Published in final edited form as:

ACS Catal. 2017 June 2; 7(6): 4221–4229. doi:10.1021/acscatal.7b01108.

Identical active sites in hydroxynitrile lyases show opposite enantioselectivity and reveal possible ancestral mechanism

Bryan J. Jones^a, Zsófia Bata^{a,b}, and Romas J. Kazlauskas^{*,a}

^aUniversity of Minnesota, Department of Biochemistry, Molecular Biology & Biophysics, 1479 Gortner Avenue, Saint Paul, MN, 55108 USA

^bBudapest University of Technology and Economics, Department of Organic Chemistry and Technology, H-1111 Budapest, 3 M egyetem rkp., Hungary

Abstract

Evolutionarily related hydroxynitrile lyases from rubber tree (*HbHNL*) and from *Arabidopsis thaliana* (*AtHNL*) follow different catalytic mechanisms with opposite enantioselectivity toward mandelonitrile. We hypothesized that the *HbHNL*-like mechanism evolved from an enzyme with an *AtHNL*-like mechanism. We created ancestor-like composite active-sites in each scaffold to elucidate how this transition may have occurred. Surprisingly, a composite active site in *HbHNL* maintained (*S*)-selectivity, while the identical set of active site residues in *AtHNL* maintained (*R*)-selectivity. Composite active-site mutants that are (*S*)-selective without the Lys236 and Thr11 that are required for the classical (*S*)-HNL mechanism suggests a new mechanism. Modeling suggested a possibility for this new mechanism that does not exist in modern enzymes. Thus, the last common ancestor of *HbHNL* and *AtHNL* may have used an extinct mechanism, not the *AtHNL*-like mechanism. Multiple mechanisms are possible with the same catalytic residues and residues outside the active site strongly influence mechanism and enantioselectivity.

Graphical abstract

*Corresponding Author: rjk@umn.edu.

Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acscatal.0000000.

Additional experimental and modeling details

Eight Tables of data

Three Figures

Notes

The authors declare no competing financial interest.

Keywords

esterase; hydroxynitrile lyase; α/β -hydrolase fold; ancestral enzyme; enantioselectivity; molecular dynamics

INTRODUCTION

Hydroxynitrile lyases (HNL's) are defense enzymes that catalyze the elimination of hydrogen cyanide from cyanohydrins. The released hydrogen cyanide kills or deters the predators. Typical cyanohydrin substrates are acetone cyanohydrin or mandelonitrile derived from valine or phenylalanine, respectively. For synthetic applications, chemists carry out the reverse reaction, an enantioselective addition of hydrogen cyanide to carbonyl compounds, Equation 1. HNL's occur mainly in plants, but bacteria and arthropods also contain HNL's^{1,2}. HNLs have independently evolved in five different protein folds³⁻⁷, and even multiple times within the α/β -hydrolase fold, suggesting that catalyzing cyanohydrin cleavage is relatively easy.

(1)

HNL's in the α/β -hydrolase fold superfamily evolved at least twice and use at least three different mechanisms to catalyze hydroxynitrile lyase cleavage. These elimination mechanisms differ in the identity and location of the cyanide-stabilizing residue. First, and most firmly established, is the mechanism for HNL from rubber tree (*Hevea brasiliensis*), *HbHNL*, which evolved from esterases in this family⁸. A combination of x-ray structure analysis, mutagenesis and kinetics^{9,10} revealed that, although the active site contains an

esterase-like catalytic triad of Ser-His-Asp, the role of the serine differs. In esterases, the catalytic serine is a nucleophile, but in *HbHNL*, it transfers protons between the substrate and catalytic histidine, Figure 1a. The positive charge of Ne of Lys 236 stabilizes the negative charge on the leaving cyanide. We will refer to this *HbHNL* mechanism as the Lys mechanism. *HbHNL* catalyzes efficient cleavage of both its natural substrate acetone cyanohydrin and of aromatic cyanohydrins like mandelonitrile, where it favors the (*S*)-enantiomer.

HNL from *Arabidopsis thaliana* (*AHNL*) also evolved from esterases, but its mechanism differs from the Lys mechanism. Instead of Thr11 and Lys236, which are essential for the Lys mechanism, *AHNL* contains Asn11 and Met236. Met236 also occurs in esterases, but Asn11 does not. Docking¹¹ and QM/MM¹² studies suggest the catalytic histidine serves as the base as in *HbHNL*, but the oxyanion hole is the cyanide stabilizing group, Figure 1b. Two main chain N–H groups donate hydrogen bonds to the cyanide. One consequence of this different mechanism is *AHNL*'s unusual catalytic promiscuity, weakly catalyzing ester hydrolysis in addition to efficient hydroxynitrile lyase cleavage. Another consequence is *AHNL*'s opposite enantioselectivity. *HbHNL* and other enzymes following the Lys mechanism favor the (*S*)-enantiomer of mandelonitrile, while *AHNL* favors the (*R*)-enantiomer.

The *AHNL* gene is transcribed upon leaf damage or senescence¹³, so the likely role of *AHNL* is plant defense, but its exact role is unknown. Since its amino acid sequence resembles esterases, it was initially annotated as a methyl esterase EST5 or MES5. Subsequent experiments detected no hydrolysis of plant signaling esters such as methyl indole-3-acetate and methyl jasmonate^{14–16}, but low esterase activity ($<2 \text{ min}^{-1}$) with unnatural substrates like naphthyl acetate and methyl pentanoate^{8,17}. *AHNL* is a good hydroxynitrile lyase with aromatic substrates like mandelonitrile^{8,17–20}, but does not react with small cyanohydrins. *Arabidopsis thaliana* does not produce mandelonitrile or acetone cyanohydrin²¹, but does produce indole cyanohydrin²². *AHNL*'s role may be to release hydrogen cyanide by lysis of indole cyanohydrin or by hydrolysis of subsequent metabolites like carbonyl nitriles. *AHNL* might also play a role in the glucosinolate-myrosinase system, a homologous plant defense pathway^{23–25}.

The third mechanism in the α/β -hydrolase-fold superfamily evolved from carboxypeptidases and occurs in HNL from *Sorghum bicolor*, *SbHNL* (mechanism not shown on Figure 1)⁵. The amino acid sequence of this HNL is most closely related to serine carboxypeptidases in the α/β -hydrolase-fold superfamily, so this HNL likely evolved independently of *HbHNL* and *AHNL*. In *SbHNL* the catalytic triad serine is too far from the substrate to serve as a base, so the proposed base is the carboxylate of the C-terminal tryptophan residue. The functional groups that stabilize the cyanide ion have not been identified. This *SbHNL* mechanism will not be further considered in this paper.

Divergent evolution of new enzymes involves catalytically promiscuous intermediates that catalyze both the old and new catalytic activities²⁶. Subsequent optimization of separate genes encoding these intermediates yields specialist enzymes for each catalytic activity. The puzzle regarding the evolution of HNLs from esterases in that the esterase and modern

lysine catalyzed HNL mechanisms are incompatible. Modern esterases do not show promiscuous lyase activity and modern hydroxynitrile lyases do not show promiscuous esterase activity. Site-directed mutagenesis to exchange key active site residues yielded intermediates with low catalytic activity^{27,28}. Two molecular reasons for the incompatibility are that esterases require access to the oxyanion hole and a uncharged alcohol binding site, while HNLs block the oxyanion hole and a key Lys converts the uncharged alcohol binding site into a charged cyanide binding site. While both enzymes contain a Ser-His-Asp catalytic triad, these HNLs also contain the key residues Thr11 and Lys236.

Ancestral enzyme reconstruction identified last common ancestors of lysine and oxyanion hole HNLs and closely related plant esterases. Several of these ancestors were catalytically promiscuous^{8,17}. HNL1 catalyzed mainly hydroxynitrile lyase cleavage, but also showed low remaining esterase activity. The active site of HNL1 is similar to modern HNL's with Thr11 Lys236 key residues. More interesting, the older ancestor EST3ml showed both esterase and HNL activity while replacing Thr11 and Lys236 by Asn11 and Met236 suggesting a different mechanism for HNL cleavage in this ancestral enzyme. EST3ml was 58% identical *HbHNL* and 78% identical to *AhHNL*, including active site residues like Met236 and Asn11. Like *AhHNL*, this ancestor was also (*R*) selective, but only weakly (*E* = 6.7). The goal of this paper is identify the EST3ml mechanism for HNL cleavage.

We hypothesized that the ancestral enzyme EST3ml followed an oxyanion mechanism like *AhHNL* and subsequent optimization switched the mechanism to the Lys mechanism. To understand the differences between the two types of HNL mechanisms in *HbHNL* and *AhHNL*, we attempted to exchange their catalytic abilities by exchanging active site residues. The choice of residues to exchange was guided by similarities to ancestral enzyme EST3ml. In this paper, we describe exchanges between active site residues of *AhHNL* and *HbHNL*. Surprisingly, we discover a possible fourth mechanism for HNL activity in the α/β -hydrolase-fold superfamily, which we call the Asn mechanism, Figure 1C.

RESULTS

Twenty-two amino acid residues lie within 6 Å of the substrate mandelonitrile in *HbHNL* (PDB:1YB6²⁹) or in the superimposed structure of *AhHNL* (PDB:3DQZ, chain A¹¹). Six of these residues are the same in both proteins (leucine 157, histidine 14, leucine 146, and tyrosine 158, as well as the residues serine 80 and histidine 235 from the catalytic triad, *HbHNL* numbering), while sixteen residues differ between the two enzymes, Figure 2 & Table S1. We hypothesized that exchanging active site residues between *AhHNL* and *HbHNL* would exchange their reaction mechanisms.

Mutating the two catalytically essential residues in *HbHNL* to the corresponding ones in *AhHNL* yielded catalytically inactive enzyme. Previously, researchers replaced Thr11 and Lys239 with catalytically inert amino acids (glycine, alanine, leucine, methionine), which eliminated 98–100% of the activity^{9,27}. Here we replaced either or both residues with corresponding asparagine or the previously tested methionine from *AhHNL*. The resulting variants lost greater than 99% of their HNL activity with mandelonitrile (<0.2/sec), indicating that simple substitution does not switch the mechanism (Table S4).

Likewise, replacing the catalytic residues Asn11, Phe80, and/or Met239 in *A*HNL with the corresponding residues in *Hb*HNL (Thr, Glu, and Lys, respectively) mostly resulted in insoluble protein. The one exception was *A*HNL with the single mutation N11T, which expressed as soluble protein, but was inactive. However, many non-catalytic single *A*HNL mutants tested were soluble and retained partial or full activity.

The other extreme was to replace sixteen residues in *Hb*HNL to make all twenty-two amino acid residues in the substrate binding region identical to those in *A*HNL. Unfortunately, neither this exchange nor the reverse – exchange of all sixteen residues in *A*HNL with the corresponding residues in *Hb*HNL – yielded soluble protein. We hypothesize that some substitutions created destabilizing interactions with the second shell of amino acid residues surrounding the active site. To create soluble variants, instead of switching all sixteen residues, we created composite identical active sites using a mix of residues from *A*HNL and *Hb*HNL to resemble the last common ancestor of these two proteins. These sites lacked the essential Thr and Lys from *Hb*HNL and therefore were expected to follow an *A*HNL-like mechanism. In this paper, we describe the construction and characterization of these variants along with modeling to explain their unusual behavior.

Various active site exchanges created thirty-six protein variants, Table S4, of which approximately half were soluble and catalytically active. To improve the soluble expression of some *Hb*HNL mutants, we added the stabilizing H103V substitution. This residue lies in a buried hydrophobic pocket outside of the active site (>8 Å from the substrate). Replacement of the charged histidine with a hydrophobic valine minimally affected catalytic activity or enantioselectivity³⁰, but increased the stability and protein expression of *Hb*HNL³¹ and the homologous HNL from *Manihot esculenta*^{31,32}.

Two variant pairs were particularly interesting because the equivalent active site in either *Hb*HNL or *A*HNL produced soluble, active protein. Starting with *Hb*HNL, substitution of either nine or ten positions (11, 12, 79, 81, 106, 148, 152, 210, and 236, with or without 178) with the corresponding residues in *A*HNL yielded *Hb*-A9-H7 and *Hb*-A10-H6 respectively. Similarly, starting with *A*HNL, substitution of the remaining six or seven amino acid residues (13, 121, 125, 128, 131, and 209, with or without 178) with the corresponding residues in *Hb*HNL yielded *At*-A9-H7 and *At*-A10-H6. These four proteins create two pairs with identical amino acid residues within the active site, but differing residues outside the active site. In these composite sites, all residues in the catalytic center (non-hydrogen atom closer than 5.4 Å to either the substrate oxygen, chiral C2 carbon, or cyanide, PDBs: 1YB6 & 3DQZ) are *A*HNL-like, including the catalytically important Asn11 and Met236. The nonpolar substrate binding pocket contains predominantly *Hb*HNL-like residues, but some *A*HNL-like residues are also present. The *Hb*HNL-like residues are closer to the phenyl carbons of the substrate, except Cys13, which is closer to the chiral C2 carbon, but is still far at 6.2 Å between heavy atoms. These four proteins had similar catalytic activity toward mandelonitrile (0.4–1.9 sec⁻¹), which is significantly lower than the catalytic activity of *Hb*HNL (25.6 sec⁻¹) or *A*HNL (27.0 sec⁻¹), Table S4. The natural substrate for *Hb*HNL is acetone cyanohydrin (126/sec), but neither *A*HNL or *Hb*-A9-H7 cleave this substrate (<0.1/sec; in 5 mM acetone cyanohydrin).

Surprisingly, the two proteins with “A10-H6” composite active sites still showed opposite enantioselectivity corresponding to the favored enantiomer of the starting protein, Figure 2 Figure S3, and Table S5. Wild type *HbHNL* favors (*S*)-mandelonitrile ($E > 39$) and is inhibited by (*R*)-mandelonitrile (Figure S2). *Hb*-A10-H6 retains this (*S*)-preference, but also catalyzed cleavage of (*R*)-mandelonitrile, lowering the enantioselectivity ($E = 3.3$). Wild type *AtHNL* favors (*R*)-mandelonitrile ($E > 39$) and *At*-A10-H6 enzyme retained the (*R*)-preference, also with lower enantioselectivity ($E = 5.2$). Thus, the same A10-H6 active site shows opposite enantioselectivity in the different scaffolds demonstrating that residues outside the active site affect enantioselectivity strongly enough to reverse the favored enantiomer. In terms of activation energy differences, this difference in enantioselectivity ($E = 3.3$ favoring *S* vs. 5.2 favoring *R*) corresponds to a $G^\ddagger$ of 1.7 kcal/mol³³.

Changing phenylalanine 178 to leucine converted the “A10-H6” site to the other composite active site, “A9-H7”, which was (*S*)-selective in both enzyme scaffolds, Figure 2, Figure S3 and Table S5. The enantioselectivity in the *HbHNL* scaffold increased from 3.3 (*S*) in *Hb*-A10-H6 to >39 (*S*) in *Hb*-A9-H7, which corresponds to a $G^\ddagger$ of at least 1.5 kcal/mol toward the (*S*) enantiomer. The enantioselectivity in the *AtHNL* scaffold also shifted toward the (*S*)-enantiomer from 5.3 (*R*) in *At*-A10-H6 to 1.6 (*S*) in *At*-A9-H7. This reversal corresponds a $G^\ddagger$ of 1.3 kcal/mol toward the (*S*)-enantiomer. Thus, the F178L substitution shifts enantioselectivity toward (*S*)-mandelonitrile by 1.3 – 1.5 kcal/mol in both scaffolds. The remaining difference in enantioselectivity (>39 (*S*) vs. 1.6 (*S*)) corresponds to a contribution of 1.9 kcal/mol from the residues outside the active site. This value for the contribution of residues outside the active site is similar to effect of residues outside the active site in the A9-H7 pair of enzymes above.

Modeling

Molecular dynamics simulations predicted a new mechanism for the composite active sites with the *HbHNL* scaffold. In the new mechanism, the cyano group of mandelonitrile accepts a hydrogen bond from H₈₂₁ of Asn11, Table 1, Figure 1c. Different binding from *AtHNL* is unexpected, as all residues in the catalytic center (as defined above) of the composite active sites are *AtHNL*-like. Catalytically active conformations (CAC) are present in 81 and 93% of the simulation time for *Hb*-A10-H6 and *Hb*-A9-H7. The cyano group hydrogen bonded almost exclusively to Asn11 in the CAC. The partially positively charged H₈₂₁ Asn11 ($q = +0.38e$, OPLS_2005 force field) interacts with the partially negatively charged nitrogen of the cyano group, fulfilling a similar role to Lys236 and the two main chain amino groups in the other two mechanisms. The stabilizing effect on the forming cyanide leaving group should be proportional the total positive charge of the protein partner. Lys236 is positively charged, and the oxyanion hole donates two hydrogen bonds, but Asn is not charged and donates only one hydrogen bond. This smaller positive charge may explain the decreased reaction rates for the composite active sites with the *HbHNL* scaffold, Table S4. The hydroxyl group of mandelonitrile, as in the other two mechanisms, donates a hydrogen bond to either Ser80 or His235.

Molecular dynamics simulations predicted an oxyanion mechanism for composite active sites with the *AtHNL* scaffold, Table 1. Simulations show the *AtHNL* variants bound the

cyano group in the oxyanion hole in all of the CAC. This similarity to wt-*A*HNL is expected, since all residues in the catalytic center of the composite active sites are *A*HNL-like, including the catalytically important Asn11 and Met236. Mandelonitrile hydrogen bonded to either Ser80 and His235 in the catalytically active frames, suggesting that either may be the proton accepting base. The supporting information includes detailed analysis of mandelonitrile binding in the wild type enzymes used as control calculations.

Modeling of composite active sites with *A*HNL scaffold explains the shift toward (*S*) enantioselectivity from wt-*A*HNL to *At*-A10-H6, but not the further shift toward (*S*) in *At*-A9-H7. In *At*-A10-H6, both enantiomers bound in CAC (*S*- 33%, *R*- 35%) in agreement with the low enantioselectivity. Altered binding of the phenyl group of mandelonitrile accounts for the increase in CAC for the slower (*S*)-enantiomer as compared to wt-*A*HNL. Adding the L178F substitution to make *A*HNL-A9-H7 reversed the enantioselectivity to favor (*S*)-MNN, but the modeling showed similar binding to wild type *A*HNL and found the slow (*R*) enantiomer more frequently in CAC. These results suggest that both (*R*) and (*S*)-mandelonitrile may be cleaved by the oxyanion hole mechanism.

Modeling of composite active sites with *Hb*HNL scaffold explains the high enantioselectivity in *Hb*-A9-H7, but misses the shift toward (*R*) in *Hb*-A10-H6 due to F178L. The favored, (*S*)-mandelonitrile, binds in CAC in 93% and 81% of the simulation time in *Hb*-A9-H7, *Hb*-A10-H6 active sites, compared to only 13% for (*R*)-mandelonitrile. The effect of F178L on the enantioselectivity is missed by the modeling, as substrate binding and the CAC ratios are similar in *Hb*-A9-H7 and *Hb*-A10-H6 active sites.

The different reactivity and enantioselectivity of composite active sites, despite identical amino acid residues, means that the structures of the active sites differ. The closest 44 residues to the substrates (Table S7) consist of 33 identical residues (either conserved or mutated to be equivalent), and 11 residues that differ in the two scaffolds. The first shell residues (16 + 5) are all identical, while the second shell contains 12 identical and 11 differing residues. To identify how the differing residues create different conformations in the identical residues, we compared the positions of these 44 residues during the MD simulations. The median distance between center of mass of each residue and the center of mass of Ser80 at each frame of the simulation indicated the typical position of the residue. The median absolute deviation of this distance measured the movement during the simulation. Seventeen of the 44 residues' positions varied more between proteins with the same composite active site than the residues moved within each protein, Figure 3, Table S7. For four residue positions (54, 81, 103 and 178) differed between proteins by more than twice the median absolute deviation within each simulation. Correlation analysis, in silico and in vitro mutational studies investigated the origins of residue position differences. Exchanging valine and phenylalanine at the second shell position 54 reduced the largest differences in the shape, but did not account for difference in reactivity and enantioselectivity. The supporting information contains details, as well as suggestions regarding the origins of the remaining differences in the consensus active sites. Position 103 was not investigated further in this study, It lies outside the active site, and as described above, changing His103 from *Hb*HNL to valine similar to leucine in *A*HNL, do not affect activity or enantioselectivity, and only affect protein stability. The modeling used histidine at

this location because the x-ray structure showed its side chain conformation, while the position of the valine side chain was less certain.

DISCUSSION

Switching a few of the sixteen residues that differ in the active sites of *HbHNL* and *AtHNL* disrupted catalysis or folding, while switching all sixteen residues disrupted folding. Switching catalytic residues Lys236 or Thr11 in *HbHNL* to corresponding residues in *AtHNL* eliminated catalytic activity. Presumably, unfavorable interactions between the active site residues of one HNL and the surrounding second shell residues of the other HNL disrupted protein folding and catalysis.

Surprisingly, switching about half of the active site residues in *HbHNL* to make the A9-H7 or A10-H6 composite active sites yielded catalytically active, stable enzymes. The first surprise was an *HbHNL* variant with eight substitutions that expressed as soluble, folded protein. The second surprise was an additional substitution (9 total) that restored HNL activity (*Hb-A9-H7*). These composite active sites resemble the active site of ancestral enzyme EST3ml. Several substitutions to make this variant more like wild type *HbHNL* (Met236 back to Lys or Phe79 back to Glu) yielded only insoluble protein. Similarly, additional substitutions that made it more like wild-type *AtHNL* (Phe125 to Pro or Trp128 to Leu) also yielded insoluble protein. Although we did not identify a path of single substitutions yielding soluble and catalytically active variants at each step, we did not test all possible combinations of these 16 substitutions.

Variant *HbHNL*'s with composite active sites (*Hb-A9-H7* and *Hb-A10-H6*) likely use a catalytic mechanism different from either wild type protein. The nine substitutions that convert wild type *HbHNL* to *Hb-A9-H7* eliminate the possibility of the (*S*)-selective catalytic Lys mechanism of *HbHNL*⁹ because these substitutions remove the essential Thr11 and Lys236 residues. While we expected that the addition of Asn11, Ala12, and Phe81 would enable the (*R*)-selective catalysis of *AtHNL*¹¹, *HbHNL-A9-H7* remained highly (*S*)-selective (>39). The ability to maintain ~3% of wild type activity without either catalytic residue suggests these composite enzymes catalyze (*S*)-selective mandelonitrile cleavage using a novel mechanism. Modeling suggested that Asn11 takes on the role of Lys236 in stabilizing the CN⁻ leaving group, hence we named this the Asn mechanism. A change in mechanism is also consistent with the "island of activity". The inability of *Hb-A9-H7* to catalyze cleavage of acetone cyanohydrin also supports the conclusion that it has lost the Lys mechanism.

Three structural elements contribute at least 1 kcal/mol to the enantioselectivity of these enzymes: six active site residues (13, 121, 125, 128, 131, & 209), residue 178, and residues outside the active site. The enantioselectivity of the two wild-type enzymes is >39 favoring opposite enantiomers. The difference corresponds to a $G^\ddagger$ between the two wild type enzymes of >4.3 kcal/mol. *AtHNL* and *At-A10-H6* differ by only six residues (13, 121, 125, 128, 131, & 209) near the aromatic ring binding region. Changing the six residues shifts the enantioselectivity from >39 to 5.3 (>1.2 kcal/mol to $G^\ddagger$), and increased the number of catalytically active conformations found *in silico* almost five fold for the slow enantiomers

while their numbers decreased slightly for the fast enantiomer. The second feature determining enantioselectivity is residue 178. Changing Phe to Leu at 178 changed enantioselectivity by $\Delta G^\ddagger$ of 1.3 and over 1.4 kcal/mol when converting from the A9-H7 to the A10-H6 active sites in the *A*HNL and *Hb*HNL scaffolds, respectively. The third feature that determines enantioselectivity is residues outside of the active site. The enantioselectivity differences between identical composite active sites in differing scaffolds correspond to 1.7 kcal/mol in the A9-H7 sites and >1.9 kcal/mol in the A10-H6 sites. While 121 residues outside of the active site differ between these two proteins, the overall structure is similar (RMS = 0.74 Å between 889 peptide backbone atoms). This finding supports the other results that, even without major structural changes, residues outside of the active site are important for determining enantioselectivity^{34–39}.

The new asparagine mechanism proposed in the composite site enzymes may have been an evolutionary stepping-stone to modern HNL mechanisms. Both *Hb*HNL and *A*HNL evolved from esterases, yet a direct transition between modern esterase and lysine mechanism HNL requires going through catalytically dead intermediates due to needing to both block the oxyanion hole and add and position a charged lysine to bind the leaving cyanide^{27,28}. However, the reconstruction of the last common ancestor of *Hb*HNL and *A*HNL, called EST3ml, catalyzed both cyanohydrin cleavage and ester hydrolysis^{8,17}. This ancestor's active site resembles the composite active site enzymes created here, including the critical Asn11. All active site residues in the ancestor either match A10-H6 (17 residues) or are residues not found in either modern enzyme (5 residues). EST3ml shows 2.3% of the HNL activity of *Hb*HNL and shows low enantioselectivity (*R*, 7) similar to the activity (1.6% of *Hb*HNL) and enantioselectivity (*S*, 3.3) of *Hb*HNL-A10-H6. Like *Hb*HNL-A10-H6, EST3ml has a Phe at both 54 and 81, which modeling suggests could obstruct the oxyanion hole, thereby preventing HNL mechanism used by *A*HNL. Since *Hb*HNL-A10-H6 likely follows the Asn mechanism, we propose that ancestral enzyme EST3ml likewise follows the Asn mechanism. This suggests that finding the new asparagine mechanism in the composite enzymes may not be due to serendipity, but instead may be uncovering the ancestral stepping-stone mechanism between ancestral esterases and modern HNLs.

A stepping stone mechanism would be compatible with both previous (esterase) and subsequent mechanisms (modern HNL mechanisms) allowing an evolutionary path from one to the other via a promiscuous enzyme, with a functional enzyme at every step. Since both the composite site enzymes and the ancestor EST3ml are significantly slower than modern, wild type HNLs, evolutionary pressure could optimize HNL function by converting from an asparagine mechanism to the modern lysine or oxyanion hole mechanisms if a viable evolutionary path exists. A few mutants along a few possible trajectories were tested and found to be dead or insoluble (Table S4), but these only represent a small sample of the 9! (>360,000) possible direct evolutionary routes to mutate the 9 residues that differ between *Hb*-A9-H6 and wild type *Hb*HNL.

EXPERIMENTAL SECTION

General

Amino acid numbering is based upon the *HbHNL* sequence. *AhHNL* contains an additional residue within the first five amino acid residues (not clear which one is the insertion), so the *AhHNL* numbering is the *HbHNL* numbering reported here plus one.

Structure comparison

Mutations focused on the active site residues that differ between *AhHNL* and *HbHNL*. The active site residues were defined as those with a heavy atom within 6 Å of the heavy atoms of the substrate in either protein. Structures used were *HbHNL* PDB: 1YB6, which contained the (*S*)-mandelonitrile substrate and *AhHNL* PDB: 3DQZ chain A, which did not contain a substrate. The two proteins were superimposed in PyMOL⁴². To identify the active site residues for *AhHNL*, the mandelonitrile position was assumed to be identical to that observed in *HbHNL*. Table S1 lists the distances to the closest 45 residues; 22 of which lie within 6 Å of the substrate binding site and are the focus of this paper.

Hydroxynitrile lyase activity

Mandelonitrile cleavage—The cleavage of racemic mandelonitrile (9 mM) in sodium citrate buffer (54 mM, pH 5.1 also containing 250 µM BES buffer) was measured in 96-well microplates (200 µL total volume per well; path length 0.6 cm) containing enzyme (20–300 pmol enzyme/well). The formation of benzaldehyde was detected over five minutes by an increase in absorbance at 280 nm measured with a SpectraMax Plus 384 plate reader (Molecular Devices). The increases were corrected for the spontaneous cleavage observed in no-enzyme control wells, and converted to turnovers per minute using extinction coefficient 1352/M cm for benzaldehyde.

Acetone cyanohydrin cleavage—The release of cyanide from acetone cyanohydrin was measured in triplicate using a modified König reaction^{43–45} in microtiter plates. The reactions contained acetone cyanohydrin (5 mM) in sodium citrate buffer (50 mM, pH 5.1) and enzyme (3–700 nM). After 5 minutes at ambient temperature, the reactions were quenched by the addition of *N*-chlorosuccinimide (2 mM, 62.5 µL also containing a 10-fold w/w excess of succinimide) to oxidize the released cyanide to a cationic species. After 2 min, barbituric acid (230 mM in 30% pyridine, 12.5 µL) was added. The reaction of the cationic species with barbituric acid was monitored for 5 min by following the increase in absorbance at 600 nm. The slope of the absorbance increase was compared to a standard curve prepared using a cyanide standard (Merck #1.19533.0500).

Enantioselectivity

The preferred enantiomer was determined by comparing the rates of cleavage of (*R*)-mandelonitrile (4.5 mM) and racemic mandelonitrile (9 mM) using the assay described above. Enantiopure *S*-mandelonitrile is not commercially available. (*S*)-Selective enzymes showed much higher activity with racemic mandelonitrile than with (*R*)-mandelonitrile. (*R*)-Selective enzymes might show similar activity with both racemic and (*R*)-mandelonitrile if the (*S*)-mandelonitrile in the racemate had no effect. In practice, (*R*)-selective enzymes

showed higher activity with (*R*)-mandelonitrile than with racemic mandelonitrile, suggesting that (*S*)-mandelonitrile in the racemate inhibited the activity.

The enantioselectivity of the HNL's toward mandelonitrile were measured by the enzyme catalyzed synthesis of mandelonitrile from benzaldehyde and excess hydrogen cyanide. The reaction (1 mL total volume) contained benzaldehyde (500 μ L of a 40 mM stock solution in *tert*-butyl methyl ether also containing 50 mM phenol as an internal standard) and HCN (400 μ L of a ~250 mM stock solution in *tert*-butyl methyl ether). An aqueous solution of enzyme (100 μ L containing 1–100 nmol of enzyme in sodium citrate buffer, 50 mM, pH 5.0, and a small amount of 4.5 mM BES buffer from the enzyme storage buffer) was added and the two-phase mixture was shaken at 700 rpm for 1 h at 15 °C. Caution! HCN is toxic and requires careful handling (see⁴⁶). Isopropanol (1 mL) was added to the reaction mixture to yield a single liquid phase. The solution was filtered through a syringe filter (0.45 μ m pores) to remove precipitated protein. The filtrate was analyzed by HPLC on a Chiralcel OD-H column (Diacel) eluted with hexane:isopropanol (97:3) at flow rate of 1.0 mL/min and monitored at 210 nm. The *S* and *R* enantiomers elute at 25 and 27 min, respectively. Absolute configuration of mandelonitrile was established by comparison with a commercial sample of (*R*)-mandelonitrile (Alfa Aesar, Ward Hill, MA). Peak areas were normalized using the phenol internal standard peak, which eluted at 15 min and were corrected for spontaneous reaction using a control sample without enzyme. The enantioselectivity, *E*, was given by $E = P/N$, where “P” is corrected amount of the preferred enantiomer produced and “N” is the corrected amount of the non-preferred enantiomer produced.

Molecular dynamics simulations

The flexible behavior of *A*HNL wild type, *Hb*HNL wild type, A9-H7 and A10-H6 composite active sites of both scaffolds was simulated by molecular dynamics using the Desmond⁴⁷ engine, in apo form and with (*R*) and (*S*)-mandelonitrile bound in the active site. The supporting information details the modeling of composite active sites and ligand docking. All proteins were simulated as their physically relevant dimers. The protein was solvated in TIP3P water molecules in an orthorhombic box with 10 Å of solvent in each direction beyond the protein. The system was neutralized using Na⁺ atoms, and ions corresponding to 0.05 M NaCl were added, to model the experimental condition. Using a combined steepest descent and conjugate gradient, the geometry of the systems was optimized until the energy changed by <1.0 kcal/mol·Å. None of the models showed significant movement at this step. The models were gradually heated to the simulation temperature (303 °K), and equilibrated before the molecular dynamics simulations using the default five step relaxation procedure in Multisim. Following the equilibration three repeats of 10 ns unrestrained molecular dynamics simulations were carried out, with different random starting velocities, in NPT conditions, using default settings in Maestro.⁴⁸ Energy conservation, constant temperature and constant pressure of the simulations was checked with the Simulation Quality Analysis module in Maestro. The simulations were stable and convergent because the RMSD of the protein structure returned to a value close to that of the starting conformation.

Ligand positions during the molecular dynamics simulation were analyzed by assigning each frame to one of four binding modes: Lys binding, Oxyanion binding, Asn binding and non catalytically active binding. The assignment was based on the hydrogen bonding of the hydroxyl group and on the interaction of the cyano group. Data for the assignment was obtained by measuring atomic distances, angles and dihedral angles between atoms in the protein and the ligand the Simulation Event Analysis engine in Maestro. Table S3 contains the complete list of the measured properties.

Residue positions characterize the shape of the active sites. The distances between the centers of mass of residues listed in Table S1 and Ser80 were measured in each frame for all trajectories with a custom tcl script in vmd.⁴⁹ Ser80 occupies a central position in the active site, and a hydrogen bond to His235 minimized movement of the Ser80 side chain. The distance between the centers of mass of mandelonitrile and Ser80 were also measured in each frame. All data analyses were performed using R.⁵⁰ The median distance between Ser80 and each residue measured the residue's position, while median absolute deviation ($MAD = \text{median}(X_i - \text{median}(X))$) measured the normal movement of each residue during the simulation. The use of median, instead of mean, minimized the effect of outlier distances. If the median residue positions in a composite active site differed in the two scaffolds by more than the median movement of these residues (MAD), then we considered the positions of these residues to differ in the two scaffolds. These differences may contribute to the different catalytic activity of the same composite active site within different scaffolds.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We thank the US National Institutes of Health (1R01GM102205-01), the National Science Foundation (CHM-1152804) and the Fulbright Visiting Student Researcher Program to ZB for financial support, the Minnesota Supercomputing Institute for use of computers and software, and Mark Lunzer and Antony Dean for helpful discussions and data for several AtHNL single substitution variants.

References

1. Dadashipour M, Asano Y. *ACS Catal.* 2011; 1:1121–1149.
2. Bracco P, Busch H, von Langermann J, Hanefeld U. *Org Biomol Chem.* 2016; 14:6375–6389. [PubMed: 27282284]
3. Dreveny I, Andryushkova AS, Glieder A, Gruber K, Kratky C. *Biochemistry.* 2009; 48:3370–3377. [PubMed: 19256550]
4. Breithaupt H, Pohl M, Bönigk W, Heim P, Schimz K-L, Kula M-R. *J Mol Catal B: Enzym.* 1999; 6:315–332.
5. Lauble H, Miehl B, Förster S, Wajant H, Effenberger F. *Biochemistry.* 2002; 41:12043–12050. [PubMed: 12356304]
6. Hajnal I, Lyskowski A, Hanefeld U, Gruber K, Schwab H, Steiner K. *FEBS J.* 2013; 280:5815–5828. [PubMed: 23981508]
7. Dadashipour M, Ishida Y, Yamamoto K, Asano Y. *Proc Natl Acad Sci U S A.* 2015; 112:10605–10610. [PubMed: 26261304]
8. Devamani T, Rauwerdink AM, Lunzer M, Jones BJ, Mooney JL, Tan MAO, Zhang Z-J, Xu J-H, Dean AM, Kazlauskas RJ. *J Am Chem Soc.* 2016; 138:1046–1056. [PubMed: 26736133]

9. Gruber K, Gartler G, Krammer B, Schwab H, Kratky C. J Biol Chem. 2004; 279:20501–20510. [PubMed: 14998991]
10. Stranzl GR, Gruber K, Steinkellner G, Zangger K, Schwab H, Kratky C. J Biol Chem. 2004; 279:3699–3707. [PubMed: 14597632]
11. Andexer JN, Staunig N, Eggert T, Kratky C, Pohl M, Gruber K. ChemBioChem. 2012; 13:1932–1939. [PubMed: 22851196]
12. Zhu W, Liu Y, Zhang R. Proteins. 2015; 83:66–77. [PubMed: 25052541]
13. Christ B, Schelbert S, Aubry S, Süßenbacher I, Müller T, Kräutler B, Hörtensteiner S. Plant Physiol. 2012; 158:628–641. [PubMed: 22147518]
14. Yang Y, Xu R, Ma C-J, Vlot AC, Klessig DF, Pichersky E. Plant Physiol. 2008; 147:1034–1045. [PubMed: 18467465]
15. Vlot AC, Liu P-P, Cameron RK, Park S-W, Yang Y, Kumar D, Zhou F, Padukkavidana T, Gustafsson C, Pichersky E, Klessig DF. Plant J. 2008; 56:445–456. [PubMed: 18643994]
16. Koo YJ, Yoon ES, Seo JS, Kim J-K, Do Choi Y. J Korean Soc Appl Biol Chem. 2013; 56:27–33.
17. Rauwerdink A, Lunzer M, Devamani T, Jones B, Mooney J, Zhang Z-J, Xu J-H, Kazlauskas RJ, Dean AM. Mol Biol Evol. 2016; 33:971–979. [PubMed: 26681154]
18. Andexer J, von Langermann J, Mell A, Bocola M, Kragl U, Eggert T, Pohl M. Angew Chem, Int Ed. 2007; 46:8679–8681.
19. Okrob D, Metzner J, Wiechert W, Gruber K, Pohl M. ChemBioChem. 2012; 13:797–802. [PubMed: 22378361]
20. Scholz KE, Okrob D, Kopka B, Grünberger A, Pohl M, Jaeger K-E, Krauss U. Appl Environ Microbiol. 2012; 78:5025–5027. [PubMed: 22544258]
21. Tattersall DB, Bak S, Jones PR, Olsen CE, Nielsen JK, Hansen ML, Høj PB, Møller BL. Science. 2001; 293:1826–1828. [PubMed: 11474068]
22. Rajniak J, Barco B, Clay NK, Sattely ES. Nature. 2015; 525:376–379. [PubMed: 26352477]
23. Kjaer, A. Chemistry in Botanical Classification: Medicine and Natural Sciences. Bendz, G., editor. Academic Press; New York: 1974. p. 229–234.
24. Rodman JE, Karol KG, Price RA, Sytsma KJ. Syst Bot. 1996; 21:289–307.
25. Wittstock U, Burow M. The Arabidopsis Book. 2010; 8:e0134. [PubMed: 22303260]
26. Bergthorsson U, Andersson DI, Roth JR. Proc Natl Acad Sci U S A. 2007; 104:17004–17009. [PubMed: 17942681]
27. Nedrud DM, Lin H, Lopez G, Padhi SK, Legatt GA, Kazlauskas RJ. Chem Sci. 2014; 5:4265–4277. [PubMed: 25346843]
28. Padhi SK, Fujii R, Legatt GA, Fossum SL, Berchtold R, Kazlauskas RJ. Chem Biol. 2010; 17:863–871. [PubMed: 20797615]
29. Gartler G, Kratky C, Gruber K. J Biotechnol. 2007; 129:87–97. [PubMed: 17250917]
30. Lim, HY. MS Thesis. University of Minnesota; St. Paul, MN: 2016. Improving the unnatural and promiscuous nitroaldolase activity of modern and ancestral HNLs.
31. Asano, Y., Akiyama, T., Yu, F., Sato, E. Substitutional variants of hydroxynitrile lyase with high specific activity and methods of use. U.S. Patent 8,030,053. Oct 4. 2011
32. Asano Y, Dadashpour M, Yamazaki M, Doi N, Komeda H. Protein Eng, Des Sel. 2011; 24:607–616. [PubMed: 21729945]
33. The free energy difference in the enantioselectivity of different enzymes ($\Delta G^\ddagger$) was given by: $\Delta G^\ddagger = -RT \ln(E_{wt}/E_{mut})$ where R is the gas constant, T is temperature (in Kelvin), E_{wt} is the enantioselectivity of the wild type enzyme, and E_{mut} is that of the mutant enzyme. The enantioselectivity values must be for the same enantiomer, thus a change from 5.3 (R) to 1.6 (S) would be indicated as a change from 5.3 (R) to 1.6⁻¹ (R). The energy value, $\Delta G^\ddagger$, has three deltas to signify (1) the Gibbs free energy of activation for the reaction, (2) the difference in reactivity between enantiomers (3) the difference in enantioselectivity between enzymes.
34. Morley KL, Kazlauskas RJ. Trends Biotechnol. 2005; 23:231–237. [PubMed: 15866000]
35. Horsman GP, Liu AMF, Henke E, Bornscheuer UT, Kazlauskas RJ. Chem—Eur J. 2003; 9:1933–1939. [PubMed: 12740839]

36. Henke E, Bornscheuer UT. *Biol Chem.* 1999; 380:1029–1033. [PubMed: 10494857]
37. Reetz MT, Wu S. *J Am Chem Soc.* 2009; 131:15424–15432. [PubMed: 19807086]
38. Liebeton K, Zonta A, Schimossek K, Nardini M, Lang D, Dijkstra BW, Reetz MT, Jaeger KE. *Chem Biol.* 2000; 7:709–718. [PubMed: 10980451]
39. Robertson DE, Chaplin JA, DeSantis G, Podar M, Madden M, Chi E, Richardson T, Milan A, Miller M, Weiner DP, Wong K, McQuaid J, Farwell B, Preston LA, Tan X, Snead MA, Keller M, Mathur E, Kretz PL, Burk MJ, Short JM. *Appl Environ Microbiol.* 2004; 70:2429–2436. [PubMed: 15066841]
40. Devamani T, Rauwerdink AM, Lunzer M, Jones BJ, Mooney JL, Tan MAO, Zhang Z-J, Xu J-H, Dean AM, Kazlauskas RJ. *J Am Chem Soc.* 2016; 138:1046–1056. [PubMed: 26736133]
41. Dadashpour M, Yamazaki M, Momonoi K, Tamura K, Fuhshuku K-I, Kanase Y, Uchimura E, Kaiyun G, Asano Y. *J Biotechnol.* 2011; 153:100–110. [PubMed: 21352863]
42. Schrodinger, LLC. The PyMOL Molecular Graphics System. 2010. www.pymol.org
43. König W. *Adv Synth Catal.* 1904; 69:105–137.
44. Selmar D, Carvalho FJ, Conn EE. *Anal Biochem.* 1987; 166:208–211. [PubMed: 3674409]
45. Andexer J, Guterl JK, Pohl M, Eggert T. *Chem Commun.* 2006; (40):4201–4203.
46. Occupational safety and health guideline for hydrogen cyanide. U.S. Centers for Disease Control and Prevention; 1978. www.cdc.gov/niosh/topics/hydrogen-cyanide/
47. Shaw, DE. Desmond Molecular Dynamics System. D. E. Shaw Research; New York, NY: 2015. www.schrodinger.com
48. Maestro, version 10.4. Schrödinger, LLC; New York, NY: 2015.
49. Humphrey W, Dalke A, Schulten K. *J Mol Graph.* 1996; 14:33–38. 27–28. [PubMed: 8744570]
50. R Core Team. R: A language and environment for statistical computing. R Foundation for Statistical Computing; 2016. www.R-project.org

Figure 1.

Different catalytically competent orientations of mandelonitrile stabilize the leaving cyanide group differently and favor different enantiomers. a) In the classical mechanism for *HbHNL*, the ϵ -ammonium group of Lys236 stabilizes the leaving cyanide. This enzyme favors the (*S*)-enantiomer. b) In the mechanism for *AhHNL*, two main chain N–H's from the oxyanion hole stabilize the leaving cyanide. This enzyme favors the (*R*)-enantiomer. This enzyme cannot use the classical mechanism because it lacks Lys236. c) In the mechanism proposed here for the (*S*)-enantioselective composite active site enzymes, derived from *HbHNL*, discussed in this paper, the δ -amide N–H of Asn11 stabilizes the cyanide leaving group. The other two mechanisms are not possible because Lys236 is missing and because distant residues cause Phe81 to block the oxyanion hole region.

Figure 2.

Four different active sites that were examined are laid out horizontally. Active site residues that differ between the enzymes are listed in yellow if *HbHNL* like and cyan if *AtHNL* like. Residues within 6 Å of the active site are shown by filled circles, while the second shell residue is represented by an empty circle. Each of these four active sites were made in both *HbHNL* (top row) and in *AtHNL* (bottom row). Completely swapping the active sites failed to produce soluble protein (red cross). The two composite active sites resulted in soluble protein in either scaffold, and the enantiomeric ratios are indicated in the boxes. Yellow indicates *HbHNL* like (*S*)-selective enzyme, while cyan indicates *AtHNL*-like (*R*)-selective enzyme. Enantioselectivities were determined by enzymatic synthesis of mandelonitrile and analysis on chiral HPLC.

Figure 3.

Among the 44 residues closest to the substrate, the positions of seventeen differ between *HbHNL* and *AaHNL* scaffolds. Residue positions are defined with respect to Ser80 (red sticks). *AaHNL*-A10-H6 (cyan) contains (*R*)-mandelonitrile (light cyan sticks), while *HbHNL*-A10-H6 (yellow) contains (*S*)-mandelonitrile (light yellow sticks). a) Overview of the seventeen differing residues (red). b) Four residues (stick representation) differ between the scaffolds by more than twice the amount that they move within each protein, while the remaining thirteen residues (thin lines) differ by more than they move within each protein. For clarity, this image is rotated approximately 180° compared to the other two images. c) Detail of the four residues with most different positions. The side chain conformation of Phe81 differs notably between the scaffolds; labels show the average dihedral angles.

Figure 4.

a) Simplified phylogenetic tree including *HbHNL* (Lys mechanism, yellow text), *AtHNL* (oxyanion hole mechanism, cyan text), HNL from *Baliospermum montanum* (*BmHNL*), esterase from *Nicotiana tabacum* (*SABP2*, blue text), as well as several reconstructed ancestral enzymes (*HNL1*, *EST2*, *EST3ml*). The numbers in black are the HNL-activity for cleavage of mandelonitrile in min^{-1} , the favored enantiomer & the enantioselectivity are in parentheses⁴⁰, and the active site residues associated with HNL activity (residues 11 and 236) are shown.^a Data for *BmHNL* are from Dadashipour et. al.⁴¹ b) The active site of *EST3ml* resembles the consensus active site A10-H6, only four of the sixteen investigated residues differ, these are shown in green. *EST3ml* contains phenylalanines at both positions 54 and 81, suggesting that the Phe81 may block the oxyanion hole, and thus require the Asn mechanism for catalysis.

Table 1

Catalytically active^a conformations (CAC) from molecular dynamics simulations^b.

	(R)-mandelonitrile				(S)-mandelonitrile			
	CAC [%]	CN-partner	% of CAC	OH-partner	% of CAC	CN-partner	% of CAC	OH-partner
wt- <i>HbHNL</i>	20	Thr11	88	Ser80	94	Lys236	99	Thr11
wt- <i>AdHNL</i>	44	Oxyanion hole	100	Ser80	95	Oxyanion hole	100	His235
<i>Af</i> -A10-H6	35	Oxyanion hole	100	Ser80	95	Oxyanion hole	100	His235
<i>Af</i> -A9-H7	54	Oxyanion hole	97	Ser80	99	Oxyanion hole	100	His235
<i>Hb</i> -A10-H6	13	Asn11	82	His235	78	Asn11	99	Ser80
<i>Hb</i> -A9-H7	13	Asn11	100	His235	85	Asn11	100	Ser80

^aA conformation is considered catalytically active if the hydroxyl group of mandelonitrile forms a hydrogen bond to a potential proton acceptor, and its cyano group interacts with a partially positive hydrogen atom. Hydrogen bond criteria (D-H...A distance 2.5 Å, D-H...A angle 90°, H...A-C angle 90°³¹. Additional criteria was employed for histidine, as the lone electron pair, participating in the hydrogen bond formation, of the sp² type nitrogen in the imidazole ring resides in the plane of the ring³². For a H-bond to His235 the torsion angle H-N-C-N was restricted to 20° out of the plane angles. For catalytically active binding with Ser80 base, Ser80 → His235 hydrogen bond must also be present.

^bThe favored enantiomers are in bold; for each protein data is from three repeated 10-ns simulations.