The active site of hydroxynitrile lyase from *Prunus amygdalus*: Modeling studies provide new insights into the mechanism of cyanogenesis

INGRID DREVENY, CHRISTOPH KRATKY, AND KARL GRUBER

Institut für Chemie, Karl-Franzens Universität Graz, Heinrichstraße 28, A-8010 Graz, Austria

(RECEIVED September 14, 2001; FINAL REVISION November 2, 2001; ACCEPTED November 2, 2001)

Abstract

The FAD-dependent hydroxynitrile lyase from almond (*Prunus amygdalus*, PaHNL) catalyzes the cleavage of *R*-mandelonitrile into benzaldehyde and hydrocyanic acid. Catalysis of the reverse reaction—the enantiospecific formation of α -hydroxynitriles—is now widely utilized in organic syntheses as one of the few industrially relevant examples of enzyme-mediated C–C bond formation. Starting from the recently determined X-ray crystal structure, systematic docking calculations with the natural substrate were used to locate the active site of the enzyme and to identify amino acid residues involved in substrate binding and catalysis. Analysis of the modeled substrate complexes supports an enzymatic mechanism that includes the flavin cofactor as a mere “spectator” of the reaction and relies on general acid/base catalysis by the conserved His-497. Stabilization of the negative charge of the cyanide ion is accomplished by a pronounced positive electrostatic potential at the binding site. PaHNL activity requires the FAD cofactor to be bound in its oxidized form, and calculations of the pK_a of enzyme-bound HCN showed that the observed inactivation upon cofactor reduction is largely caused by the reversal of the electrostatic potential within the active site. The suggested mechanism closely resembles the one proposed for the FAD-independent, and structurally unrelated HNL from *Hevea brasiliensis*. Although the actual amino acid residues involved in the catalytic cycle are completely different in the two enzymes, a common motif for the mechanism of cyanogenesis (general acid/base catalysis plus electrostatic stabilization of the cyanide ion) becomes evident.

Keywords: Cyanohydrins; C–C couplings; biocatalysis; enzyme mechanism; electrostatics

The degradation of α -cyanohydrins into hydrocyanic acid (HCN) and the respective aldehyde or ketone (Fig. 1) constitutes the key step in a process known as cyanogenesis in higher plants, ferns, bacteria, fungi, and insects (Conn 1981). This cleavage reaction is catalyzed by hydroxynitrile lyases (HNL), and the release of HCN supposedly serves as a defense mechanism against herbivores and microbial attack (Nahrstedt 1985).

Already in 1908, an HNL containing enzyme preparation from almonds was utilized as a catalyst in asymmetric syn-

theses (Rosenthaler 1908). More recently, hydroxynitrile lyases have again received increased attention as industrial biocatalysts for the synthesis of chiral compounds by exploiting the reverse of the biological reaction, that is, the stereospecific addition of HCN to aldehydes or ketones (Effenberger 1994; Griengl et al. 1997). In fact, this reaction represents one of the few industrially relevant cases of enzyme-mediated carbon–carbon bond formation (Fessner and Walter 1996).

HNLs from a variety of mainly plant species have been characterized and classified, depending on the presence or absence of an FAD cofactor (Hickel et al. 1996; Wajant and Effenberger 1996). Although FAD-independent HNLs (EC 4.1.2.11, 4.1.2.37, and 4.1.2.39) are quite heterogeneous with regard to their biological substrate, mass, and sequence, all FAD-dependent HNLs (EC 4.1.2.10) exclu-

Reprint requests to: Karl Gruber, Institut für Chemie, Karl-Franzens Universität Graz, Heinrichstraße 28, A-8010 Graz, Austria; e-mail: karl.gruber@uni-graz.at; fax: (43 316) 380-9850.

Article and publication are at <http://www.proteinscience.org/cgi/doi/10.1110/ps.38102>.

Fig. 1. Reaction catalyzed by the hydroxynitrile lyase from almond (*Prunus amygdalus*, PaHNL).

sively accept *R*-mandelonitrile (**1**) as their natural substrate, and are distantly related (with up to 30% sequence identity) to members of the glucose–methanol–choline (GMC) oxidoreductase family (Kiess et al. 1998). The HNL bound cofactor can be generated in three oxidation states (oxidized, semiquinone, and reduced form; Fig. 2), of which only the oxidized form **3** shows enzymatic activity (Bärwald and Jaenicke 1978; Jorns 1979). As in oxidases, a sulfite adduct can be formed, and the reduced cofactor species are oxidized by air (Massey et al. 1969). Despite these similarities, however, FAD-dependent HNLs, like only very few other flavoenzymes, appear not to exploit the unique redox properties of the cofactor (Jorns 1979), and the FAD was proposed to have a mere “structural role” (Gerstner et al.

Fig. 2. Chemical constitution of the isoalloxazine portion of the FAD cofactor in various redox and protonation states.

1968; Jorns 1979; Cheng and Poulton 1993) or to be an evolutionary remnant of a truly flavin dependent ancestor (Jorns 1979).

Recently, the first X-ray crystal structure of an FAD-dependent HNL (the enzyme from almond, *Prunus amygdalus*, PaHNL) was determined at high resolution (Dreveny et al. 2001). In line with the moderate sequence homology, this analysis revealed a close structural similarity of PaHNL to two enzymes from the family of GMC–oxidoreductases, glucose oxidase (Hecht et al. 1993; Wohlfahrt et al. 1999), and cholesterol oxidase (Vrieling et al. 1991; Yue et al. 1999). The similarities were most evident in domains responsible for FAD binding, but also extended to the active site region of the oxidases. It was, therefore, inferred that the active site of PaHNL is located in the corresponding region of this enzyme, close to the isoalloxazine ring of the cofactor and to the side chain of a histidine residue (His-459) that is strictly conserved among all GMC–oxidoreductases and FAD-dependent HNLs (Dreveny et al. 2001).

So far, however, neither mutational data nor structures of PaHNL substrate (or inhibitor) complexes are available. To obtain independent evidence for the location of the active site and—even more important—information on the mode of substrate binding, we performed systematic docking simulations using the biological substrate mandelonitrile **1** as a ligand (Fig. 1). Because our aim was to induce as little bias as possible regarding the location of the active site, we did not restrict the calculations to certain regions of the enzyme, but rather searched the whole volume occupied by the protein.

The docking simulations clearly located the active site of PaHNL, which indeed appears to be in close vicinity to the FAD. In addition, the calculations identified amino acid residues that are most likely involved in the catalytic cycle of the enzyme. Analysis of the electrostatics at the substrate binding site provided a rationalization of the enzymatic activity and a likely explanation for the observed absolute requirement of oxidized FAD as cofactor.

Results

Location of the active site

We carried out docking calculations to identify low-energy binding modes of the *R*-enantiomer of mandelonitrile **1** (Fig. 1) to the FAD-dependent HNL from almonds (PaHNL). This search was split up into separate simulations within 45 smaller boxes that were systematically spread over the whole protein volume. The pooled results obtained with two different versions (2.4 and 3.0) of the program AutoDock (Morris et al. 1996, 1998) are summarized in Table 1. These two versions of the program differ with respect to the empirical force field that is employed. The v2.4 force field consists of parameters for van der Waals,

Table 1. Docking results (kcal/mol) for the binding of R-mandelonitrile **1** to PaHNL

Rank	Autodock 2.4	Autodock 3.0
1	-36.62^a	-7.84 (2.2) ^b
2	-30.89	-6.40 (21.0)
3	-30.29	-6.35 (-4.0)^a
4	-30.14	-6.35 (38.8)
5	-30.14	-5.85 (1.2)
6	-29.56	-5.81 (23.2)
7	-29.37	-5.75 (3.7)
8	-29.28	-5.66 (4.3)
9	-29.17	-5.66 (19.8)
10	-29.10	-5.65 (5.1)

^a Bold energy values correspond to substrate binding as shown in Figure 3.

^b Energy values given in parentheses were calculated using the Carnal module from the AMBER 6.0 suite of programs (Case et al. 1997) after minimization.

electrostatic, and H-bonding interactions, while in AutoDock v3.0 additional empirical solvation parameters are included, which allow the estimation of binding free energy values (Morris et al. 1998). It has to be noted that these different parameters yield energy values that are not comparable on an absolute scale.

The simulations with AutoDock v2.4 (Morris et al. 1996) yielded a very clear binding site that is well separated in energy from alternative sites. The next best solution is about 6 kcal/mol less favorable, whereas the energies of the remaining top 10 solutions are lying within only 2 kcal/mol. The thus identified substrate binding site is situated close to the FAD cofactor of PaHNL (Fig. 3A). Possible hydrogen bonding interactions exist between the hydroxy group of the substrate and the side chains of residues His-497, Tyr-457, and Cys-328. The cyano group of **1** points towards the N5-atom of the flavin and the side chain of residue His-459, whereas the phenyl group is bound in a mostly hydrophobic pocket built up by the side chains of residues Phe-72, Ala-111, Arg-300, Val-316, Phe-330, Phe-342, His-357, Ala-359, and Trp-458.

To cross check our docking results we repeated the calculations with AutoDock v3.0 (Morris et al. 1998). Here, the third best solution (Table 1) closely corresponds to the binding mode identified previously with AutoDock v2.4 (Fig. 3A), while the top two sites are located near the biphosphate group of the cofactor. The remaining top 10 sites are distributed over four locations on the surface of the protein.

Only one binding mode scored among the top results in both calculations (Fig. 3A). On the contrary, the binding sites close to the biphosphate were not found among the top 20 results obtained with AutoDock v2.4, and yielded an at least 9 kcal/mol less favorable docking energy compared to the best solution. The top 10 results from the calculations with AutoDock v3.0 were further analyzed for their respec-

tive molecular mechanics interaction energy between substrate and protein. This analysis yielded a revised order of the energy of these complexes (Table 1), in which the binding mode of **1** as depicted in Figure 3A ranks first, and shows a difference in interaction energy of at least 5 kcal/mol compared to the remaining solutions.

pK_a of enzyme-bound HCN

We estimated the shift of the pK_a value of PaHNL-bound HCN relative to an aqueous solution (pK_a = 9.2) using the program DELPHI (Nicholls and Honig 1991), and investigated the influences of the redox and protonation state of the FAD cofactor (Fig. 2), as well as of charged residues in the vicinity of the active site. All calculations were performed with and without benzaldehyde bound in the active site, and were repeated for two different positions of HCN/CN⁻ obtained from the docked complex of **1** and the modeled complex of the enzyme with benzaldehyde **2** and CN⁻ (Fig. 3). The results of these calculations are shown in Table 2.

In the native enzyme (containing oxidized FAD **3**) the electrostatic contributions of surrounding charged residues appear to be sufficient to counterbalance the desolvation penalty that is due to the transfer of HCN/CN⁻ from bulk water ($\epsilon_r = 80$) to the protein interior ($\epsilon_r = 4$). As a consequence, the pK_a of enzyme-bound HCN is virtually unchanged. Moreover, the cyanide ion is stabilized by about 2.7 kcal/mol in the presence of benzaldehyde, leading to a decrease of the pK_a (an increase in the acidity) of HCN by two units. For the semiquinone and the reduced FAD, the outcome of the calculations heavily depended on the assumed protonation state of the respective cofactor derivative (Fig. 2). Although very similar results compared to the native enzyme were obtained for the neutral compounds **4** and **6**, the anionic species **5** and **7** were predicted to drastically destabilize the cyanide ion indicated by a large positive pK_a shift of more than 7 units (Table 2). Figure 4 shows the electrostatic potential calculated with the program GRASP (Nicholls 1993) mapped onto the surface of enzyme-bound mandelonitrile in the presence of FAD species **3**, **5**, or **7**. These more qualitative pictures clearly illustrate the strong positive potential in the active site of the native enzyme and the transition to a significantly negative potential (especially at the position of the cyano group) in the presence of the anionic compounds **5** and **7**.

The influences of charged residues located within 14 Å of the active site on the pK_a of HCN were investigated, and the results are shown in the lower part of Table 2. For this purpose the partial atomic charges of the respective side chains were modified to yield a neutral species, while possible structural changes were not considered. Some of these charged residues are predicted to have a significant influence on the pK_a of enzyme-bound HCN. Removing the positive charge of amino acid residues Arg-194, Arg-300, or

Fig. 3. Stereo views of the active site of PaHNL in the modeled complexes with mandelonitrile **1** (A), and benzaldehyde **2** plus CN^- (B). C-, N-, O-, and S-atoms are depicted as gray, blue, red, and yellow spheres, respectively. Potential hydrogen-bonding interactions are represented by green dotted lines. The bound substrates are shown in magenta, the FAD cofactor in light blue. This figure (as well as Fig. 4) was prepared using Molscript (Kraulis 1991) and Raster3D (Merritt and Bacon 1997).

Lys-361 destabilizes CN^- relative to HCN, and results in pKa shifts of 2–3 units, while for the other residues intermediate (Arg-182) or negligible (Arg-116 and Asp-96) effects are obtained. Removing the negative charge of Asp-297 is predicted to result in a considerable stabilization of the cyanide ion (Table 2). In all cases, the estimated pKa shift is less positive (or more negative) when benzaldehyde is bound to the active site.

Discussion

The currently known examples of hydroxynitrile lyases (HNLs) suggest that enzymes with cyanogenic activity are the result of convergent evolution originating from at least

five unrelated ancestors (Wajant and Effenberger 1996). It appears that—given the relative simplicity of the reaction—the development of cyanohydrin cleavage activity is an “easy” task, and it is an intriguing question as to whether common mechanistic motifs have developed for these enzymes. To date, only the HNLs from *Hevea brasiliensis* and *Manihot esculenta*, which both show close homology to α/β -hydrolases, have been thoroughly characterized each with respect to their three-dimensional structure and their catalytic mechanism (Wagner et al. 1996; Gruber et al. 1999; Zuegg et al. 1999; Gruber 2001; Lauble et al. 2001a, 2001b).

Another completely unrelated class of HNLs comprises enzymes that require FAD as a cofactor. Despite a wealth of

Table 2. Calculated pK_a -shifts (relative to an aqueous solution) of PaHNL bound HCN

	ΔpK_a (–benzaldehyde)	ΔpK_a (+benzaldehyde)
Cofactor ^a		
Oxidized (3)	0.0	–2.0
Semiquinone, neutral (4)	–0.2	–2.1
Semiquinone, anion (5)	7.4	6.9
Reduced, neutral (6)	0.6	–1.1
Reduced, anion (7)	6.8	6.3
“Mutated” residue ^b		
Arg-300	3.7	1.8
Lys-361	3.2	2.0
Arg-194	1.9	0.2
Arg-182	1.2	–0.5
Arg-116	0.0	–1.8
Asp-96	–0.4	–2.4
Asp-297	–1.9	–4.2

^a Partial atomic charges calculated for the different cofactor derivatives were mapped onto the observed structure of oxidized FAD bound to PaHNL (Dreveny et al. 2001).

^b “Mutation” involved the modification of the partial atomic charges of the side chain atoms to yield a neutral residue.

biochemical data on these FAD-dependent HNLs (Jorns 1979; Yemm and Poulton 1986) the exact location of the active site and the role of the cofactor are still a matter of debate. Based on the structural similarity of one member of this class (the enzyme from almond, PaHNL) to glucose oxidase (Hecht et al. 1993; Wohlfahrt et al. 1999) and cholesterol oxidase (Vrieling et al. 1991; Yue et al. 1999), the active site was located near the flavin cofactor (Dreveny et al. 2001). This tentative assignment is also in line with the observed enzyme inactivation upon formation of an FAD–sulfite adduct (Jorns 1979) and with the occurrence of three mercury binding sites (mercury ions are also known to inhibit the HNL reaction; Yemm and Poulton 1986) in this region (Dreveny et al. 2001).

In general, structural information on enzyme substrate complexes is a prerequisite for an understanding of the catalytic mechanism on a molecular level. In the absence of such data for PaHNL, systematic and unbiased docking simulations provided independent evidence regarding the location of the active site and—for the first time—identified the amino acid residues that are most likely involved in catalysis. Although this approach did not rely on any prior knowledge, it yielded very clear results regarding the location of the active site (Table 1, Fig. 3) and corroborated previous hypotheses.

Inhibition studies of PaHNL had indicated that the active site of the enzyme contains a hydrophobic pocket and a positively charged residue (Jorns 1980). Very early, it had also been recognized that cyanogenic activity in general requires a base to deprotonate the cyanohydrin hydroxyl group and a positive charge to stabilize the emerging cy-

nide ion (Becker and Pfeil 1966). In PaHNL, His-497 most likely acts as the base aided by hydrogen bonds from the side chains of Tyr-457 and Cys-328 (Fig. 3A). In the reverse reaction, the same residues supposedly donate hydrogen bonds to the carbonyl oxygen atom (Fig. 3B), and thereby activate benzaldehyde for nucleophilic attack. Although His-497 and Tyr-457 are strictly conserved among FAD-dependent HNLs, the position of Cys-328 is occupied by valine and isoleucine in some cases (Dreveny et al. 2001), indicating that this cysteine residue might not be essential for catalysis.

In glucose oxidase, a histidine residue equivalent to His-497 was also suggested as the base required in the catalytic mechanism of this enzyme (Hecht et al. 1993; Meyer et al. 1998). Although the imidazole group of this residue is hydrogen bonded to a glutamate side chain in glucose oxidase, His-497 in PaHNL appears not to be particularly activated by the hydrogen bond to the side chain of Ser-496, which is itself interacting with a water molecule. The analysis of the pH dependence of the enzymatic reaction yielded a pK_a value of around 6 (Becker and Pfeil 1966). The same value was also obtained from the pH dependence of thiocyanate binding to the enzyme (Jorns 1979). Both observations are at least consistent with a more or less unperturbed histidine residue present in the active site. In the proposed mechanism of cholesterol oxidase, another histidine residue (corresponding to His-459 in PaHNL) was suggested to act as a general base mediated by a water molecule (Li et al. 1993). An equivalent water molecule was also observed at this position in the crystal structure of PaHNL (Dreveny et al. 2001). In the modeled complexes (Fig. 3), His-459 itself cannot directly deprotonate the hydroxyl group of the substrate, and the position of the water molecule is occupied by the cyano group. However, because all water molecules were removed before the docking simulations, the involvement of a His-459/water pair as a general base in the catalytic cycle cannot be excluded based on these calculations alone. On the other hand, the interaction of the histidine with the side chain of Lys-361 (Fig. 3) renders this option rather unlikely.

No positively charged residues are in direct contact with the docked substrates (Fig. 3). Instead, the second presumed requirement for HNL activity—the stabilization of the negatively charged cyanide ion (Becker and Pfeil 1966)—appears to be met by a strong positive electrostatic potential at the substrate binding site of PaHNL (Fig. 4A). This is also evident from the estimated pK_a shifts of enzyme bound HCN (Table 2). In the native enzyme, with oxidized FAD as a cofactor, electrostatic interactions counterbalance the desolvation energy and yield a virtually unshifted pK_a value. Almost all charged residues in the neighborhood of the active site are conserved among FAD-dependent HNLs. A notable exception is Lys-361, which is mutated to a glutamine in one HNL isoenzyme from each *Prunus amygdalus*

Fig. 4. Electrostatic potential mapped onto the molecular surface of PaHNL bound mandelonitrile **1** in the presence of oxidized FAD **3** (A) and the anionic forms of the semiquinone **5** (B) and the reduced FAD **7** (C). The electrostatic potential is represented by a color ramp from blue (positive) to red (negative). The molecular surface and the potential were calculated using the program GRASP (Nicholls 1993).

and *Prunus serotina* (Suelves and Puigdomenech 1998; Dreveny et al. 2001). Together with Arg-300 and Arg-194, this residue contributes most to the electrostatic potential, and removal of its charge would result in a significant destabilization of the cyanide ion (Table 2). In both isoenzymes, the mutation of Lys-361 is accompanied by the replacement of the close-by Thr-493 by glutamate. In the absence of structural information it can only be speculated that His-459 is protonated in these cases favored by an interaction with the newly introduced glutamate residue. The charge on His-459 would still provide a strong positive electrostatic potential in the substrate binding site. Only in the *Prunus amygdalus* isoenzyme the negatively charged Asp-297 is also replaced by an asparagine (Suelves and Puigdomenech 1998), which has a profound influence on the pKa of HCN (Table 2), outweighing the effect of removing the positive charge of Lys-361.

Although both degradation and formation of cyanohydrins does not include a redox step, it has been inferred that the bound FAD is somehow involved in catalysis, because reduction of the cofactor causes enzyme inactivation (Bärwald and Jaenicke 1978; Jaenicke and Preun 1984). Our calculations yielded no significant differences in the pKa value of HCN when oxidized FAD was replaced by the neutral forms of the semiquinone or the reduced FAD (Table 2). However, spectroscopic data indicated that the red semiquinone anion **5** is formed instead of the blue neutral semiquinone **4** upon photochemical reduction of the PaHNL bound cofactor (Massey and Palmer 1966). In the presence of the anion **5** the cyanide ion is then drastically destabilized, and the pKa of HCN is shifted to much higher values. Reduced FAD is also most likely bound as the anion **7** to PaHNL. This assumption is based on NMR measurements (Sanner et al. 1991) and modeling studies (Meyer et al. 1998) on the structurally homologous enzyme glucose oxidase and on the fact that especially the FAD environment is very similar in the two proteins (Dreveny et al. 2001). The direction and magnitude of the pKa shift in the presence of compound **7** is comparable to the one seen for the semiquinone anion **5** (Table 2). The docking and pKa calculations

suggest that FAD is not directly involved in catalysis, and that the observed enzyme inactivation upon cofactor reduction is largely the result of the reversal of the electrostatic potential within the active site (Fig. 4). It has to be noted, however, that possible concomitant conformational changes of the protein were not considered. Additional effects on the enzymatic activity due to such structural changes cannot be excluded at the moment, and are even necessary for an explanation of the reduced activity of PaHNL with the neutral 5-deaza flavin as a cofactor (Bärwald and Jaenicke 1978; Jorns 1979).

We propose an enzymatic mechanism for PaHNL (Fig. 5), in which His-497 acts as a general acid/base deprotonating the hydroxyl group of the substrate. In addition, the emerging negative charge of the cyano group is stabilized by the positive electrostatic potential at the substrate binding site. In a second step, His-497 supposedly protonates the cyanide ion to yield HCN. In the direction of cyanohydrin synthesis, HCN is deprotonated by His-497, the resulting cyanide ion then attacks the carbonyl carbon of the aldehyde or ketone, and the carbonyl oxygen is protonated by His-497. In this sense, His-497 may be regarded as a proton shuttle between the hydroxy group/carbonyl oxygen and HCN/CN⁻. Enzyme kinetics data are only consistent with an ordered uni-bi mechanism for the PaHNL reaction (Jorns 1980). In such a mechanism, the carbonyl compound binds first followed by HCN, whereas, in the degradation direction, HCN would leave the active site first. Our pKa calculations showed that the cyanide ion is more stabilized in the presence of benzaldehyde (Table 2), which is most likely due to a favorable interaction with the carbonyl carbon atom. These results correlate well with the kinetic data, and indicate that, in the synthesis direction, the presence of the carbonyl compound is necessary for facilitating the deprotonation of HCN and for increasing its nucleophilicity.

The proposed mechanism bears strong resemblance to the one suggested for the FAD-independent HNL from *Hevea brasiliensis* (Zuegg et al. 1999; Gruber 2001). In the latter enzyme a catalytic triad consisting of a serine, a histidine, and an aspartate side chain is supposed to act as the general

Fig. 5. Proposed catalytic mechanism of PaHNL.

acid/base. Stabilization of the cyanide ion is accomplished by the positive charge of an active site lysine that is in direct contact with the cyano group of the substrate (Gruber 2001). A related mechanism was also proposed for the highly homologous HNL from *Manihot esculenta* (Lauble et al. 2001a, 2001b), but without recognizing the essential role of the positive charge of the lysine residue (Gruber 2001). Although the actual amino acid residues involved in catalysis are different in the two HNL classes, a common motif for the mechanism of cyanogenesis starts to emerge that had been predicted already about 40 years ago (Becker and Pfeil 1966).

Materials and methods

Docking

The model of the substrate mandelonitrile **1** (Fig. 1) was built and optimized using the program Sybyl v6.5 (Tripos Inc.). Partial atomic charges for **1** were calculated using the RESP protocol (Cieplak, et al. 1995). Parameters for the oxidized FAD cofactor were generously provided by G. Wohlfahrt (Wohlfahrt et al. 1999). Protein coordinates were taken from the X-ray crystal structure of PaHNL, determined at 1.5-Å resolution (PDB-code: 1JU2) (Dreveny et al. 2001). Because the two independent copies of the enzyme in the asymmetric unit were observed to be structurally very similar (with a root-mean-square deviation of 0.5 Å for all C α -atoms), only one of them was used in the subsequent calculations. Although all crystallographic water molecules were discarded, the FAD cofactor and the first sugar moiety (N-acetylglucosamine) at all four identified glycosylation sites were kept. Asp, Glu, Arg, and Lys residues were treated as charged; protonation and tautomerization states of His residues were chosen, which resulted in sensible hydrogen bonding networks. Hydrogen atoms were added to the structure, followed by a geometry optimization using AMBER 6.0 (Case et al. 1997), applying harmonic restraints on the positions of all heavy atoms. Only polar hydrogen atoms of the protein and the ligands were retained for the docking simulations.

The coordinates of PaHNL were transformed in such a way that the macromolecule was centered at the origin and had its principal axes aligned with the coordinate axes. Forty-five orthogonal boxes (each containing 61 $\times$ 59 $\times$ 65 grid points with a spacing of 0.42 Å) were then positioned on a regular 5 $\times$ 3 $\times$ 3 grid within the bounding box of the protein (78.5 $\times$ 51.5 $\times$ 50.5 Å, including a 3-Å cushion). The number of boxes and their dimensions were chosen to achieve a considerable overlap: the total volume of the small boxes was about 360% of the volume of the bounding box. In each of these boxes, the *R*-enantiomer of **1** was docked to the enzyme using the programs AutoDock v2.4 (Morris et al. 1996) and v3.0 (Morris et al. 1998). In both cases, the protein was kept rigid, and the position and orientation of **1** as well as two torsion angles were allowed to vary during the simulations.

In the calculations with AutoDock v2.4 low-energy binding modes were identified from 100 randomly generated initial structures that were each subjected to 75 cycles of Monte Carlo simulated annealing. The initial value for *RT* and the cooling factor were set to 1000 and 0.9, respectively. In each cycle a maximum of 1000 accepted or rejected moves were allowed. In the calculations with AutoDock v3.0 a hybrid genetic algorithm with phenotypic local search (designated as a Lamarckian genetic algorithm) (Morris et al. 1998) was applied. Ten independent simulations with populations consisting of 50 random structures were performed. The number of energy evaluations was limited to 250,000, resulting in about 90–100 generations on average. The best individual of each generation automatically survived. The probability for performing a local search consisting of up to 300 iterations of a pseudo Solis&Wets optimization (Solis and Wets 1981) was 10%.

In all cases, the low-energy structures found in each “docking box” were clustered with an RMS-tolerance of 2.5 Å. These structures were collected from all boxes, and were again clustered using the same cutoff. In addition, the top 10 binding sites of **1** as obtained in the simulations with AutoDock v3.0 were subjected to a restrained minimization using the program AMBER 6.0 (Case et al. 1997). The Cornell et al. all-atom force field (Cornell et al. 1995) augmented by parameters for cyanohydrins (Gruber 2001) was used, and a distance dependent dielectric function, an 8-Å cutoff for nonbonded interactions as well as harmonic positional restraints for all protein atoms were applied. The molecular mechanics interaction energies of **1** with the protein were computed using the Carnal module of AMBER.

Estimation of pKa values of HCN

Electrostatic potentials at the active site of PaHNL were calculated using the finite-difference Poisson-Boltzmann (FDPB) method as implemented in the program DELPHI (Nicholls and Honig 1991). The two values for the bulk dielectric constant were set to 4.0 (protein) and 80.0 (solvent), and an ionic strength of 0.145 M was assumed throughout. The scale of the grid was $2 \text{ \AA}^{-1}$. The ion exclusion radius was set to $2 \text{ \AA}$ and the probe radius (for surface calculations) to $1.4 \text{ \AA}$. Partial atomic charges for oxidized FAD (**3** in Fig. 2), as well as for the neutral and anionic forms of the semiquinone (**4** and **5**) and the reduced FAD (**6** and **7**) were obtained from RESP calculations (Cieplak et al. 1995) on the respective derivatives of 7,8,10-trimethylisalloxazine (lumiflavin). The protonation states of amino acid residues were the same as used in the docking simulations.

The calculations of the pKa shifts of PaHNL bound HCN included contributions of electrostatic interactions as well as desolvation energies (Yang et al. 1993). The reported ΔpK_a values are averages over calculations assuming two different positions of HCN/CN⁻ that were deduced (a) from the docked mandelonitrile **1** (Fig. 3A), and (b) from the modeled complex of PaHNL with benzaldehyde and CN⁻ (Fig. 3B). The latter structure was built based on the docking results with **1**, and its geometry was optimized by a restrained energy minimization (with only the substrate molecules being mobile) using AMBER 6.0. To assess the influence of the redox state of the cofactor, the calculations were performed with all five FAD derivatives (Fig. 2) bound to the enzyme. For this purpose, only the varying partial atomic charges were mapped onto the crystallographically observed cofactor (Dreveny et al. 2001), while structural changes that possibly occur upon reduction were not considered. Additional calculations were carried out on single-point "charge mutations" in which all charged residues within a radius of $14 \text{ \AA}$ around the substrate binding site were neutralized in turn. All pKa calculations were performed with and without benzaldehyde bound in the active site. The coordinates for the aldehyde were taken from the modeled complex mentioned above (Fig. 3B).

Acknowledgments

This work was supported by the Österreichische Akademie der Wissenschaften (APART-fellowship 614) and by the Fonds zur Förderung der wissenschaftlichen Forschung (SFB Biokatalyse). We thank the members of the Sepzialforschungsbereich Biokatalyse, specifically Herfried Griengl and Helmut Schwab, for their continued interest, encouragement, and support.

The publication costs of this article were defrayed in part by payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 USC section 1734 solely to indicate this fact.

References

- Bärwald, K.R. and Jaenicke, L. 1978. D-Hydroxynitrile lyase: Involvement of the prosthetic flavin adenine dinucleotide in enzyme activity. *FEBS Lett.* **90**: 255–260.
- Becker, W. and Pfeil, E. 1966. Über das Flavinenzym D-Oxynitrilase. *Biochem. Z.* **346**: 301–321.
- Case, D.A., Pearlman, D.A., Caldwell, J.W., Cheatham III, T.E., Ross, W.S., Simmerling, C.L., Darden, T.A., Merz, K.M., Stanton, R.V., Cheng, A.L., Vincent, J.J., Crowley, M., Ferguson, D.M., Radmer, R.J., Seibel, G.L., Singh, U.C., Weiner, P.K., and Kollman, P.A. 1997. *AMBER 6.0*. University of California, San Francisco, CA.
- Cheng, I.P. and Poulton, J.E. 1993. Cloning of cDNA of *Prunus serotina* (R)-(+)-mandelonitrile lyase and identification of a putative FAD-binding site. *Plant Cell Physiol.* **34**: 1139–1143.
- Cieplak, P., Cornell, W.D., Bayly, C.I., and Kollman, P.A. 1995. Application of the multimolecule and multiconformational RESP methodology to biopolymers—Charge derivation for DNA, RNA, and proteins. *J. Comput. Chem.* **16**: 1357–1377.
- Conn, E.E. 1981. Secondary plant products. In *The biochemistry of plants: A comprehensive treatise* (eds. P.K. Stumpf and E.E. Conn), vol. 7, pp. 479–500. Academic Press, New York.
- Cornell, W.D., Cieplak, P., Bayly, C.I., Gould, I.R., Merz, K.M., Ferguson, D.M., Spellmeyer, D.C., Fox, T., Caldwell, J.W., and Kollman, P.A. 1995. A 2nd generation force-field for the simulation of proteins, nucleic-acids, and organic molecules. *J. Am. Chem. Soc.* **117**: 5179–5197.
- Dreveny, I., Gruber, K., Glieder, A., Thompson, A., and Kratky, C. 2001. The hydroxynitrile lyase from almond: A lyase that looks like an oxidoreductase. *Structure* **9**: 803–815.
- Effenberger, F. 1994. Synthesis and reactions of optically-active cyanohydrins. *Angew. Chem. Int. Ed.* **33**: 1555–1564.
- Fessner, W.D. and Walter, C. 1996. Enzymatic C—C bond formation in asymmetric synthesis. *Top. Curr. Chem.* **184**: 97–194.
- Gerstner, E., Mätzke, V., and Pfeil, E. 1968. Zur chemischen und biologischen Systematik der Rosaceen: Untersuchung des Flavinsystems D-Oxynitrilase. *Naturwissenschaften* **55**: 561–563.
- Griengl, H., Hickel, A., Johnson, D.V., Kratky, C., Schmidt, M., and Schwab, H. 1997. Enzymatic cleavage and formation of cyanohydrins: A reaction of biological and synthetic relevance. *J. Chem. Soc. Chem. Commun.* 1933–1940.
- Gruber, K. 2001. Elucidation of the mode of substrate binding to hydroxynitrile lyase from *Hevea brasiliensis*. *Proteins* **44**: 26–31.
- Gruber, K., Gugganig, M., Wagner, U.G., and Kratky, C. 1999. Atomic resolution structure of hydroxynitrile lyase from *Hevea brasiliensis*. *Biol. Chem.* **380**: 993–1000.
- Hecht, H.J., Kalisz, H. M., Hendle, J., Schmid, R. D. and Schomburg, D. 1993. Crystal structure of glucose oxidase from *Aspergillus niger* refined at $2.3 \text{ \AA}$ resolution. *J. Mol. Biol.* **229**: 153–172.
- Hickel, A., Hasslacher, M., and Griengl, H. 1996. Hydroxynitrile lyases: Functions and properties. *Physiol. Plantarum* **98**: 891–898.
- Jaenicke, L. and Preun, J. 1984. Chemical modification of hydroxynitrile lyase by selective reaction of an essential cysteine-SH group with α,β -unsaturated propiophenones as pseudo-substrates. *Eur. J. Biochem.* **138**: 319–325.
- Jorns, M.S. 1979. Mechanism of catalysis by the flavoenzyme oxynitrilase. *J. Biol. Chem.* **254**: 12145–12152.
- Jorns, M.S. 1980. Studies on the kinetics of cyanohydrin synthesis and cleavage by the flavoenzyme oxynitrilase. *Biochim. Biophys. Acta* **613**: 203–209.
- Kiess, M., Hecht, H.J. and Kalisz, H.M. 1998. Glucose oxidase from *Penicillium amagasakiense*. Primary structure and comparison with other glucose-methanol-choline (GMC) oxidoreductases. *Eur. J. Biochem.* **252**: 90–99.
- Kraulis, P.J. 1991. MOLSCRIPT: A program to produce both detailed and schematic plots of protein structures. *J. Appl. Crystallogr.* **24**: 946–950.
- Lauble, H., Förster, S., Miehlisch, B., Wajant, H., and Effenberger, F. 2001a. Structure of hydroxynitrile lyase from *Manihot esculenta* in complex with substrates acetone and chloroacetone: Implications for the mechanism of cyanogenesis. *Acta Crystallogr.* **D57**: 194–200.
- Lauble, H., Miehlisch, B., Förster, S., Wajant, H., and Effenberger, F. 2001b. Mechanistic aspects of cyanogenesis from active-site mutant Ser80Ala of hydroxynitrile lyase from *Manihot esculenta* in complex with acetone cyanohydrin. *Protein Sci.* **10**: 1015–1022.
- Li, J., Vrielink, A., Brick, P., and Blow, D.M. 1993. Crystal structure of cholesterol oxidase complexed with a steroid substrate: Implications for flavin adenine dinucleotide dependent alcohol oxidases. *Biochemistry* **32**: 11507–11515.
- Massey, V. and Palmer, G. 1966. On the existence of spectrally distinct classes of flavoprotein semiquinones. A new method for the quantitative production of flavoprotein semiquinones. *Biochemistry* **5**: 3181–3189.
- Massey, V., Muller, F., Feldberg, R., Schuman, M., Sullivan, P.A., Howell, L.G., Mayhew, S.G., Matthews, R.G., and Foust, G.P. 1969. The reactivity of flavoproteins with sulfite. Possible relevance to the problem of oxygen reactivity. *J. Biol. Chem.* **244**: 3999–4006.
- Merritt, E.A. and Bacon, D.J. 1997. Raster3D: Photorealistic molecular graphics. *Methods Enzymol.* **277**: 505–524.
- Meyer, M., Wohlfahrt, G., Knablein, J., and Schomburg, D. 1998. Aspects of the mechanism of catalysis of glucose oxidase: A docking, molecular mechanics and quantum chemical study. *J. Comput. Aided Mol. Des.* **12**: 425–440.
- Morris, G.M., Goodsell, D.S., Halliday, R.S., Huey, R., Hart, W.E., Belew, R.K., and Olson, A.J. 1998. Automated docking using a Lamarckian genetic

- algorithm and an empirical binding free-energy function. *J. Comput. Chem.* **19**: 1639–1662.
- Morris, G.M., Goodsell, D.S., Huey, R., and Olson, A.J. 1996. Distributed automated docking of flexible ligands to proteins: Parallel applications of AutoDock 2.4. *J. Comput. Aided Mol. Des.* **10**: 293–304.
- Nahrstedt, A. 1985. Cyanogenic compounds as protecting agents for organisms. *Pl. Syst. Evol.* **150**: 35–47.
- Nicholls, A.J. 1993. *GRASP: Graphical representation and analysis of surface properties*. Columbia University, New York.
- Nicholls, A. and Honig, B. 1991. A rapid finite difference algorithm utilizing successive over-relaxation to solve the Poisson-Boltzmann equation. *J. Comp. Chem.* **12**: 435–445.
- Rosenthaler, L. 1908. Durch Enzyme bewirkte asymmetrische Synthesen. *Biochem. Z.* **14**: 238–253.
- Sanner, C., Macheroux, P., Ruterjans, H., Muller, F., and Bacher, A. 1991. ¹⁵N- and ¹³C-NMR investigations of glucose oxidase from *Aspergillus niger*. *Eur. J. Biochem.* **196**: 663–672.
- Solis, F.J. and Wets, R.J.B. 1981. Minimization by random search techniques. *Math. Oper. Res.* **6**: 19–30.
- Suelves, M. and Puigdomenech, P. 1998. Molecular cloning of the cDNA coding for the (R)-(+)-mandelonitrile lyase of *Prunus amygdalus*: Temporal and spatial expression patterns in flowers and mature seeds. *Planta* **206**: 388–393.
- Vrielink, A., Lloyd, L.F., and Blow, D.M. 1991. Crystal structure of cholesterol oxidase from *Brevibacterium sterolicum* refined at 1.8 Å resolution. *J. Mol. Biol.* **219**: 533–554.
- Wagner, U.G., Hasslacher, M., Griengl, H., Schwab, H., and Kratky, C. 1996. Mechanism of cyanogenesis: The crystal structure of hydroxynitrile lyase from *Hevea brasiliensis*. *Structure* **4**: 811–822.
- Wajant, H. and Effenberger, F. 1996. Hydroxynitrile lyases of higher plants. *Biol. Chem.* **377**: 611–617.
- Wohlfahrt, G., Witt, S., Hendle, J., Schomburg, D., Kalisz, H.M., and Hecht, H.J. 1999. 1.8 and 1.9 Å resolution structures of the *Penicillium amagasakiense* and *Aspergillus niger* glucose oxidases as a basis for modeling substrate complexes. *Acta Crystallogr.* **D55**: 969–977.
- Yang, A.-S., Gunner, M.R., Sharp, K., and Honig, B. 1993. On the calculation of pKa's in proteins. *Proteins* **15**: 252–265.
- Yemm, R.S. and Poulton, J.E. 1986. Isolation and characterization of multiple forms of mandelonitrile lyase from mature black cherry (*Prunus serotina* Ehrh.) seeds. *Arch. Biochem. Biophys.* **247**: 440–445.
- Yue, Q.K., Kass, I.J., Sampson, N.S., and Vrielink, A. 1999. Crystal structure determination of cholesterol oxidase from *Streptomyces* and structural characterization of key active site mutants. *Biochemistry* **38**: 4277–4286.
- Zuegg, J., Gruber, K., Gugganig, M., Wagner, U.G., and Kratky, C. 1999. Three-dimensional structures of enzyme-substrate complexes of the hydroxynitrile lyase from *Hevea brasiliensis*. *Protein Sci.* **8**: 1990–2000.