Identification of Potential Active-site Residues in the Hydroxynitrile Lyase from *Manihot esculenta* by Site-directed Mutagenesis*

(Received for publication, March 13, 1996, and in revised form, August 7, 1996)

Harald Wajant†‡ and Klaus Pfizenmaier

From the Institute of Cell Biology and Immunology, University of Stuttgart, Allmandring 31, 70569 Stuttgart, Germany

The hydroxynitrile lyase from cassava (*Manihot esculenta* Crantz) (EC 4.1.2.37) catalyzes the decomposition of the achiral α-hydroxynitrile acetone cyanohydrin into HCN and acetone during cyanogenesis of damaged plants. This enzyme can also be utilized for stereoselective synthesis of a wide array of (S)-cyanohydrins and by addition of HCN to aldehydes or ketones. Optically active cyanohydrins are interesting intermediates for the synthesis of α-hydroxy acids, α-hydroxy ketones, or β-ethanolamines, all of which are important building blocks in organic synthesis. Inhibition of hydroxynitrile lyase from *M. esculenta* (MeHNL) by serine- and histidine-modifying reagents suggests involvement of active site seryl and histidyl residues. Furthermore, serine 80 of MeHNL is part of the active site motif Gly-X-Ser-Gly/Ala, often considered as the hallmark of catalytic triads having independently evolved in four groups of enzymes: the αβ hydrolase fold enzymes, subtilisins, the cysteine proteases, and the eukaryotic serine proteases. By site-directed mutagenesis, three residues critical for enzyme activity have been identified: serine 80, aspartic acid 208, and histidine 236. These residues may be directly involved in MeHNL-catalyzed decomposition of cyanohydrins, providing evidence for a catalytic triad in HNLs, too. The order of the catalytic triad residues in the primary sequence of MeHNL is nucleophile-histidine-acid, suggesting that MeHNL belongs to the αβ hydrolase fold group of enzymes. In contrast to all other enzymes having a catalytic triad, HNLs catalyze no net hydrolytic reactions.

Hydroxynitrile lyases, which catalyze the dissociation of α-hydroxynitriles (cyanohydrins) into carbonyl and HCN (Fig. 1), are described for several plant species of widely differing phylogenetic origin (1). In recent years these enzymes have received increasing attention due to their potential use as catalysts for the stereoselective synthesis of chiral cyanohydrins, providing evidence for a catalytic triad in HNLs, too. To date, HNLs from various dicotyledons (7, 8) as well as from *Sorghum bicolor* (ShHNL) (9, 10), a monocotyledone, and *Phlebodium aureum* (PhaHNL) (11), a fern, have been biochemically characterized. HNLs comprise a heterogenous group of enzymes with homo- and heteromers as the active enzyme, glycoproteins and nonglycoproteins, as well as flavoproteins. The molecular mass of HNL subunits ranges from 20 to 60 kDa (7–11). The heterogeneity of their properties suggests that most of these enzymes have independently evolved (convergent evolution). In contrast, the lack of homologies among the amino acid sequences of the recently cloned HNLs from *Prunus serotina* (PsHNL) (12), *Manihot esculenta* (MeHNL) (13), and *S. bicolor* (14) emphasize this idea. While PsHNL and MeHNL show none or only limited homology to other known proteins (12, 13), ShHNL has astonishingly high sequence homology with serine carboxypeptidases, especially with those from wheat. It appears of particular relevance that ShHNL shares the catalytical triad Ser, Asp, and His with these enzymes (14). The catalytical triad motif evolved in the carboxypeptidase, chymotrypsin, and subtilisin group of proteases by convergent evolution but is also found in a series of other hydrolytic enzymes (15–17). The above-mentioned similarities between ShHNL and serine carboxypeptidases provoked the question of whether other HNLs also use the catalytical triad motif for catalysis. In support of this idea, almost all HNLs are inhibited by serine/cysteine and/or histidine-modifying agents (11, 14, 18).

In the present study, we have used site-directed mutagenesis to demonstrate that serine 80, which is part of the typical serine protease Gly-X-Ser-Gly/Ala consensus motif, is essential for enzyme activity of MeHNL. In addition, site-directed mutagenesis further shows that histidine 236 and aspartic acid 208 are important for enzyme activity. These observations suggest that the catalytical triad motif is also utilized in the active site of MeHNL. This work provides the first fundamental evidence for use of a catalytic triad by a member of the hydroxynitrile group of enzymes.

**EXPERIMENTAL PROCEDURES**

**Expression Cloning of MeHNL**—The MeHNL gene was cloned in the expression vector pQE4 (Quiagen). In brief, total RNA was isolated from fresh leaves of *M. esculenta* using a guanidine chloride method (19). Subsequently, the coding region of the MeHNL gene was amplified from oligo(dT)-primed first strand cDNA using primer with 5′ overhangs encompassing *Bam*HI restriction sites (forward primer, GCA GGC CGG CAT CCC ATT TCC AAA ATG GTA ACT GCA CA; reverse primer, GCA GGC CGG CAT CCC ATT TCC AAA ATG GTA ACT GCA CA; reverse primer, GCA GGC CGG CAT CCC ATT TCC AAA ATG GTA ACT GCA CA; reverse primer, GCA GGC CGG CAT CCC ATT TCC AAA ATG GTA ACT GCA CA). The in frame insertion of the *Bam*HI digested polymerase chain reaction fragment into the *Bam*HI site of pQE4 results in the plasmid pQE4-MeHNLLwt, which was transformed in *Escherichia coli* M15[pREP4] cells (Quiagen) for overexpression of MeHNL.

**Site-directed Mutagenesis**—Mutations were introduced into pQE4-MeHNLLwt using the Chameleon double-stranded, site-directed mutagenesis kit (Stratagene) according to the manufacturer’s recommendations. In brief, the double-stranded target plasmid DNA was heat denatured, and a selection primer, which removes a nonessential

---

*This work was supported by the Bundesministerium für Forschung und Technologie, Germany (Grant A03-U-ZSP Stuttgart) and by Deutsche Forschungsgemeinschaft Grant WA 1025/1–1. The publication of this article was defrayed in part by the payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

† To whom correspondence should be addressed. Tel.: 49-711-685-7446; Fax: 49-711-685-7484.

‡ The abbreviations used are: ShHNL, hydroxynitrile lyase from *S. bicolor* (EC 4.1.2.11); DFP, diisopropyl fluorophosphate; MeHNL, hydroxynitrile lyase from *M. esculenta* (EC 4.1.2.37); PsHNL, hydroxynitrile lyase from *P. serotina* (EC 4.1.2.11).
unique restriction site, as well as a primer defining the mutation of interest (Table were annealed to one strand. These primers were extended around the plasmid using T7 DNA polymerase and T4 DNA Ligase. Subsequently, the plasmid DNA was digested with the restriction enzyme that corresponds to the nonessential restriction site, which is still present in unmutated plasmids. Undigested, mutated plasmid DNA was favored, compared to digested parental plasmid DNA, in a subsequent transformation into a repair-deficient E. coli strain due to the greater transformation efficiency of circular plasmid DNA compared to linear plasmid DNA. The transformed bacteria were grown in a liquid culture overnight. To select further for mutated plasmids, DNA was isolated from the overnight culture, digested with the restriction enzyme corresponding to the nonessential unique restriction site, removed, and transformed into competent XL1-Blue cells. Mutants were identified by introduction of a new restriction site and/or by sequence analysis using the dideoxy chain termination method of Sanger et al. (20). All primers were 5'-phosphorylated before use. The single codon mutated plasmid was listed in Table 1, in which the introduced substitution, its location, and the resultant amino acid are also summarized. A 27mer oligonucleotide having the sequence 5'-CAT CAT TGG AAA ACG CTC TTC GGG GCG-3' from the Amp' gene of pQE4.

Expression and Purification of Wild Type and Mutated MeHNL—Wild type and mutant MeHNL genes were expressed in E. coli strain M15[pREP4]. For induction of recombinant proteins, an overnight culture was performed at 37°C in LB medium supplemented with ampicillin (100 μg/ml) and kanamycin (25 μg/ml), diluted 1:10 with LB medium supplemented only with ampicillin (100 μg/ml), and grown at 30°C. 90 min after dilution isopropyl-β-D-thiogalactopyranoside was added to a final concentration of 1 mM for induction of gene expression. The cells were harvested by centrifugation after 4 h, resuspended in 50 mM sodium acetic acid, pH 5.4, and disintegrated by 20 pulses at 60% with a Sonoulus HD 200 (Bandelin, Berlin, Germany). After centrifugation (16,000 rpm, 20 min, 4°C, Ja 20 rotor), wild type and mutant MeHNLs were purified from the supernatant as described previously for MeHNLs from the natural source (18). Briefly, the supernatant was applied to Q-Sepharose FF at pH 5.4, and the enzyme eluted with NaCl. Final purification of MeHNL was obtained by gel exclusion chromatography on Superdex 200 (Pharmacia) and anion exchange chromatography on Mono-Q (Pharmacia) at pH 7.5.

Enzyme Assays—The assay for MeHNL based on the measurement of the HCN formed during dissociation of acetyl cyanohydrin. The release of HCN was determined as described by Selmar et al. (21). One unit of HNL activity is defined as the amount of enzyme that can catalyze the decomposition of 1 μmol of substrate/min under the conditions described by Selmar et al. (21). The calculation of the specific activity of total protein was determined with the BCA Protein Assay kit according to the manufacturer's recommendations. For inhibition studies with DFP (Sigma), phenylmethylsulfonyl fluoride (Boehringer Mannheim), and 4-(2-aminoethyl)-benzenesulfonyl fluoride hydrochloride (Boehringer Mannheim), 5 μg of protein were incubated in the indicated concentration of the reagent for 15–30 min at room temperature. The remaining activity was then determined as described above.

### Site-directed Mutagenesis Identifies the Catalytical Triad Ser80-Asp208-His236 in MeHNL—The above data obtained with serine-modifying reagents indicated a critical functional role of a serine residue in MeHNL activity. For a molecular identification of catalytically important residues, site-directed mutagenesis was used. The rationale for performing site-directed mutagenesis of serine 80 was based on sequence analysis showing serine 80 as the best candidate for a serine of catalytic importance because it is the only serine of MeHNL within a Gly-X-Ser-X-Gly/Ala consensus motif. This motif defines the serine of the catalytic triad in serine proteases. The putative functional importance of serine 80 is further underscored by the fact that the consensus motif is conserved (maximum: 33.2% identity in a 258-amino acid overlap) between MeHNL and two proteins of unknown function from rice (25) (Fig. 2), whose sequence is deposited in the EMBL data bank (accession numbers Z34270 and Z34271). We selected alanine as a substitute for serine 80. Expression of the S80A mutant enzyme in E. coli revealed no differences to wild type MeHNL according to oligomeric structure, molecular weight, and behavior in the standard purification procedure (Table II, Fig. 3). However, the
Potential Active-site Residues in the Hydroxynitrile Lyase

TABLE I
Mutations generated by site-directed mutagenesis

| Sequence of mutagenic oligonucleotide 5′–3′ | Location of oligonucleotides on MeHNL cDNA | Amino acid substitution |
|--------------------------------------------|--------------------------------------------|------------------------|
| CAT TGT TGG TGA GCC CTG TGC AGG GC         | 250–276                                     | S80A                   |
| GGT AAC TGC AGC TAA AGT TCT G             | 28–55                                       | H5A                    |
| GCA CAT TTT GTT CTG ATC GCG GGT            | 35–72                                       | H10A                   |
| GCA TGG ATT TGG C                         | 347–379                                     | H112A                  |
| GCC AGA CAC CGT CCG GAG CCC ATC           |                                            |                        |
| TTA CAA TG                              |                                            |                        |
| CAG GTC CAA GGT GGA GAC GCG AAG           | 815–858                                     | H236A                  |
| CTC GAG CTT ACA AAA ACT G                 | 283–331                                     | D95A                   |
| GCT ATT GCT GAT AGA TAC TGT               |                                            |                        |
| GCC AAA ATT GCA GCT GTT TTC               | 392–436                                     | D130A                  |
| GAG TCG TTT CCT GAC TGG AGG GCC           |                                            |                        |
| ACA GAG TAT TGG ACC GAT CAG GCC           | 716–762                                     | D208A                  |
| AAA ATA TTA TTA CCA GAC TTT C             |                                            |                        |

* Nucleotide number is based on the published MeHNL sequence (13).

TABLE II
Comparison of wild type and mutated forms of MeHNL

The kinetic parameters $K_m$ and $V_{max}$ were calculated from a Lineweaver-Burk plot using acetone cyanohydrin as substrate. The coefficient of determination value of the linear regression was above 0.95 in all experiments.

| Mutant | Elution volume in Sephadex 200 | Enzyme activity in crude extract | Specific activity of purified enzyme | $K_m$ value$^b$ | $V_{max}$ | Inhibition by 2 mM DFP | NaCl$^d$ |
|--------|--------------------------------|---------------------------------|------------------------------------|-----------------|-----------|-----------------------|---------|
| Wild type | Leaves: 192 ± 0.5 | 0.38 | 92 ± 4.1 | 120 | 125 | >80 | 120 |
|         | Recombinant: 192 ± 0.6 | 18 | 86 ± 5 | 101 | 138 | >80 | 110 |
| S80A | 192.3 ± 0.4 | <0.5 | <1$^d$ | — | — | — | 115 |
| H5A | 192.5 | 20 | 95 | 125 | 140 | >70 | 110 |
| H10A | 192 | 13.4 | 90 | 130 | 140 | >80 | 110 |
| H112A | 193.5 | 18 | 90 | 95 | 130 | >70 | 120 |
| H236A | 192 ± 0.6 | <0.5 | <1$^d$ | — | — | — | 115 |
| D95A | 193 | 18 | 90 | 120 | 130 | >80 | 120 |
| D130A | 191.5 | 14 | 85 | 120 | 120 | >80 | 110 |
| D208A | 192.5 ± 0.8 | 4.2 | 18.3 ± 1.9 | — | — | >70 | 115 |

* Elution volumes were determined spectrophotometrically for purified enzymes and by enzyme assay of 1-ml fractions for crude extracts.

$^b$ Kinetic studies were performed with purified enzymes.

$^d$ [NaCl] by which MeHNL was eluted during anion exchange chromatography on Mono Q at pH 5.4.

$^e$ No detectable enzyme activity.

$^f$ Activity to a low to adequately calculate $K_m$ and $V_{max}$ values.

FIG. 2. Alignment of amino acid sequences of MeHNL and two proteins of unknown function, which can be induced by Pseudomonas syringae in Oryza sativa. Amino acid residues are numbered on the right. Conserved residues are indicated by asterisks, and conservative substitutions are indicated by points. Residues of the putative catalytic triad of MeHNL are indicated by black arrows.

...mutated enzyme was completely inactive in the standard acetone cyanohydrin cleaving assay (Table II).

Serine 80 is part of the consensus motif Gly-X-nucleophile-X-Gly-Ala, which is considered as the hallmark of proteins having the catalytic triad nucleophile–histidine–acid. To identify the histidine residue involved in a putative catalytic triad of MeHNL, we have systematically replaced histidine residues by alanine. We have restricted this alanine scanning to histidines conserved between MeHNL and the above-mentioned rice proteins (Asp95, H5A, H10A). Comparison of wild type and mutated forms of MeHNL with respect to the before-mentioned rice proteins (Asp95, H5A, H10A) gives compelling evidence for involvement of a catalytic triad in this enzyme. All known non-proteolytic enzymes having a catalytic triad belong to the α/β hydrolase fold group of enzymes. In all these enzymes, the linear order of the catalytic residues in the primary sequence is nucleophile–histidine–acid. To identify the catalytic triad acid of MeHNL, we have constructed mutants with alanine substituting for aspartic acid residues located between serine 80 and histidine 236 in the primary sequence and which are conserved with respect to the before-mentioned rice proteins (Asp95, Asp100, and Asp208). A greater 80% inhibition of enzyme activity was found when aspartic acid 208 was replaced by alanine...
me NaCl (Table II). Furthermore, the $K_m$ values of all mutants not affected in the putative catalytic triad residues are comparable to the $K_m$ value of the wild type protein (Table II). Therefore, we assume that the introduced substitutions do not change general protein integrity. Therefore, as mutations S80A, D208A, and H236A all dramatically reduce enzyme activity, serine 80, aspartic acid, and histidine 236 are likely to be directly involved in catalysis or represent important sites of substrate recognition.

**Proposed Reaction Mechanism of MeHNL—Chymotrypsin is a well studied example of an enzyme using a catalytic triad for catalysis. The enzymatic activity of chymotrypsin critically depends on histidine 57 and serine 195. Both are located near each other in the active site. Analogous to the reaction mechanism proposed for the hydrolysis of peptide bonds by chymotrypsin, one can formulate a possible general base catalyzed mechanism for cyanohydrin cleavage as follows: In free MeHNL, the hydroxyl group of serine 80 is hydrogen-bonded to the imidazole nitrogen of histidine 236, which in turn is stabilized by a hydrogen bond with aspartic acid 208. If the enzyme encounters a substrate molecule, the proton of serine 80 is rapidly transferred to an imidazole nitrogen of histidine 236. The resulting oxanion of serine 80 functions as a strong base in a nucleophilic attack of the substrate hydroxyl proton, leading to a negatively charged oxanion on the substrate. This oxanion could be stabilized by an oxygen hole formed by amide nitrogen of the backbone of serine 80 and glycine 78 (Fig. 4). Both residues belong to the consensus motif, which is typical for nucleophils located in a catalytic triad. The stabilized oxanion could further increase the local positive charge on the $\alpha$-C atom of the cyanohydrin, allowing the cyanide group to leave the molecule.

**Cyanohydrin Cleavage by MeHNL Defines a Novel Type of Catalytic Triad-catalyzed Reaction—**To date, four groups of enzymes have been found to utilize the catalytic triads as active site principle: subtilisins, the cysteine proteases, the eukaryotic serine proteases, and the $\alpha/\beta$ hydrolase fold enzymes (17). It is commonly accepted that this stable active site principle was evolved by convergent evolution (17). All members of the first three groups are proteases, whereas the recently defined group of $\alpha/\beta$ hydrolase fold enzymes also contains lipases, dehalogenases, and esterases (17). The members of this group of enzymes are of widely differing phylogenetic origin but share the structural core motif of eight $\beta$ sheets connected by $\alpha$ helices. Ollis et al. (17) suggested that the $\alpha/\beta$ hydrolase fold group of enzymes represents molecules, where the catalytic subsite framework was conserved during evolution. These enzymes have therefore the same linear order of active site residues in their primary sequences.
We have identified the components of a putative catalytic triad in MeHNL using site-directed mutagenesis. The order of these residues is identical to those in enzymes of the α/β hydrolase fold group, suggesting that MeHNL also belongs to this group of molecules. Cleavage of cyanohydrins into HCN and carboxylate compounds by MeHNL defines a new reaction executed by α/β hydrolase fold enzymes. Interestingly, this reaction is not a net hydrolytic reaction, like all other reactions catalyzed by enzymes having a catalytic triad. The recently cloned HNL from S. bicolor (SbHNL) shows extended sequence homologies to serine carboxypeptidase from wheat, which is also a member of the α/β hydrolase fold enzymes (14, 17). Thus, SbHNL is likely to be a member of this enzyme group. MeHNL and SbHNL share no significant sequence homologies. Moreover, the high sequence homologies between SbHNL and serine carboxypeptidase from wheat indicate that SbHNL evolved only recently. In fact, the SbHNL type of HNLs is restricted to the sorghum group of gramineae.2 In contrast, the putative common ancestor of all α/β hydrolase fold enzymes should be placed early in molecular evolution due to lack of sequence homologies among these enzymes. Therefore, MeHNL and SbHNL would have to be independently evolved from two different members of the α/β hydrolase fold group. In other words, SbHNL and MeHNL appear to be the result of convergent evolution on the background of molecules, which were in turn evolved by divergent evolution.

The well known HNL from P. serotina, which is unusual in containing a FAD without catalyzing a net oxido/reduction reaction, was also recently cloned (12). And in fact, this enzyme also possesses exactly one nucleophilic residue, which is part of the before-mentioned consensus motif, defining the nucleophile in a catalytic triad.

REFERENCES
1. Seigler, D. S. (1992) in Herbivores, Their Interactions with Secondary Plant Metabolites (Rosenthal, G. A., and Berenbaum, M. R., eds) Vol. I, pp. 35–77, Academic Press, Inc., San Diego
2. Effenberger, F., Ziegler, T., and Förster, S. (1987) Angew. Chem. Int. Ed. Engl. 26, 458–460
3. Effenberger, F., Hirsch, B., Förster, S., and Ziegler, T. (1990) Tetrahedron Lett. 31, 1249–1250
4. Smitskamp-Wilms, E., Brussese, J., van der Gen, A., van Scharrenburg, G. J. M., and Sloothaak, J. B. (1991) Recl. Trav. Chim. Phys.-Bio 110, 209–215
5. Albrecht, J., Jansen, I., and Kula, M. R. (1993) Biotechnol. Appl. Biochem. 17, 191–203
6. Klempier, N., Griengl, H., and Hayn, M. (1993) Tetrahedron Lett. 34, 4769–4772
7. Poulton, J. E. (1988) in Cyanide Compounds in Biology (Evered, D., and Harrett, S., eds), Ciba Foundation Symposium No. 140, pp. 67–91, John Wiley and Sons, Chichester, UK
8. Kurek, G. W., and Conn, E. E. (1989) Proc. Natl. Acad. Sci. U.S.A. 86, 6978–6981
9. Boye, C., and Conn, E. E. (1961) J. Biol. Chem. 236, 207–210
10. Wajant, H., and Mundry, K.-W. (1993) Plant Sci. 99, 127–133
11. Wajant, H., Förster, S., Selmar, D., Effenberger, F., and Pfizenmaier, K. (1995) Plant Physiol. 100, 1211–1218
12. Cheng, I.-P., and Poulton, J. E. (1993) Plant Cell Physiol. 34, 1139–1143
13. Hughes, J., Carvalho, J. P. de C., and Hughes, M. A. (1994) Arch. Biochem. Biophys. 311, 496–502
14. Wajant, H., Mundry, K.-W., and Pfizenmaier, K. (1994) Plant Mol. Biol. 26, 735–746
15. Liao, D. I., and Remington, S. J. (1996) J. Biol. Chem. 271, 6528–6531
16. Liao, D. I., Breddam, K., Sweet, R. M., Bollock, T., and Remington, S. J. (1992) Proc. Natl. Acad. Sci. U.S.A. 89, 6978–6981
17. Ollis, D. L., Cheah, E., Cyliger, M., Dijkstra, B., Frolow, F., Franken, S. M., Harel, M., Remington, S. J., Silman, I., Schrag, J., Sussman, J. L., Verschooren, K. H. G., and Goldman, A. (1992) Protein Eng. 5, 197–211
18. Wajant, H., Förster, S., Bottinger, H., Effenberger, F., and Pfizenmaier, K. (1995) Plant Sci. 108, 1–11
19. Logemann, J., Schell, J., and Willmitzer, L. (1987) Anal. Biochem. 163, 16–20
20. Sanger, F., Nicklen, S., and Coulson, A. R. (1977) Proc. Natl. Acad. Sci. U.S.A. 74, 5463–5467
21. Selmar, D., Carvalho, J. P. de C., and Conn, E. E. (1987) Anal. Biochem. 166, 208–211
22. Laemmli, U. K. (1970) Nature 227, 680–685
23. Blum, H., Beier, H., and Gross, H. J. (1987) Electrophoresis 8, 93–99
24. Towbin, H., Staehelin, T., and Gordon, J. (1979) Proc. Natl. Acad. Sci. U.S.A. 76, 4350–4354
25. Reimmann, C., Hofmann, C., and Dudler, R. (1995) Physiol. Mol. Plant Pathol. 46, 71–81
