 18.05%, R_{free} = 22.68% for 25637 reflections in the resolution range 46.01-2.05Å. In the Ramachandran plot are 97%of the residues in favoured regions and 3% are located in allowed regions (Table 1). The SeMet-DtHNL structure was used as model for molecular replacement.

The second refined structure of DtHNL soaked with R-mandelonitrile (DtHNL-MXN II) is constructed of 3029 non H atoms containing 349 of 368 residues in both chains and 349 water molecules in the asymmetric unit. In both chains (A & B) the residues 1 to 9 are missing. Alternative conformations for the residues S54A, M100A, E61B, T65B, S81B, M100B, L165 and for the ligand in chain B were made to provide better density fitting. The r factors for the final model are Rwork = 17.89% and Rfree = 20.43% for all 49862 reflections in the given resolution range 36.72-1.63Å. Validation of the

Ramachandran plot shows that 97% of the residues are in favoured regions and 3% are in allowed regions. (Table 1) The SeMet-DtHNL structure was used as model for molecular replacement.

Both MXN soaked DtHNL structures consists like the native and the selenomethionine derivate of an 8-stranded antiparallel β-sheet which bends around a long C-terminal α-helix and two consecutive shorter α-helices n-terminal. A cavity is located between the β-sheet and the 3 α-helices that extend through the structure with two openings on the protein surface.

In the first (R)-mandelonitrile soaked structure (F(o) - F(c)) difference density appeared at one site in each chain. These positive density blobs are located in the hydrophobic centre of the cavities. Strict inspections of the densities indicated that it come from the soaked substrate (R)-mandelonitrile. The locations of the two sites are shown in Figure 17.
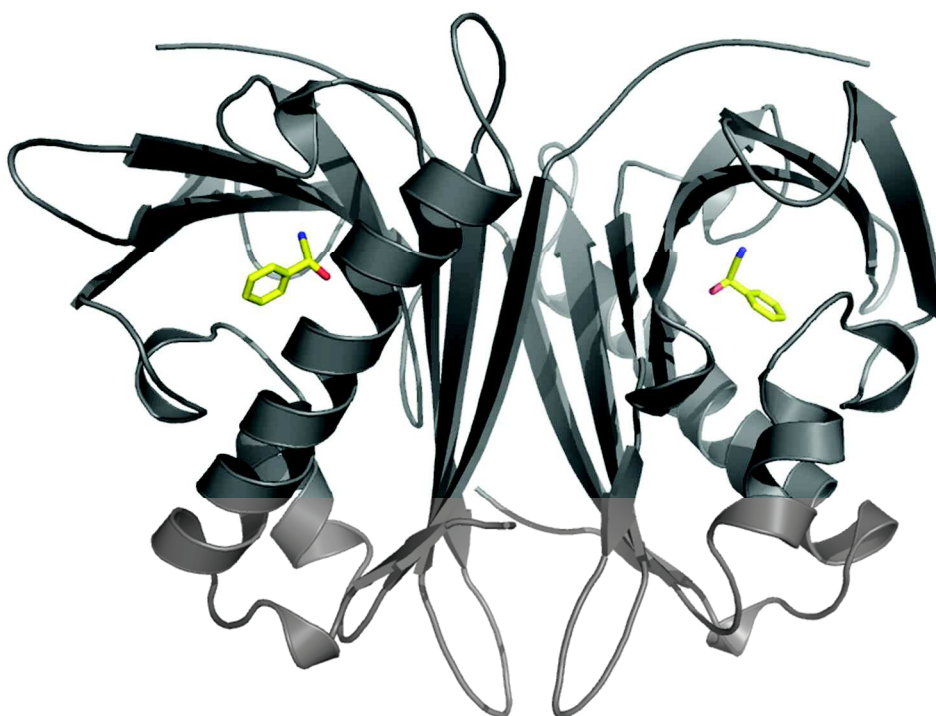
Figure 17. Overall structure of the (R)-mandelonitrile soaked DtHNL I*.* The enzyme is illustrated as cartoon in grey and the soaked (R)-mandelonitrile, located in the active site, is coloured in yellow.

In the assumed active site, a network of hydrogen bonding partners is present, hydrogen bonds exist between the formyl group of the ligand MXN and the hydroxyl group of Y117 as well the hydroxyl group of Y101. Y101 is also in hydrogen bonding interaction with a water molecule, which sits at the
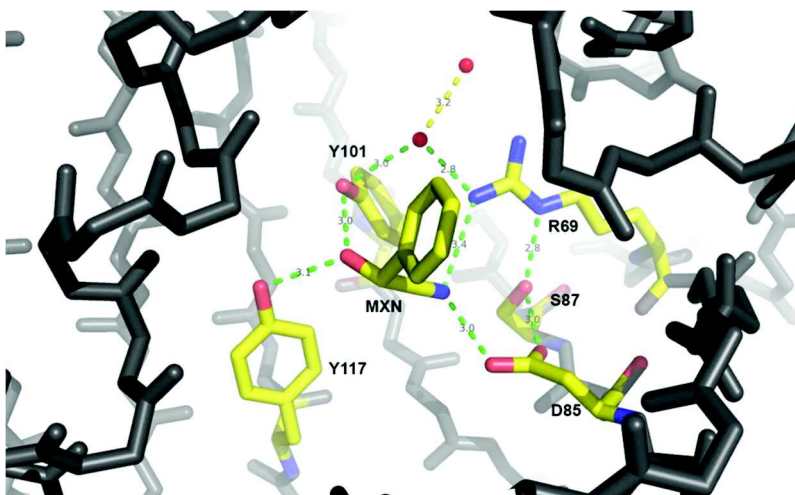
Figure 18. Stick presentation of several amino acids, water molecules and the soaked (R)-mandelonitrile in the active of native DtHNL. The distances between the members of the hydrogen network are measured and illustrated.

end of a water channel with connection to the surface. That water is in turn interacting with the next water molecule in the channel as well with the guanidinium group of R69. R69 is also interacting with the nitrile group of MXN and S87 which is in turn interacting with the carboxyl group of D85. D85 builds also a hydrogen bond to the nitrile group of MXN. (Figure 18)

The aromatic ring of (R)-mandelonitrile is embedded in a hydrophobic pocket consisting of the residues V44, V48, V51, F71, C73, I108, W138 and L160.

In the DtHNL-MXN II structure the previously described (F(o) - F(c)) difference density also appeared at one site in each chain. Detailed examinations of the densities in both chains indicated that the electron density in chain B derive from the product of the enzymatic reaction (benzaldehyde) and not from the substrate (R-mandelonitrile). Nevertheless, the presence of uncleaved substrate at low occupancy cannot be excluded. It is not unusual to observe a product molecule in the active site, since native enzyme was used for crystallisation and the soaking experiments. In chain B a water molecule took over the position of the nitrogen atom in the cyano group of the substrate mandelonitrile. An alternate conformation for the benzaldehyde ligand was set to achieve better density fitting. The density in chain A indicates that it derives from the soaked (R)-mandelonitrile. The obtained structure with both ligands is shown in Figure 19.

Figure 19. Overall structure of the (R)-mandelonitrile soaked dimeric HNL from *Davallia tyermanii.* The enzyme is illustrated as cartoon in grey and the substrate (R)-mandelonitrile, located in the active site of chain A, and the product benzaldehyde, located in the active site of chain B, are coloured in yellow.

The hydrogen network in chain A is analogue to the DtHNL-MXN I (3.3 DtHNL complex with (R)-mandelonitrile) described network. The active sites of chain A and chain B with its hydrogen bonds is shown in Figure 20.

In chain B hydrogen bonds exist between benzaldehyde (conformation A) and the hydroxyl groups of Y101 and Y117 while benzaldehyde (conformation B) only builds a hydrogen bond to the nitrogen compensating water molecule which also builds interactions with R69 and D85. In both conformations Y101 builds a hydrogen bond to the water molecule, located in the previously described water channel. This water molecule interacts with the next water in the channel and also with the guanidinium group of R69. R69 interacts also with S87 which is in interaction with D85.

Comparison with the native structure revealed in both structures no rearrangements between the orientations of the active site residues.
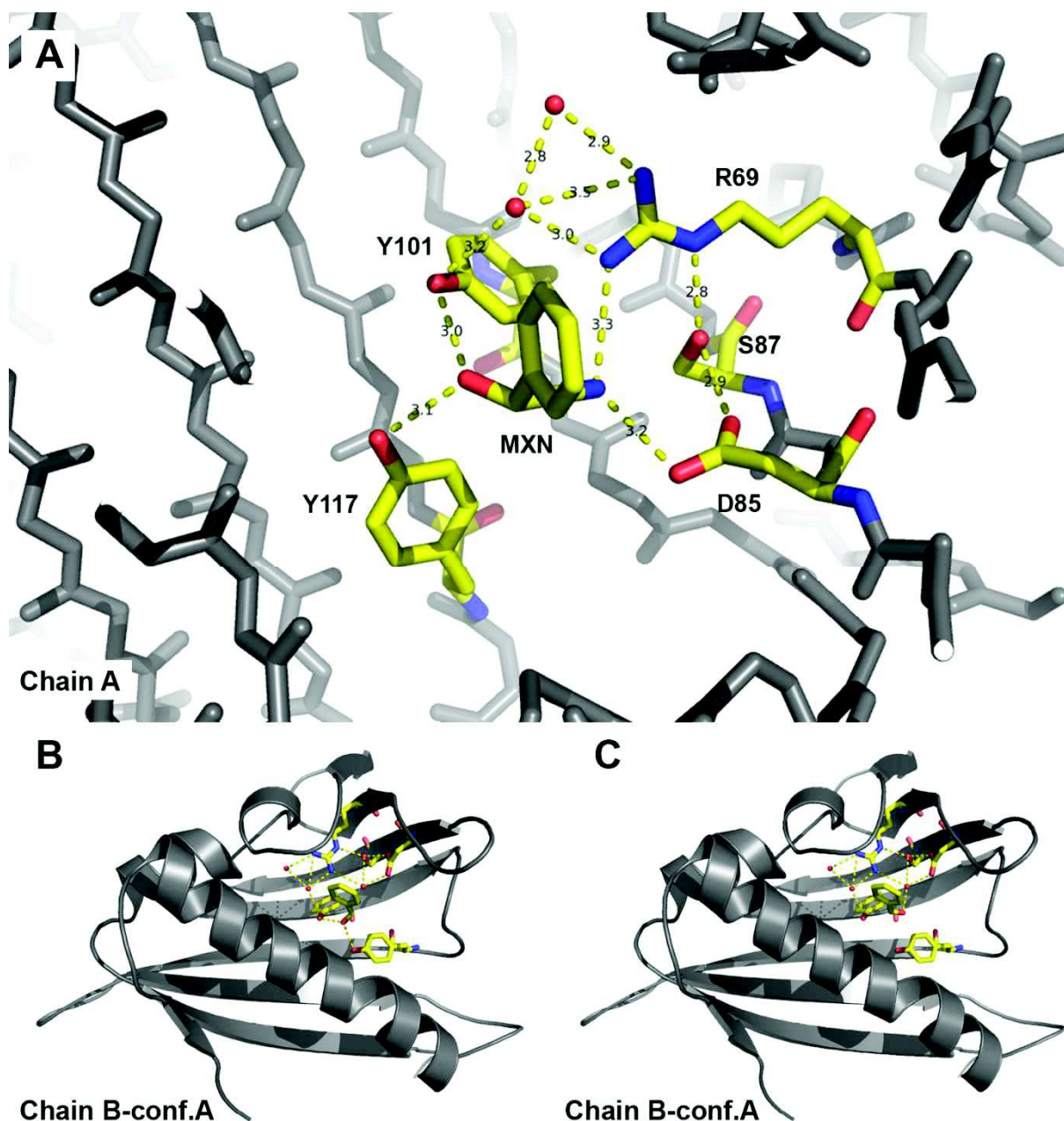
Figure 20. Stick presentation of several amino acids, water molecules and the ligands in the active of native DtHNL. The distances between the members of the hydrogen network are illustrated.
(A) Chain A active site with the substrate (R)-mandelonitrile. (B) Chain B active site with the product benzaldehyde in configuration A. (C) Chain B active site with the product benzaldehyde in configuration B.

3.4 DtHNL complex with benzoic acid

The refined structure of the DtHNL complex with benzoic acid (DtHNL-BEZ) consists of 3076 non H protein atoms containing 349 of 368 residues in both molecules in the asymmetric unit. The first 9 residues in both chains and residue 184 in chain A are not present. The N-terminal ends of the monomers did not provide enough electron density

to place the missing residues. 274 water molecules were placed. Alternative conformations for the residues S54A, M100A, L165A, E61B, T65B, S81B and M100B and the ligand benzoic acid in chain B were made to gain better electron density fitting of the structure. Successively performed refinement cycles converged to R factors of R_{work} = 18.36%, R_{free} = 22.7% for 40351 reflections in the resolution range 46.01-2.05Å. In the Ramachandran plot are 97%of the residues in favoured regions and 3% are located in allowed regions (Table 1). The SeMet-DtHNL structure was used as model for molecular replacement.

The BEZ soaked DtHNL structure consists like the native and the selenomethionine derivate of an 8-stranded antiparallel β-sheet which bends around a long C-terminal α-helix and two consecutive shorter α-helices n-terminal. A cavity is located between the β-sheet and the 3 α-helices that extend through the structure with two openings on the protein surface.

After benzoic acid soaking positive densities in the (F(o) - F(c)) density map appeared in the two cavities of the dimer. The examination of the densities with the benzoic acid soaking background indicated that it come from the soaked inhibitor molecule benzoic acid. An alternative conformation for the chain B benzoic acid was set to achieve better density fitting. (Figure 21) shows the location of the two ligands in the overall structure.
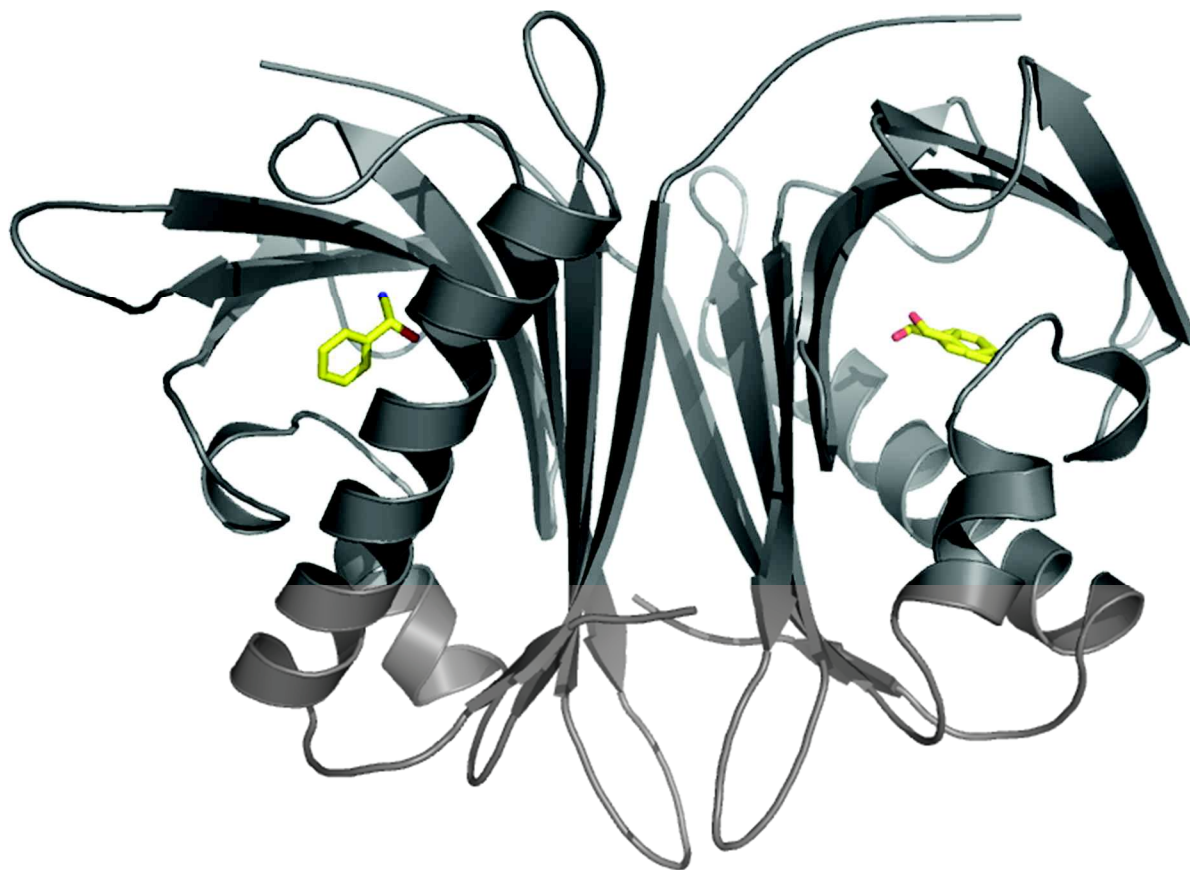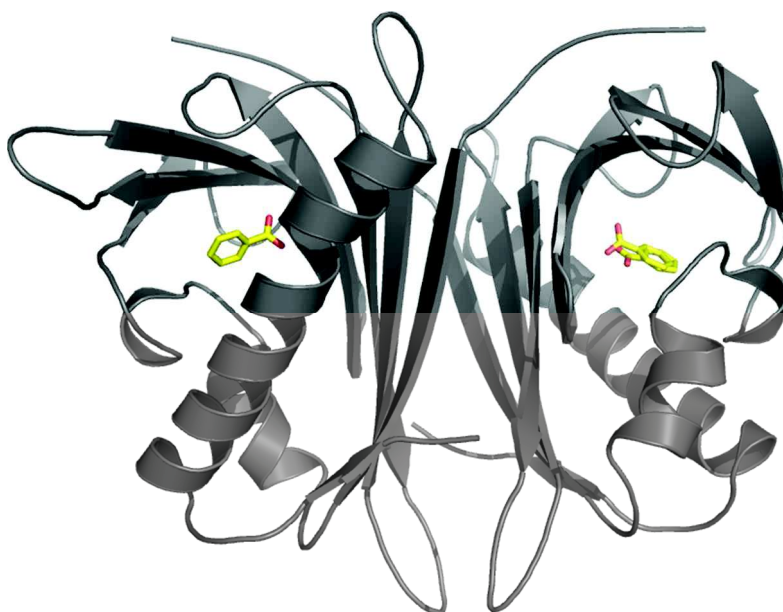

Figure 21. Overall structure of the benzoic acid soaked dimeric HNL from *Davallia tyermanii*. The enzyme is illustrated as cartoon in grey and the ligand benzoic acid, located in the active sites of chain A and B.

In the ligand binding site of chain A hydrogen bonds from the benzoic acid to the hydroxyl group of Y117 as well the hydroxyl group of Y101 are present. Y101 also interact with a water molecule, which sits at the end of a water channel with connection to the surface. That water is also interacting with the next water molecule in the channel

as well with the guanidinium group of R69. R89 also interacts with S87 which builds a hydrogen bond with D85. A water molecule took over the MXN-nitrile group equivalent position in the MXN soaked structure. This water molecule interacts with the hydroxyl group of BEZ, with D85 and with the guanidinium group of R69. (Figure 22A)

The hydrogen bond network in chain B shows slightly different distances. The first conformation (conf A) is comparable to chain A. The nitrogen substituting water molecule is in this chain and conformation 0.3Å closer to the carboxyl group of the BEZ ligand. (Figure 22B)

In conformation B the carboxyl group of the BEZ molecule shows a twist directed toward the residue Y101 which results in alternative bond lengths. (Figure 22C)

Figure 22. Stick presentation of several amino acids, water molecules and the ligand benzoic acid in the active of native DtHNL. The distances between the members of the hydrogen network are illustrated. (A) Chain A active site. (B) Chain B active site with the ligand benzoic acid in configuration A. (C) Chain B active site with the ligand benzoic acid in configuration B.

The given electron density does not support one conformation alone. Therefore, the occupancy of alternative conformation A in chain B was set to 55% and conformation B was set to 45%.

3.5 Implications for the mechanism

On basis of the obtained crystal structures it was possible to formulate three possible mechanisms for the degradation direction under the determination that after the cyanohydrin cleavage only one configuration is consistent with the structural data. (Figure 23)

After the substrate (R)-mandelonitrile is positioned in the active site, the cyano group is facing the amino acids R69, D85 and S87 while the hydroxyl group is hydrogen bonded and stabilized through Y101 and Y161. The negative charge emerging at the cyano group is very likely stabilized by the positive charge of R69 and the hydrogen bond from the protonated D 85.

Y101 takes the proton from the hydroxyl group of the (R)-mandelonitrile, this leads to a new electron distribution and the result is that the cyano group is splited and CN (cyanide) and benzaldehyde is formed.

The protonation of cyanide to form HCN could either involve D85 or possibly a proton relay via R69 and a bridging water molecule between the OH-group of Y101 and the guanidinium group of R69.

These mechanisms were proposed on the basis of the obtained structure and the possible proton distributions in this region of the enzyme. Actually these mechanisms or configurations only differ with respect to the proton distribution in the triad consisting of Y101, R69 and a bridging water molecule.

The possible proton distribution leads to the following three configurations of the triad:

I.   Y101 O$^-$ (deprotonated hydroxyl group) / H2O / R69$^+$
II. Y101 OH (protonated hydroxyl group) / OH$^-$ (hydroxide ion) / R69$^+$
III. Y101 OH (protonated hydroxyl group) / H2O / R69 neutral

After the cyanohydrin cleavage only one configuration is consistent with the structural data:

Y101 OH (protonated hydroxyl group) / H2O / R69$^+$

The developed negative charge at the cyano group is most likely stabilized by the positively charged R69 and a hydrogen bond with the OH group of D85.

To verify these implications, mutations studies were made. The amino acids Y101, Y117 and Y161 were exchanged for a phenylalanine. Y101F led to a total loss of function, while Y117F and Y161F resulted in an activity loss above 90%.

The amino acids D85 and S87 were exchanged for an alanine. The activity loss was also above 90%.

A

B

C
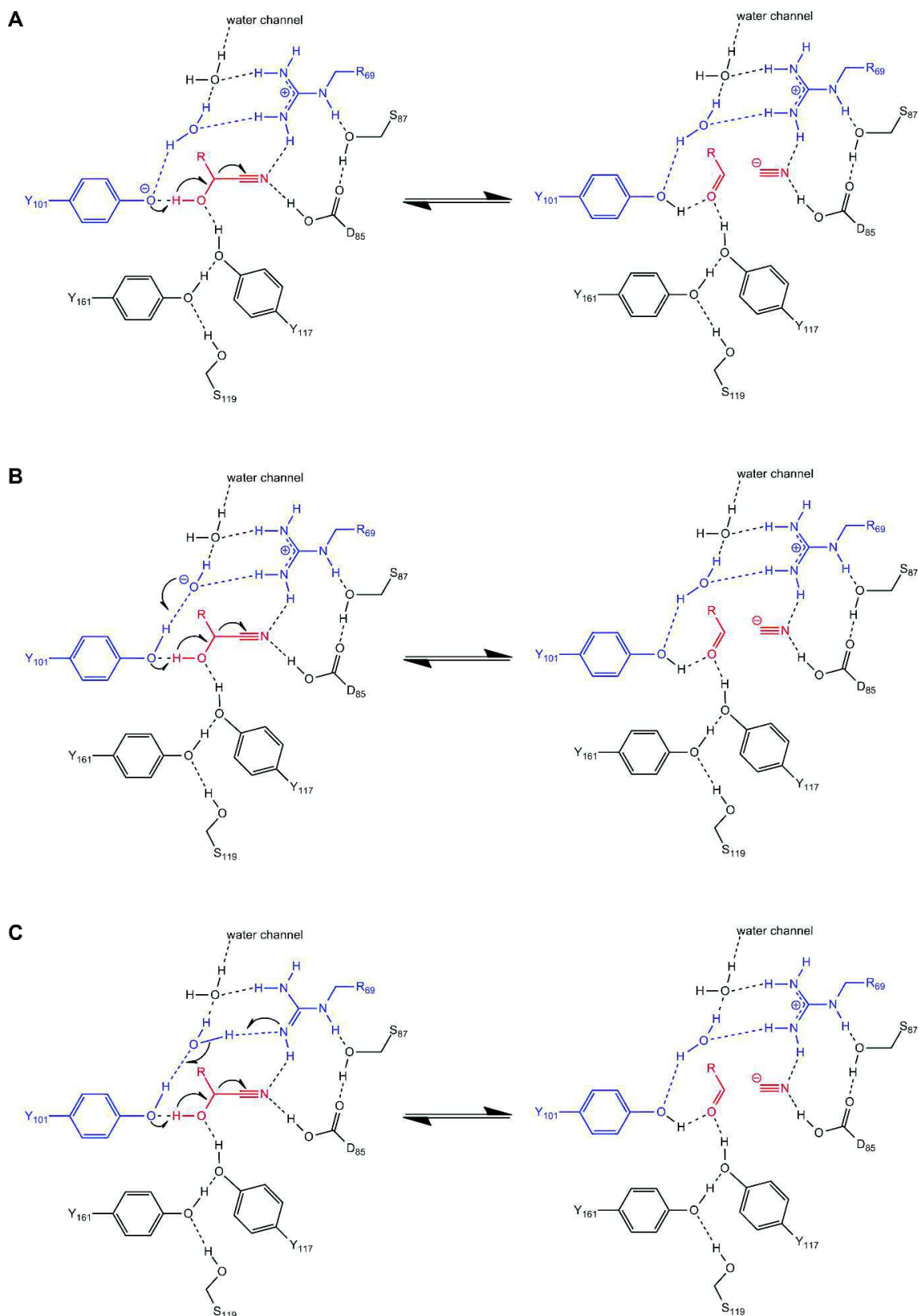
Figure 23. Proposed mechanisms based on the obtained structures and the possible proton distributions in this part of the enzyme.
(A) Configuration I: Y101 O⁻ (deprotonated hydroxyl group) / H2O / R69⁺
(B) Configuration II: Y101 OH (protonated hydroxyl group) / OH⁻ (hydroxide ion) / R69⁺
(C) Configuration III: Y101 OH (protonated hydroxyl group) / H2O / R69 neutral
After the cyanohydrin cleavage only the shown configuration is consistent with the structural data: Y101 OH (protonated hydroxyl group) / H2O / R69⁺

4 Disscussion

In this structural investigation, it was possible to resolve the structure of hydroxylnitrile lyase from *Davallia tyermanii*, which is the first structure of this HNL. The structural models determined by single anomalous diffraction and molecular replacement revealed that DtHNL consists of an 8-standed antiparallel β-sheet (β1-8) and α-helices (α1-3). The β-sheet bends around the long C-terminal α-helix (α3) while the two shorter helices, α1 and α2; separate the C-terminal part of α3 and the β-sheet from each other. This enables the formation of a cavity which stretches through each monomer. CaSox accessibility analysis revealed that this cavity possesses two exits, the smaller one located between α1 and α2 and the larger one between α3 and β3-5. In the crystallographic structure the smaller tunnel is filled with water molecules which are under suspicion to fulfil a role in the catalytic mechanism. This tunnel could also serve as HCN enter or exit tunnel. This would be unique among the known hydroxynitrile lyase structures but still has to be confirmed

The structure of DtHNL revealed that the protein is structurally homologous to the Bet V1 superfamily of proteins. Members of this family share a similar fold, with a 7-standed β-sheet and three α-helices. [66] However structural differences exist, as the DtHNL has an eighth β strand to compose the β sheet.

The inhibitor (DtHNL-BEZ) and substrate soaked structures (DtHNL-MXN I & II) confirmed the suspicion that the active site of DtHNL is located in the centre of the previously described cavity. The inhibitor and product shows, despite alternative conformations and smaller size, compared to the substrate (R)-mandelonitrile similar arrangements. The hydroxyl group of MXN and BEZ or the carbonyl oxygen in HBX is stabilized through the tyrosine pair 101 and 117. In MXN the cyanide nitrogen is also hydrogen bonded with R69 and D85. The soakings with cyanide free ligands and the MXN soaking with a product molecule in the active site revealed that the position of the cyanide nitrogen is occupied with a water molecule.

Interestingly the hydrogen network of the interactive residues in the active site depends on a bridging water molecule located in the smaller surface tunnel. There are reports about water molecules in the active site of HNLs which are involved in the catalytic mechanism [41] but this kind of hydrogen network maintaining water molecules are not described in HNL at the moment. Tunnel blocking mutations could be made to verify the importance of these water molecules.

As previously mentioned the electron density of the ligands in the benzoic acid and (R)-mandelonitrile soaked crystal structures shows possible alternative conformations (Figure 20 and Figure 22). The obtained electron density maps do not support one conformation alone. Due to the missing nitrile group the size of the product and the inhibitor is not as cavity filling as the substrate. The resulting space allows the smaller ligands to obtain alternative conformations within the possible chemical angles in the given active site cavity.

Unexpectedly the crystal structures of SeMet-DtHNL and native DtHNL provided a positive Fo-Fc density in each monomer cavity. The density in the SeMet-DtHNL electron map is very distinctive while the density in the native DtHNL electron map is comparatively not so explicit (Figure 12 and Figure 16). This inhomogeneity could be the result of the quality variation of the two datasets and resulting crystal structures (see Table 1). Anyway it has been reported that proteins from the Bet V1 fold family bind a wide range of molecules [66]. It is possible that a ligand entered the cavity during the protein expression or cell lysis. However, the appearances of unidentified ligands in the active site are not inconsistent with the results from the soaking experiments. It was tried to place the soaked molecules in the positive Fo-Fc densities from the SeMet-DtHNL and native DtHNL, but the refinements did not provide admitting results. The unidentified ligands could be some substrate or product related molecules. Unfortunately, the densities could not be clearly assigned to one certain ligand.

It was possible to draft three possible mechanisms for the degradation of the (R)-mandelonitrile under the determination that after the cyanohydrin cleavage only one configuration is consistent with the structural data. These mechanisms were proposed on basis of the possible proton distributions in the active site region considered from the obtained structures of the enzyme.

The cleavage of the cyanhydrin is triggered by the deprotonation of its hydroxyl group by a triad consisting of Y101, R69 and a bridging water molecule. The given data allow different possible proton distributions of this triad but the amino acid Y101 is in all cases the receiver of the hydroxyl group proton. The deprotonation of the hydroxyl group is enabled through the secondary hydrogen bond with Y181 and leads to a new proton distribution in the ligand. The cyano group is ultimately splited and cyanide and benzaldehyde is formed.

In order to form hydrogen cyanide the CN has to be protonated. The cyanide is facing the amino acids R69, D85 and S87 and the protonation could either involve D85 or a proton relay via R69.

The mutation study supports the assumption that the amino acids R69, D85; S87, Y101, Y117 and Y161 participate in the substrate binding and the enzymatic activity and the results are granting interesting first insights in a potentially new hydroxynitrile lyase catalysed cyanohydrin cleavage.

5   Conclusions

In summary, we have determined structures of DtHNL in the ligand-free form and in complex with MXN, BEZ and HBX. The structures revealed that the DtHNL dimer is a structural homolog of the birch pollen allergen Bet V1. On the basis of the soaking experiments the active site was identified and first implications on the catalytic mechanism were made.

These results significantly contribute to our understanding of the diverse hydroxynitrile lyase family and provide vuluable information's for possible enzyme modifications of other Bet V1 protein members in order to generate HNL activity. The obtained structure allows considerations about DtHNL active site cavity modifications in favour of the accessibility of bigger substrates in industrial applications.

6 References

[1] F. Wöhler and J. Liebig, "Ueber die bildung des bittermandeloels," *Annalen der Physik,* no. 22, pp. 1-24, 1837.

[2] A. Hickel, M. Hasslacher and H. Griengl, "Hydroxynitrile lyases: Functions and properties," *Physiol. Plantarum,* no. 98, p. 891–898, 1996.

[3] L. Rosenthaler, "Enzyme effected asymmetrical synthesis," *Biochem. Zeitschrift,* no. 14, pp. 238-253, 1908.

[4] J. Poulton, "Cyanogenesis in plants," *Plant Physiol.,* no. 94, pp. 401-405, 1990.

[5] M. Zagrobelny, S. Bak and B. Moeller, "Cyanogenesis in plants and arthropods," *Phytochemistry,* no. 69, pp. 1457-1468, 2008.

[6] J. Andexer, J. von Langermann, A. Mell, M. Bocola, U. Kragl, T. Eggert and M. Pohl, "An R-selective hydroxynitrile lyase from Arabidopsis thaliana with an alpha/beta-hydrolase fold," *Angew Chem Int Ed,* no. 46, pp. 8679-8681, 2007.

[7] M. Winkler, A. Glieder and K. Steiner, " C–X Bond Formation: Hydroxynitrile Lyases: From Nature to Application," in *Carreira E.M. and Yamamoto H. (eds.) Comprehensive Chirality*, vol. 7, Amsterdam, Elsevier, 2012, pp. 350-371.

[8] F. Effenberger, "Synthesis and reactions of optically active cyanohydrins," *Angew Chem Int Ed Engl,* no. 33, pp. 1555-1564, 1994.

[9] R. Gregory, "Cyanohydrins in nature and the laboratory, biol-ogy preparations and synthetic applications," *Chemical Reviews,* no. 99, pp. 3649-3682, 1999.

[10] K. CG., "Chiral cyanohydrins—their manufacture & utility as chiral building blocks," in *Collins AN, Shildrake GN, Crosby J, editors. Chirality in Industry.*, New York, John Wiley & Sons, 1992, pp. 279-299.

[11] E. E. Conn, "In The Biochemistry of Plants: A Comprehensive Treatise. (Stumpf, P.K. & Conn, E.E., eds)," *Secondary plant products,* vol. 7, pp. 479-500, 1981.

[12] H. Wajant and F. Effenberger, "Hydroxynitrile lyases of higher plants," *Biological Chemistry,* no. 377, pp. 611-617, 1996.

[13] R. Lieberei, D. Selmar and B. Biehl, "Metabolization of cyanogenic glucosides in Hevea brasiliensis," *Plant Syst Evol,* no. 150, pp. 49-63, 1985.

[14] D. Selmar, "Apoplastic occurrence of cyanogenic β-glucosidases and consequences for the metabolism of cyanogenic glucosides," *ACS symposium series,* no. 533, pp. 191-204, 1993.

[15] H. Wajant and K. Pfizenmaier, "Identification of potential active-site residues in the hydroxynitrile lyase from Manihot esculenta by site-directed mutagenesis," *Journal of Biological Chemistry,* no. 271, p. 25830–25834, 1996.

[16] M. Hasslacher, M. Schall, M. Hayn, H. Griengli, . S. D. Kohlwein and H. Schwab, "Molecular Cloning of the Full Length cDNA of (S)-Hydroxynitrile Lyase from Hevea brasiliensis, Functional Expression in E.coli and Identification of an Active Site Residue," *Journal of Biological Chemistry,* no. 271, pp. 5884-5891, 1996.

[17] I. P. Cheng and J. Poulton , "Cloning of cDNA of Prunus serotina (R)-(+)-mandelonitrile lyase and identification of a putative FAD-binding site," *Plant Cell Physiol.,* no. 34, p. 1139–1143, 1993.

[18] Z. Hu and J. E. Poulton, "equencing, genomic organization, and preliminary promoter analysis of a black cherry (R)-(+)-mandelonitrile lyase gene," *Plant Physiol.,* no. 115, p. 1359–1369, 1997.

[19] H. Wajant, K. W. Mundry and K. Pfizenmaier, "Molecular cloning of hydroxynitrile lyase from Sorghum bicolor (L.). Homologies to serine carboxypeptidases," *Plant Mol Biol.,* no. 26, p. 735–746, 1994.

[20] K. Trummler and H. Wajant, "Molecular cloning of acetone cyanohydrin lyase from flax (Linum usitatissimum). Definition of a novel class of hydroxynitrile lyases.," *Journal of Biological Chemistry,* no. 272, pp. 4770-4774, 1997.

[21] Z. Hussain, R. Wiedner, K. Steiner, T. Hajek, M. Avi, B. Hecher, A. Sessitsch and H. Schwab, "Characterization of Two Bacterial Hydroxynitrile Lyases with High Similarity to Cupin Superfamily Proteins," *Applied and Environmental Microbiology,* no. 78, p. 2053–2055, 2012.

[22] K. Gruber and C. Kratky, "Biopolymers for Biocatalysis: Structure and Catalytic Mechanism of Hydroxynitrile Lyases," *Journal of Polymer Science Part A: Polymer Chemistry,* no. 42, p. 479–486, 2004.

[23] M. Schall, "Isolation and Characterization of a (S)-Hydroxynitrile Lyase from Hevea brasiliensis," *Ph.D. Thesis University of Graz,* 1996.

[24] U. Hanefeld, G. Stranzl, A. J. J. Straathoff, J. J. Heijnen, A. Bergmann, R. Mittelbach, O. Glattler and C. Kratky, "Electrospray ionization mass spectrometry, circular dichroism and SAXS studies of the (S)-hydroxynitrile lyase from Hevea brasiliensis," *Biochimica et biophysica acta. Protein structure and molecular enzymology,* no. 1544, pp. 133-142, 2001.

[25] H. Hasslacher, M. Schall, M. Hayn, H. Griengl, H. Schwab and S. D. Kohlwein, "(S)-Hydroxynitrile lyase from Hevea brasiliensis," *Ann. NY Acad. Sci.,* no. 799, pp. 702-712, 1996.

[26] N. Klempier and H. Griengl, "Aliphatic (S)-cyanohydrins by enzyme catalyzed synthesis," *Tetrahedron Lett.,* no. 34, pp. 4769-4772, 1993.

[27] N. Klempier, U. Pichler and H. Griengl, "Synthesis of a β-unsaturated (S)-cyanohydrins using the oxynitrilase from Hevea brasiliensis.," *Tetrahedron: asymmetry,* no. 6, p. 845–848, 1995.

[28] D. Johnson and H. Griengl, "The Chemoenzymatic Synthesis of (S)-13-Hydroxyoctadecadienoic Acid using the hydroxynitrile lyase from Hevea brasiliensis.," *Tetrahedron 53,* no. 53, pp. 617-624, 1997.

[29] U. Wagner, M. Hasslacher, H. Griengl, H. Schwab and C. Kratky, "Mechanism of cyanogenesis: the crystal structure of hydroxynitrile lyase from Hevea brasiliensis.," *Structure,* no. 4, pp. 811-822, 1996.

[30] J. Zuegg, K. Gruber, M. Gugganig, U. Wagner and C. Kratky, "Three-dimensional structures of enzyme-substrate complexes of the hydroxynitrile lyase from Hevea brasiliensis," *Protein science,* no. 8, pp. 1990-2000, 1999.

[31] K. Gruber, "Elucidation of the mode of substrate binding to hydroxynitrile lyase from Hevea brasiliensis.," *Poteins Struct. Funct. Genet.,* no. 44, pp. 26-31, 2001.

[32] S. Bornemann, "Flavoenzymes that catalyse reactions with no net redox change," *Natural product reports,* no. 19, pp. 761-772, 2002.

[33] M. S. Jorns, "Mechanism of catalysis by the flavoenzyme oxynitrilase," *Journal of Biological Chemistry,* no. 254, pp. 12145-12152, 1979.

[34] I. Dreveny, K. Gruber, A. Glieder, A. Thompson and C. Kratky, "The hydroxynitrile lyase from almond: A lyase that looks like an oxidoreductase," *Structure,* no. 9, p. 803–815, 2001.

[35] M. Kiess, H. J. Hecht and H. M. Kalisz, "Glucose oxidase from Penicillium amagasakiense. Primary structure and comparison with other glucose-methanol-choline (GMC) oxidoreductases.," *European Journal of Biochemistry,* no. 252, pp. 90-99, 1998.

[36] R. K. Wierenga, J. Drenth and G. E. Schulz, "Comparison of the three-dimensional protein and nucleotide structure of the FAD-binding domain of p-hydroxybenzoate hydroxylase with the FAD-as well as NADPH-binding domains of glutathione reductase," *Journal of Molecular Biology,* no. 167, pp. 725-739, 1983.

[37] I. Dreveny, C. Kratky and K. Gruber, "The active site of hydroxynitrile lyase from Prunus amygdalus: modeling studies provide new insights into the mechanism of cyanogenesis.," *Protein science,* no. 11, pp. 292-300, 2002.

[38] H. Wajant and K. W. Mundry, "Hydroxynitrile lyase from Sorghum bicolor - a glycoprotein heterotetramer," *Plant science,* no. 89, pp. 127-133, 1993.

[39] C. Bove and E. E. Conn, "Metabolism of aromatic compounds in higher plants. II. Purification and properties of the oxynitrilase of Sorghum vulgare," *Journal of Biological CHemistry,* no. 236, pp. 207-210, 1961.

[40] D. I. Liao, K. Breddam, R. M. Sweet, T. Bullock and S. T. Remington, "Refined atomic model of wheat serine carboxypeptidase II at 2.2-Å resolution," *Biochemistry,* no. 31 , pp. 9796-9812, 1992.

[41] H. Lauble, B. Miehlich, S. Foerster, H. Wajant and F. Effenberger, "Crystal structure of hydroxynitrile lyase from Sorghum bicolor in complex with the inhibitor benzoic acid: a novel cyanogenic enzyme," *Biochemistry,* no. 41, pp. 12043-12050, 2002.

[42] L. L. Xu, B. K. Singh and E. E. Conn, "Purification and characterization of acetone cyanohydrin lyase from Linum usitatissimum," *Archives of biochemistry and biophysics,* no. 263, pp. 256-263, 1988.

[43] J. Albrecht, I. Jansen and M. R. Kula, "Improved purification of an (R)-oxynitrilase from Linum usitatissimum (flax) and investigation of the substrate range," *Biotechnology and applied biochemistry,* no. 17, pp. 191-203, 1993.

[44] S. F. Altschul, W. Gish, W. Miller, E. W. Myers and D. J. Lipman, "Basic local alignment search tool," *Journal of Molecular Biology,* no. 215, p. 403–410, 1990.

[45] T. Luque, S. Atrian, O. Danielsson, H. Jörnvall and R. Gonzàlez-Duarte, "Structure of the Drosophila melanogaster Glutathione-Dependent Formaldehyde Dehydrogenase/Octanol Dehydrogenase Gene (Class III Alcohol Dehydrogenase)," *European Journal of Biochemistry,* no. 225, p. 985–993, 1994.

[46] H. W. Sun and B. V. Plapp, "Progressive sequence alignment and molecular evolution of the Zn-containing alcohol dehydrogenase family," *Journal of Molecular Evolution,* no. 35, pp. 522-535, 1992.

[47] H. Breithaupt, M. Pohl, W. Boenigk, P. Heim, K. L. Schimz and M. R. Kula, "Cloning and expression of (R)-hydroxynitrile lyase from Linum usitatissimum (flax)," *Journal of Molecular Catalysis B: Enzymatic,* no. 6, p. 315–332, 1999.

[48] G. Agarwal, M. Rajavel, B. Gopal and N. Srinivasan, "Structure-Based Phylogeny as a Diagnostic for Functional Characterization of Proteins with a Cupin Fold," *PLoS ONE,* no. 4, p. e5736, 2009.

[49] J. M. Dunwell, A. Culham, C. E. Carter, C. R. Sosa-Aguirre and P. W. Goodenough, "Evolution of functional diversity in the cupin superfamily.," *Trends in biochemical sciences,* no. 26, pp. 740-746, 2001.

[50] J. E. Dunwell, A. Purvis and S. Khuri, "Cupins: the most functionally diverse protein superfamily?," *Phytochemistry,* no. 65, pp. 7-17, 2004.

[51] I. Hajnal, A. Łyskowski, U. Hanefeld, K. Gruber, H. Schwab and K. Steiner, "Biochemical and structural characterization of a novelbacterial manganese-dep endent hydroxynitrile lyase," *FEBS Journal,* no. 280 , p. 5815–5828, 2013.

[52] F. A. Chavez , A. Banerjee and B. Sljivic , "Modeling the metal binding site in cupin proteins," in *On Biomimetics*, Rijeka,Croatia, InTech, 2011, p. 3–28.

[53] E. Kiljunen and L. T. Kanerva, "Approach to (R)- and (S)-ketone cyanohydrins using almond and apple meal as the source of (R)-oxynitrilase.," *Tetrahedron: Asymmetry,* no. 8, p. 1551–1555, 1997.

[54] E. Lanfranchi, K. Steiner, M. Winkler, A. Glieder, T. Pavkov-Keller and M. Diepold, "New hydroxynitrile lyases". Patent 14180247.0 - 1410, 2014.

[55] W. Kabsch, "XDS," *Acta Cryst,* no. D66, pp. 125-132, 2010.

[56] T. Terwilliger, P. Adams, R. Read, A. McCoy, N. Moriarty, R. Grosse-Kunstleve, P. Afonine, P. Zwart and L. Hung, "Decision-making in structure solution using Bayesian estimates of map quality: the PHENIX AutoSol wizard," *Acta Cryst,* no. 65, pp. 582-601 , 2009.

[57] T. Terwilliger, R. Grosse-Kunstleve, P. Afonine, N. Moriarty, P. Zwart, L. Hung, R. Read and P. Adams, "Iterative model building, structure refinement and density modification with the PHENIX AutoBuild wizard," *Acta Cryst,* no. 64, pp. 61-69, 2008.

[58] P. Adams, R. Grosse-Kunstleve, L. Hung, T. Ioerger, A. McCoy, N. Moriarty, R. Read, J. Sacchettini, N. Sauter and T. Terwilliger, "PHENIX: building new software for automated crystallographic structure determination," *Acta Cryst,* no. 58, pp. 1948-1954, 2002.

[59] P. Emsley, B. Lohkamp, W. G. Scott and K. Cowtan, "Features and development of Coot," *Acta Cryst,* no. 66, pp. 486-501, 2010.

[60] T. G. G. Battye, L. Kontogiannis, O. Johnson, H. R. Powell and A. G. W. Leslie, "iMOSFLM: a new graphical interface for diffraction-image processing with MOSFLM," *Acta Cryst.,* no. D67, pp. 271-281, 2011.

[61] P. Evans, "Scaling and assessment of data quality," *Acta Cryst,* no. D62, pp. 72-82, 2006.

[62] A. J. McCoy, R. W. Grosse-Kunstleve, P. D. Adams, M. D. Winn, L. C. Storoni and R. J. Read, "Phaser crystallographic software," *Journal of Applied Crystallography,* no. 40, pp. 658-674, 2007.

[63] B. W. Matthews, "Solvent content of protein crystals," *Journal of Molecular Biology,* no. 33, pp. 491-497, 1968.

[64] G. Kleywegt and A. Brünger, "Checking your imagination: applications of the free R value.," *Structure,* no. 4, pp. 897-904, 1996.

[65] Z. Marković-Housley, M. Degano, D. Lamba, E. von Roepenack-Lahaye, S. Clemens, M. Susani, F. Ferreira, O. Scheiner and H. Breiteneder, "Crystal structure of a hypoallergenic isoform of the major birch pollen allergen Bet v 1 and its likely biological function as a plant steroid carrier," *Journal of molecular biology,* no. 325(1), pp. 123-133, 2003.

[66] C. Radauer, C. Lackner and H. Breiteneder, "The Bet v 1 fold: an ancient, versatile scaffold for binding of large, hydrophobic ligands," *BMC Evolutionary Biology,* no. 8, 286, 2008.

[67] M. Estonius , J. O. Hoeoeg and O. Danielsson, "Residues specific for class III alcohol dehydrogenase. Site-directed mutagenesis of the human enzyme," *Biochemistry,* no. 33 , p. 15080–15085, 1994.