Atomic Resolution Crystal Structures and Quantum Chemistry Meet to Reveal Subtleties of Hydroxynitrile Lyase Catalysis*

Andrea Schmidt†, Karl Gruber‡, Christoph Kratky§, and Victor S. Lamzin‡

From the †EMBL Hamburg c/o Deutsches Elektronen Synchrotron, Notkestrasse 85, D-22607 Hamburg, Germany and §Institut fur Chemie, Karl-Franzens Universita¨t Graz, Humboldtstrasse 50/III, A-8010 Graz, Austria

Hydroxynitrile lyases are versatile enzymes that enantiospecifically cope with cyanohydrins, important intermediates in the production of various agrochemicals or pharmaceuticals. We determined four atomic resolution crystal structures of hydroxynitrile lyase from Hevea brasiliensis: one native and three complexes with acetone, isopropyl alcohol, and thiocyanate. We observed distinct distance changes among the active site residues related to proton shifts upon substrate binding. The combined use of crystallography and ab initio quantum chemical calculations allowed the determination of the protonation states in the enzyme active site. We show that His235 of the catalytic triad must be protonated in order for catalysis to proceed, and we could reproduce the cyanohydrin synthesis in ab initio calculations. We also found evidence for the considerable pKa shifts that had been hypothesized earlier. We envision that this knowledge can be used to enhance the catalytic properties and the stability of the enzyme for industrial production of enantiomerically pure cyanohydrins.

Hydroxynitrile lyases (HNLs),2 EC 4.2.1.39, catalyze the cleavage of cyanohydrins into hydrocyanic acid (HCN) and the corresponding aldehyde or ketone (Fig. 1). In nature, the liberation of HCN serves as a defense mechanism against herbivores and microbial attack in a large variety of plants (1–3) and is initiated by a β-glycosidase-mediated degradation of cyanoglycosides, yielding a sugar and the α-hydroxynitrile (cyanohydrin). HNL is therefore one of the key targets for disruption of this pathway in order to prevent the HCN liberation and thus to detoxify food crops. In vitro, the (reverse) cyanohydrin synthesis reaction is favored (4), and HNLs have already been efficiently used for industrial syntheses of optically pure chiral cyanohydrins (4–7), important intermediates for the production of a wide range of pharmaceuticals and agrochemicals (8). An overview of HNL function and application has been given by Johnson et al. (9) and references therein.

Hydroxynitrile lyase from Hevea brasiliensis (Hb–HNL) is an FAD-independent HNL (2), an unglycosylated protein of the α/β hydrolase fold family with a molecular mass of 29.2 kDa that occurs as a homodimer in neutral aqueous solution (10). In vitro Hb–HNL accepts a variety of aliphatic, aromatic, and heterocyclic aldehydes or ketones for the (S)-specific synthesis of the corresponding α-cyanohydrins (4, 6, 11). Although the three-dimensional structures of the native enzyme (12, 13) and of a number of complexes (14–16) have been known for some time, fine details of substrate interaction, such as the protonation state or subtle active site rearrangement, or structural data on a ternary complex with both substrates have so far been elusive. One problem was the limitation in crystallographic resolution for the complexes. Another problem was posed by the “stickiness” of the active site, which resulted in unwanted complexes at high resolution with histidine (used as a buffer in the first samples (12)) or glycerol (used as a cryoprotectant (13)). Hence, no truly “native” cryotemperature structure to high resolution with an empty active site has been available to which those complexes could be compared.

Here we present the analysis of four atomic resolution structures of Hb–HNL: native and in complex with acetone, isopropyl alcohol (synthesis substrate and substrate analog), and thiocyanate (inhibitor). We primarily focus on the subtleties of ligand binding revealed from the contact patterns in the active site and from complementary ab initio modeling of the protonation state on those highly accurate protein models. From the structures, we deduce the subtleties of substrate recognition and protein-substrate interaction as well as the importance of the structural flexibility of this enzyme. Analysis of the thiocyanate structure provides information on the protonation state and the electron distribution in the binding site. In combination with the structures of the acetone and isopropyl alcohol complexes, we could model the relative positioning of the two substrates in the protein. The combination of accurate structural data and ab initio quantum chemical calculations aided in understanding the mechanism of active site adaptation, the requirements for catalysis, and the influence of pH.

EXPERIMENTAL PROCEDURES

Sample Preparation and Crystallization—Hb–HNL was obtained from the company DSM and purified as described elsewhere (17). A solution of Hb–HNL with a concentration of 12 mg/ml in 50 mM potassium phosphate buffer at pH 7.5 was
Atomic Resolution Structures of Hydroxynitrile Lyase

FIGURE 1. HNL catalysis. Shown is the chemical reaction catalyzed by HNL. The degradation occurs in nature, and the synthesis reaction is used in industry.

used for crystallization by hanging drop vapor diffusion. Crystallization drops were set up in Linbro plates on siliconized glass coverslides. The wells were sealed with immersion oil. The reservoir solution contained 2% polyethylene glycol 400 and 2 M ammonium sulfate in 100 mM Hepes-Na buffer at pH 7.5. Drops of 4 μl of protein solution mixed with 4 μl of reservoir solution were equilibrated against 1000 μl of reservoir solution at 20 °C. Streak seeding using a cat's whisker was applied ~2 h after setting up the drops. This was essential for producing high quality crystals, which appeared within a few days and rapidly grew to an average size of 350 × 200 × 100 μm³.

The ligands for complexes, acetone, isopropyl alcohol, and thiocyanate (a known inhibitor) were added by soaking. Of the volatiles acetone and isopropyl alcohol, 20 μl were added directly to the reservoir solution. After 10 h of equilibration, another 50 μl were added to the reservoir for overnight equilibration. For the thiocyanate soak, a 500 mM solution of NaSCN was prepared in reservoir solution. 1 μl of this solution was then added directly to the drop containing native crystals. These crystals were left to incubate for 24 h.

In contrast to previous methods (18) and in order to avoid the formation of unwanted complexes with glycerol, the crystals were briefly transferred into a cryoprotectant of 15–20% polyethylene glycol 200 in reservoir solution prior to data collection and shock-frozen to 100 K directly on the beamline goniostat in the nitrogen stream.

Data Collection and Processing—Crystallographic data were collected on the EMBL Hamburg beamlines equipped with a single axis goniostat and Marresearch 165-mm CCD and 345-mm image plate detectors. The data collection was carried out in two passes to a high resolution cut-off of 0.9 and 2 Å. For the low resolution passes, the exposure time was adjusted to avoid overloads. The final, merged data sets were 100% complete and extended to 1.25 Å for the complex with acetone and to 1.05 Å for the other three structures. Data were processed using the HKL suite (19). Conversion of the diffraction intensities to structure factor amplitudes and other necessary file conversions were carried out using the programs of the CCP4 suite (20). The statistics are listed in Table 1.

Structure Refinement—The known crystal structure of Hb-HNL (Protein Data Bank code 1QJ4 (13)) was first rigid body-refined against the measured data using REFMAC (21). This was followed by cycles of isotropic refinement with REFMAC against structure factor amplitudes and automatic solvent building using ARP/WARP (22). Corrections to the model and the addition of alternate conformers were carried out manually using the graphics program O (23). Ligand molecules were inserted after the remaining solvent building had converged. Further refinement was performed with SHELXL (24) against diffraction intensities. At this stage, hydrogen atoms at their calculated positions were included. Finally, the models were refined with individual anisotropic atomic displacement parameters. This procedure was applied to all structures and resulted in final R-values between 0.10 and 0.13 (Table 1). The coordinate error estimates were calculated using SFCHECK (25). The coordinates and diffraction data have been deposited in the Protein Data Bank with accession codes 3C6X (NAT), 3C6Y (ACET), 3C6Z (ISOP), and 3C70 (SCN).

Quantum Chemical (QC) Calculations—Ab initio calculations were done using the GAMESS (26) software. A 6−31G** basis set and a Hückel model for the construction of the initial orbitals was used in Hartree-Fock type calculations. In order to minimize the number of atoms, the amino acids involved in substrate contact were modeled as follows: His235 as imidazole, Asp207 as acetic acid, Ser80 as methanol, and Lys236 as aminoethane. Their protonation states were varied and are listed in Fig. 2. The coordinates for these residues and the ligands acetone and thiocyanate/isothiocyanic acid were taken from the refined crystal structures. In order to mimic the predetermined placement of these residues in a larger molecule, some of the coordinates were fixed in space. This was done by constraining those distances (one per residue) that were not interfering with the contacts of interest. In case of multiple conformers, the one with the higher occupancy (corresponding to the ligand-bound
state by analysis of contacts) was used. Model runs were set up using a combination of the structures with thiocyanate and acetone, in an attempt to mimic the substrate interaction. For these calculations, the sulfur atom was removed from thiocyanate to model the natural substrate cyanide (hydrocyanic acid). Acetone was placed into the thiocyanate structure from an overlay of the two crystal structures. Various protonation states were tried for all models (Fig. 2).

RESULTS

We determined four structures of Hb-HNL to atomic resolution: native (NAT) and complexes with acetone (ACET), isopropyl alcohol (ISOP), and thiocyanate (SCN) (Table 1). In all cases, the electron density was of very high quality, permitting the assignment of atom types, unambiguous determination of side chain orientations, and estimation of occupancies for residues with static disorder. In the structures determined to 1.05 Å, numerous hydrogen atoms were visible even in the first maps; however, they showed mostly on the main chain and aliphatic side chains. A few residues showed two distinct main chain conformers. Side chain disorder occurred primarily in surface lysines and glutamates, whereas only very few other residues showed distinct multiple conformers. Among the latter were Ser<sup>80</sup> in the ACET and SCN structures and His<sup>235</sup> in ISOP, both residues of the catalytic triad. In addition to the ligands in the active site in the structures ACET, ISOP, and SCN, several sulfate ions from the crystallization buffer were found in all structures on the protein surface and in intermolecular contact regions. In ACET, NAT, and ISOP, additional residual density was found that was attributed to a polyethylene glycol molecule from the crystallization buffer. Also, Cys<sup>94</sup> appeared to be modified by attachment of β-mercaptoethanol in those structures.

Overall, the four structures superimpose almost perfectly with r.m.s. deviations that are of the same order of magnitude as the coordinate variations (Tables 1 and 2), but the arrangement of the active site residues shows distinct and significant shifts upon ligand binding. Most pronounced are the changes in distances between Ser<sup>80</sup> Oγ and Cys<sup>81</sup> Sγ, Ser<sup>80</sup> Oγ and His<sup>235</sup> Ne2, and Lys<sup>236</sup> Nε and His<sup>235</sup> Ne2 (Table 3). Interestingly, in the NAT, ISOP, and ACET structures, the distance between Ser<sup>80</sup> Oγ and His<sup>235</sup> Ne2 is too long for a hydrogen bond, in contrast to other catalytic triads commonly found in a variety of hydrolases (Table 3). The active site lies in a deep cavity that is only accessible through a very narrow tunnel formed by the residues of the so-called cap domain (Fig. 3) and thus is well protected from the solvent, as was also reported earlier (12). It features a catalytic triad consisting of Ser<sup>80</sup>, His<sup>235</sup>, and Asp<sup>207</sup>. These residues are always very well defined in the electron density in single conformation, with the two exceptions mentioned above. Further residues within contact range to the triad are Lys<sup>236</sup>, Thr<sup>11</sup>, and Cys<sup>81</sup>. In the native structure, the active site contains a small cluster of very well defined water molecules. All except one of these waters are replaced by the ligands in the other structures. This remaining water is located in a pocket just underneath the catalytic His<sup>235</sup> (Fig. 4).

Acetone and isopropyl alcohol bind in a very similar manner. They are held in place by Ser<sup>80</sup> and the catalytic triad and nearby Thr<sup>11</sup> through hydrogen bonding contacts to the oxygen atom of the ketone or alcohol functional groups (Figs. 4 and 5). In the ISOP structure, His<sup>235</sup> occurs in two conformations, although the bound isopropyl alcohol molecule is fully occupied.

In the ACET and SCN structures, Ser<sup>80</sup> appears in two conformations. In ACET, this seems to correlate with the total occupancy of 0.80 for the acetone ligand, which itself appears in two conformations with occupancies of 0.50 and 0.30 (adding up to 0.80). We think that the conformer of Ser<sup>80</sup> with the

### Table 1

Data collection and refinement statistics for the HNL structures

| Structure short name | Native | ISOP | ACET | SCN |
|----------------------|--------|------|------|-----|
| Resolution range (Å) | 35–1.05| 30–1.05| 15–1.25| 35–1.05|
| Completeness (%)     | 145,064| 140,831| 83,123| 145,202|
| Redundancy           | 4.1    | 3.9  | 3.5  | 6.2 |
| No. of atoms (total) | 2,857  | 2,865| 2,789| 2,775|
| No. of heteroatoms   | 609    | 610  | 535  | 612 |
| Overall B protein (Å<sup>2</sup>) | 12 | 11 | 13 | 12 |
| DPI                   | 0.032  | 0.032 | 0.054 | 0.032 |
| Estimated minimum/maximum coordinate error (Å) | 0.003/0.016 | 0.004/0.016 | 0.010/0.025 | 0.003/0.015 |

### Table 2

r.m.s. deviation values (Å) for the superposition of the protein structures (all protein atoms)

|        | NAT | ISOP | ACET | SCN |
|--------|-----|------|------|-----|
| r.m.s. deviation 1–2 distances | 0.027 | 0.016 | 0.012 | 0.015 |
| r.m.s. deviation 1–3 distances | 0.034 | 0.030 | 0.027 | 0.030 |

- **a** Outermost shell.
- **b** Diffraction data precision indicator.
Atomic Resolution Structures of Hydroxynitrile Lyase

TABLE 3
Contact distances (Å) and their variations (r.m.s. deviation from mean) in the HNL active site
The largest changes occurred in the contacts between residues of the catalytic triad.

|          | Asp<sup>207</sup>-His<sup>235</sup>, O<sub>61</sub>-Ne2 | His<sup>235</sup>-Ser<sup>80</sup>, Ne2-O<sub>y</sub> | Ser<sup>80</sup>-Cys<sup>81</sup>, O<sub>y</sub>-Sy | Cys<sup>81</sup>-Thr<sup>11</sup>, Sy-O<sub>y</sub> | Glu<sup>17</sup>-Lys<sup>236</sup>, Oe2-N<sub>η</sub> | Lys<sup>236</sup>-His<sup>235</sup>, N<sub>η</sub>-Ne2 |
|----------|---------------------------------------------|---------------------------------------------|---------------------------------------------|---------------------------------------------|---------------------------------------------|---------------------------------------------|
| NAT      | 2.70                                        | 3.77                                        | 4.08                                        | 3.81                                        | 2.71                                        | 4.67                                        |
| ISOP     | 2.78                                        | 3.23                                        | 4.50                                        | 3.77                                        | 2.74                                        | 4.50                                        |
| ACET     | 2.66                                        | 3.73                                        | 4.15                                        | 3.81                                        | 2.72                                        | 4.56                                        |
| SCN      | 2.70                                        | 3.00                                        | 4.60                                        | 3.92                                        | 2.73                                        | 4.74                                        |
| r.m.s. deviation | 0.05                                        | 0.38                                        | 0.26                                        | 0.06                                        | 0.01                                        | 0.11                                        |

occupancy of 0.80 relates to a “loaded” state, since it points to the acetone molecule and makes a well defined hydrogen bonding contact. The conformer with the occupancy of 0.20 would then correspond to an “open” state and indicates elevated flexibility of that residue. The two conformers of acetone are rotated by a small amount with respect to each other around an axis perpendicular to the molecule plane but leaving the oxygen atom position the same.

Rhodanide binds deeper in the active site pocket and makes contacts to both Lys<sup>236</sup> and His<sup>235</sup> via its nitrogen end. The nitrogen atom occupies a position taken by a water molecule in the NAT and ACET structures. The sulfur atom occupies the position taken by the central carbon atoms of acetone or isopropyl alcohol in the ACET and ISOP structures (Fig. 4). The electron density for thiocyanate clearly shows the very short bond of 1.16 Å between the carbon and nitrogen atom, whereas the sulfur atom is well separated in the map at a distance of 1.74 Å from the carbon. This indicates that of the two possible iso-

carbon atom is well separated in the map at a distance of 1.74 Å from the carbon. This indicates that of the two possible iso-

carbon atom is well separated in the map at a distance of 1.74 Å from the carbon. This indicates that of the two possible iso-

carbon atom is well separated in the map at a distance of 1.74 Å from the carbon. This indicates that of the two possible iso-
state of the active site favors catalysis. This setup required Ser^{80} to be protonated (QC2, QC5, and QC6 in Table 6), since this is the only polar contact the acetone molecule makes in the protein apart from the second hydrogen bond to Thr^{11} (Fig. 2 and Table 6). Settings with a protonated His^{235} (QC2 and QC5 in Table 6) seem favorable and result in a structure that contains a cyanohydrin molecule closely linked to Ser^{80} via a hydrogen bridge. Thus, the acetone molecule, which in the overlay of structures lay with its plane roughly perpendicular to the SCN ligand, now flips over to form a tetrahedral structure with a strong covalent bond to the CN moiety (Fig. 8 and QC5 in Table 6). On the other hand, in a setting with SCN-H and a deprotonated His^{235} (QC6 in Table 6), the original ligands HCN and acetone are retained, and the cyanohydrin does not form during the course of the calculations, indicating a pivotal role of the correct protonation state of His^{235}. These results show that the ligands in the SCN and ACET structures occupy the positions required for catalysis and that the interactions with the few active site residues included in the calculations are sufficient to trigger the chemical reaction, provided they have the correct protonation states.

**DISCUSSION**

An overlay of the substrate binding sites shows that the ligands in the SCN and ACET structures are positioned almost perfectly as required for catalysis, both relative to each other and with respect to the active site residues. The similarity of the ACET and ISOP structures is intriguing, since the two ligands have different stereochemistry and hydrogen bonding properties: acetone is planar and can accept two hydrogen bonds on its oxygen atom, whereas isopropyl alcohol is puckered and can accept two and donate one hydrogen bond on its OH group. The orientation of acetone and isopropyl alcohol is very similar, but neither shows ideal hydrogen bonding geometry. This already hints at some flexibility and the capability of the binding site to accommodate the slightly different shape of their ligands (flat or puckered) as well as both types of oxygen atoms (carbonyl and alcohol). The QC results show that the few hydrogen bonding contacts are sufficiently strong to allow for the chemical reaction to occur and support a reaction mechanism suggested by Gruber et al. (13) (see also Ref. 9). In this mechanism, Ser^{80} and Thr^{11} provide key contacts and trigger the nucleophilic attack of CN by inducing electron withdrawal from the acetone carbonyl carbon atom. The lack of other specific contacts explains why a wide variety of ketone or aldehyde compounds are accepted by Hb-HNL, which makes this enzyme so versatile and particularly suitable for industrial applications.

The position of the SCN ligand differs only little from the previously reported location in a medium resolution structure (14). The sulfur atom of SCN takes the space of the central carbon atom of acetone or isopropyl alcohol in the ACET and ISOP structures (Fig. 4), which structurally explains why SCN is a good competitive inhibitor for HNL. The SCN sulfur atom is located at a distance of more than 4.6 Å from the S^2/H9253 of Cys^{81}, and this prohibits close interaction with that residue, despite the fact that it had earlier been identified by site-directed mutagenesis as a key residue for catalysis (28). In previously proposed mechanisms, Cys^{81} was thought to abstract the proton from the hydrocyanic acid bound to the active site (12, 15). Although we cannot determine whether the ligand(s) interact...
with Cys81 during entry into the active site, the distance between the Cys81 $\gamma$ and the SCN C in the bound state is 6.0 Å, too long for any interaction. The only nearby residues are His235 and Ser80. His235 makes a strong hydrogen bonding contact to the N atom of SCN, and the Ser80 O is at a distance of 3.13 Å from the SCN C atom. These contacts confirm the proposed role for Ser80, which is to activate the ketone compound (15), and for His235, which “activates” the cyanide ligand and is responsible for its reprotonation in the cyanohydrin cleavage reaction (Tables 4–6).

The visual appearance of SCN in the electron density and the observed bond lengths strongly suggest a triple C=N bond. Isothiocyanic acid SCN-H is the isoform that is usually formed upon protonation (the proton resides on N). Thus, we mimicked this protonation state in QC for both cyanide compounds, in order to distinguish between what could be attributed to chemical difference between CN and SCN (which includes a substantial difference in $pK_a$) or a truly remarkable capability of the enzyme active site to adapt to different substrates. From QC, it appears that SCN binds deprotonated with a proton shared with His235 at the N-end (QC1 in Table 4). The bonding distances exactly reproduce the crystal structure, and bond orders for the C–N and S–C bonds are 2.6 and 1.2, respectively, which concurs with the appearance in the electron density. Interestingly, the bonding distances (1.16 and 1.74, respectively) differ somewhat from the average values (1.15 and 1.64 for C–N and C–S, respectively), as extracted from an analysis of SCN compounds in the Cambridge Structural Database (result not shown), and the molecule appears slightly bent. This indicates considerable involvement of hydrogen bonding contacts to the sulfur atom (by Ser80 and Thr11), which elongate the C–S bond.

The N atom of the SCN ligand is involved in a tight hydrogen bond to His235 and a weaker contact to Lys236, as also suggested in the mechanism proposed by Gruber et al. (13). Given this and the limitations of space in the active site, we assume that the locations of the C and N atoms of SCN are exactly those needed for catalysis. Modeling both ligands, CN and acetone, into the active site gives a distance between C of CN and the carbonyl C of acetone of 1.55 Å, almost precisely what it should be for a covalent carbon–carbon bond. The CN orientation is almost perpendicular to the acetone plane, ideal for a nucleophilic attack on the carbonyl carbon.

The protonation state of the cyanide ligand determines the roles of the surrounding residues as proton donors, nucleophiles, or hydrogen bonding partners. From our results, it appears that the CN ligand is bound in its anionic state; however, the protonation state of the ligand during entry into the
TABLE 4
Results from the ab initio calculations with CN as a ligand and compared with the SCN crystal structure
The numbers are bond distances (Å) as well as bond orders (italic type, in parentheses). r.m.s. deviations are calculated for the distances between nonhydrogen atoms listed in the table.

| SCN  | QC1 | QC2 | QC3 | QC4 | QC5 | QC6 | Crystal structure |
|------|-----|-----|-----|-----|-----|-----|-------------------|
| S_cn-C_cn | 1.737 (1.16) | 1.651 (1.31) | 1.669 (1.22) | 1.692 (1.30) | 1.548 (1.97) | 1.604 (1.59) | 1.74 |
| C_cn-N_cn  | 1.159 (2.56) | 1.158 (2.49) | 1.250 (1.69) | 1.264 (1.64) | 1.209 (1.76) | 1.161 (2.18) | 1.16 |
| N_cn-H_cn  | NA  | NA  | NA  | NA  | NA  | NA  | NA  |
| N_cn-H2    | 2.070 (0.08) | 2.088 NA | 1.025 (0.86) | 2.896 NA | 2.402 NA | 1.041 (0.63) | 0.90/2.07 |
| His235     | N235-H2  | 0.913 (0.79) | 0.999 (0.79) | 2.262 NA | 1.006 (0.87) | 0.994 (0.84) | 1.691 (0.12) | 0.90/2.07 |
| Lys236     | H236-N236 | 2.241 NA | 2.110 (0.07) | 2.199 NA | 1.756 (0.15) | 2.160 NA | 2.862 NA | 2.24 |
| Ser230     | H230-O230 | NA  | 0.942 (0.87) | NA  | NA  | 0.944 (0.86) | 0.944 (0.86) | 0.82 |
| Asp207     | O207-C207 | 1.807 (0.17) | 1.948 NA | 2.059 NA | 2.123 NA | 0.968 (0.74) | 1.445 (0.22) | 1.80 |
| Overall r.m.s. deviation to crystal structure |
| Nonhydrogen atoms | 0.012 | Match | Small shifts | Large shifts | Large shifts | Small shifts | Large shifts | NA |
| Comment | NA  | NA  | NA  | NA  | NA  | NA  | NA  | NA  |

* Proton-shifted between His and SCN (i.e. H235 = H2); for the H-SCN His-H+ cases, additional proton at 0.90 Å at appropriate dihedral angle.
* NA, included but not available.

TABLE 5
Results from the ab initio calculations with CN as a ligand and compared with the SCN crystal structure
The numbers are bond distances (Å) as well as bond orders (italic type, in parentheses). r.m.s. deviations are calculated for the distances between nonhydrogen atoms listed in the table.

| CN  | QC1 | QC2 | QC3 | QC4 | QC5 | QC6 | Crystal structure |
|-----|-----|-----|-----|-----|-----|-----|-------------------|
| C_cn-N_cn | 1.161 (2.60) | NA  | NA  | 1.177 (2.41) | 1.280 (1.57) | 1.178 (2.26) | 1.16 |
| N_cn-H_cn  | NP  | NP  | NP  | 1.268 (0.30) | 1.029 (0.88) | 1.142 (0.47) | 1.14 |
| N_cn-H2    | NA  | 2.106 (0.08) | NA  | 2.170 (NA) | 2.073 (0.12) | 1.013 (0.72) | 0.90/2.07 |
| His235     | N235-H2  | 1.005 (0.78) | NA  | 1.002 (0.78) | 1.014 (0.74) | 2.093 (0.07) | 0.90/2.07 |
| Lys236     | H236-N236 | 3.110 (NA) | NA  | 3.135 (NA) | 3.084 (NA) | 3.105 (NA) | 2.97 |
| Ser230     | H230-O230 | 0.949 (0.812) | NA  | 0.956 0.56 | 0.784 0.87 | 0.940 NA | 0.82 |
| Asp207     | O207-C207 | 1.807 (0.17) | 1.948 NA | 2.059 NA | 2.123 NA | 0.968 (0.74) | 1.445 (0.22) | 1.80 |
| Overall r.m.s. deviation to crystal structure |
| Nonhydrogen atoms | NA  | Disintegrates | 0.115 | Small shifts | NA  | Disintegrates | 0.105 | Small shifts | 0.136 | Small shifts | 0.092 | Small shifts | NA  |
| Comment | NA  | NA  | NA  | NA  | NA  | NA  | NA  | NA  |

* NP, not present in this setup.
* NA, included but not available.

active site is still not yet clear. Given the charge distribution in the active site, it seems likely that the substrates enter in their neutral state (H-CN rather than CN−). Indeed, hydrocyanic acid has a pK_a of 9.4, and thus its protonated form is strongly dominating, especially at the pH optimum of ~5 for HNL. We thus think that H-CN enters protonated and is then stripped off of its proton during binding.

The calculations indicate that the scenario with Ser230 deprotonated, His235 protonated, and SCN deprotonated (isothiocyanic acid has a pK_a of ~1.3) is the most favorable case for the SCN structure with an excellent match to the crystal structure (QC1 in Table 4) determined at pH 7.5. For the setup with Ser230 deprotonated, HCN, and a protonated His235 (QC2 in Table 5), the proton shifts from the HCN molecule onto the Ser230 (QC4 in Table 5), which is indicated by the distances and bond orders. For the CN models in the QC calculations, the picture is generally not very clear (Table 5), but we can conclude that for cyanide binding either the incoming ligand or Ser230 must carry a proton. Given the electronic configuration in the SCN ligand, the negative charge would be located at the sulfur atom and in part be compensated by the hydrogen bonding contacts to Thr11 but, on the other hand, also through the contact to the
Atomic Resolution Structures of Hydroxynitrile Lyase

TABLE 6
Results from the ab initio calculations with CN and acetone as ligands compared with the SCN crystal structure

The numbers are bond distances (Å) or angles (degrees) as well as bond orders (italic type, in parentheses). r.m.s. deviations are calculated for the distances between nonhydrogen atoms listed in the table.

|       | QC1       | QC2       | QC3       | QC4       | QC5       | QC6       | Crystal structure |
|-------|-----------|-----------|-----------|-----------|-----------|-----------|------------------|
| CN    |           |           |           |           |           |           |                  |
| C_{cn}^{-}\text{N}_{cn} | X^a       | X         | X         | 1.240 (1.88) | 1.158 (2.42) | 1.16        |
| N_{cn}^{-}\text{H}_{cn} | X         | NP^b      | X         | 1.026 (0.86) | NP         | 0.90        |
| N_{cn}^{-}\text{H}_{cn} | X         |           |           | 2.237 (NA) | 1.025 (0.64) | 0.90/2.07^c  |
| His^{235} |           |           |           |           |           |           |                  |
| N_{ct}^{-}\text{H}_{ct} | X         | 0.999 (0.81) | X         | 0.996 (0.83) | 1.786 (0.14) | 0.90/2.07^c  |
| N_{ct}^{-}\text{N}_{ct} | X         | 3.162     | X         | 3.226     | 2.808     | 2.97       |
| Lys^{236} |           |           |           |           |           |           |                  |
| H_{ct}^{-}\text{N}_{ct} | X         | 1.971 (0.68) | X         | 2.330 (NA) | 2.237 (NA) | 2.24        |
| H_{ct}^{-}\text{N}_{ct} | X         | 1.018 (0.76) | X         | 1.010 (0.81) | 1.014 (0.79) | 0.89        |
| Ser^{80} |           |           |           |           |           |           |                  |
| H_{ct}^{-}\text{O}_{ct} | X         | 0.981 (0.68) | X         | 0.946 (0.81) | 0.951 (0.78) | 0.82        |
| O_{ct}^{-}\text{C}_{ct} | X         | 3.233     | X         | 3.170     | 2.475     | 3.15        |
| Asp^{207} |           |           |           |           |           |           |                  |
| O_{ct}^{-}\text{H}_{ct} | X         | 1.480 (0.24) | X         | 1.409 (0.29) | 1.682 (0.19) | 1.80        |
| O_{ct}^{-}\text{N}_{ct} | X         | 2.566     | X         | 2.526     | 2.717     | 2.70        |
| Acetone |           |           |           |           |           |           |                  |
| C_{acet}^{-}\text{C}_{acet} | X         | 1.561 (0.76) | X         | 1.487 (0.92) | 2.475 (NA) | 1.55        |
| C_{acet}^{-}\text{O}_{acet} | X         | 1.331 (1.12) | X         | 1.407 (0.81) | 1.206 (1.68) | 1.23        |
| O_{acet}^{-}\text{H}_{acet} | X         | 1.629 (0.17) | X         | 1.982 (NA) | 1.840 (0.07) | NP         |
| O_{acet}^{-}\text{O}_{acet} | X         | 2.608     | X         | 2.909     | 2.788     | 2.65        |
| C_{acet}^{-}\text{C}_{acet} | X         | 105.8^d   | X         | 109.5^d   | 90.3^d    | 77.3^e      |
| C_{acet}^{-}\text{C}_{acet} | X         | 110.3^d   | X         | 94.4^d    | 87.3^f    |            |
| Overall r.m.s. deviation to crystal structure | X         | 0.139     | X         | 0.149     | 0.434     | NP         |
| Nonhydrogen atoms | Omitted | Small shifts | Omitted | Small shifts | Large shifts | NP |

^a X, setting omitted.
^b NP, not present in this setup.
^c NA, included but not available.
^d Proton-shifted between His and SCN (i.e., H_{cn} = H_{h2}); for the H-SCN His-H^+ cases, additional proton at 0.90 Å at appropriate dihedral angle.

The observation of the protonated His^{235} in SCN (Fig. 7) concurs with earlier findings that showed that the pK_a of His^{235} was indeed elevated to ~8 in a complex with rhodanide. This means that upon binding of the cyanide compound, His^{235} would turn into a powerful agent for the deprotonation of HCN. The pK_a of His^{235} is ~4 in the free enzyme (15, 28, 29), and indeed this residue appears deprotonated in the NAT structure. This indicates that the incoming ligand(s) severely affect the pK_a of the catalytic residues and, together with the apparent high mobility of the protons in the active site observed from the QC results, supports the idea that the reaction mechanism functions via a protonation-deprotonation sequence (15).

QC indicated the formation of a short, strong hydrogen bond between His^{235} and Asp^{207} with distances of 2.57 and 2.53 Å (2.70 Å in the crystal structure) for the parent atoms and bond orders of 0.24 and 0.29 in settings QC2 and QC5 (Table 6), respectively. Such a strong, short hydrogen bond has previously been both calculated (15) and experimentally monitored (29). In our models, it is formed only in those settings where cyanohydride is synthesized (QC2 and QC5 in Table 6), confirming its crucial role.

His^{235}/Lys^{236} pair. Since the sulfur atom is missing in the CN model, this may explain why a different protonation scheme with a protonated serine is favored for this model (QC2 in Table 5), which leaves His^{235} as the deprotonating base.

The contact distances around Ser^{80} are among the most variable ones in the active site (Table 3), and this residue shows two conformers in two of the structures. Thus, we propose that movement of Ser^{80} may be a crucial factor in the substrate binding mechanism of the enzyme. This is also supported by the fact that Ser^{80}, like other catalytic serines in hydrolases, consistently appears in the disallowed regions of the Ramachandran plot, also in the Hb-HNL structures determined earlier, indicating considerable steric strain in that residue.
The protonated state of His\textsuperscript{235} seems essential for the formation of cyanohydrin (Table 6, QC with protonated His\textsuperscript{235}, QC2 and QC5) results in the cyanohydrin complex, regardless of whether the ligand CN or HCN is present. When His\textsuperscript{235} is modeled deprotonated, the original compounds are retained (QC6 in Table 6). This emphasizes the importance of the correct protonation state of His\textsuperscript{235} and shows its capability to adapt upon ligand binding (in NAT it is deprotonated). From an analysis of the bonding distances and angles in the imidazol ring, we can divide the four structures into two groups: NAT and ISOP with deprotonated His\textsuperscript{235}, SCN and ACET with protonated His\textsuperscript{235}. Considering the change of the angle around Ne2 by more than 2° in the SCN structure (Fig. 7), it seems likely that the SCN ligand elevates the pK\textsubscript{a} for His\textsuperscript{235} even higher than the reported value of 8 (15). This also supports the function of His\textsuperscript{235} as the deprotonating base for the cyanide compound. It seems that the true substrate analog acetone affects the pK\textsubscript{a} of this histidine, making the active site ready for catalysis, whereas isopropyl alcohol has no such effect.

Lys\textsuperscript{236} was suggested earlier as an important contact for the incoming H-CN ligand and as a driving force for deprotonation (15). The negative charge on the SCN or CN ligand would be compensated by Lys\textsuperscript{236}, which always forms a weak but distinct contact in our QC calculations. Additionally, Lys\textsuperscript{236} forms a salt bridge to Glu\textsuperscript{79}, and this charge compensation can be envisioned to enable protonation of the neighboring His\textsuperscript{235}, since two residues carrying positive charges at close distance would not be energetically favorable. We can envision that the active site would await the first ligand HCN in a state where Ser\textsuperscript{80} is deprotonated. Upon binding of HCN, its proton is then abstracted and shifted to Ser\textsuperscript{80}, turning the first ligand into a powerful nucleophile that forms a strong hydrogen bond to His\textsuperscript{235} and a further contact to Lys\textsuperscript{236}, which compensates the charge.

A previously discussed reaction mechanism (12) suggested that when the acetone molecule binds, its interaction prepares the active site for the subsequent binding of cyanide. However, the cyanide binding site lies deeper in the active site cavity. It would therefore seem obvious that cyanide would prearrange the active site for binding of the ketone (or aldehyde) component (i.e. so that Ser\textsuperscript{80} and His\textsuperscript{235} carry their protons; Tables 4 and 5). Existing kinetic data, however, belie this and show that hydrocyanic acid does not inhibit cyanohydrin binding (30) and can be made. Although at the present time we cannot resolve this discrepancy, in either case, the enzyme must display sufficient flexibility in order to admit the two ligands. This is already indicated by higher mobility of tryptophan 128 as shown by higher ADPs and distinct anisotropy. The Trp\textsuperscript{128} side chain forms a “gate” to the tunnel, and investigation of the anisotropic atomic displacement parameters on that residue shows that there is a tendency for rotation of the side chain around the C\textsubscript{2}–C\textsubscript{8} bond. Such movement would widen the entrance to the tunnel and could permit access for the small HCN compound. Furthermore, the adjacent Lys\textsuperscript{239} is tightly linked to the side chain of Glu\textsuperscript{158}, which is part of a short helix whose residues Leu\textsuperscript{153} and Leu\textsuperscript{157} form the hydrophobic wall of the active site cavity. It can be hypothesized that movement of the “gate” would be relayed to the active site. Further investigation of the molecular flexibility should therefore prove interesting in the future, and a study on this topic is under way.

A known feature of Hb-HNL is its ability to accommodate large compounds (8), which has already been exploited for the industrial production of insecticides. However, the active site cavity is rather small (Fig. 3), and therefore conformational adaptations additional to the movement of Trp\textsuperscript{128} must occur for the admission of the substrate and also for the release of the products. Structurally, the region around Trp\textsuperscript{128} shows already a strained conformation, as indicated by Lys\textsuperscript{239}, which, like Ser\textsuperscript{80}, is consistently located in the disallowed regions of the Ramachandran plot. Thus, movement is likely to occur. Knowing about the destabilization of the enzyme at low pH (~4), it can be hypothesized that the low pH optimum is to some extent caused by increased flexibility that Hb-HNL may already have under these conditions. The ab initio calculations and the distance changes in the active site (Fig. 5 and Tables 4 – 6) suggest that adaptations in the structures are likely to occur, and further investigation of the internal motions of this molecule will be the subject of future research.

REFERENCES

1. Conn, E. E. (1981) The Biochemistry of Plants: A Comprehensive Treatise (Stumpf, P. K., and Conn, E. E., eds) Academic Press, London
2. Hickel, A., Hasslacher, M., and Griengl, H. (1996) Physiol. Plant. 98, 891–898
3. Wajant, H., and Effenberger, F. (1996) Biol. Chem. 377, 611–617
4. Klemptier, N., Griengl, H., and Hayn, M. (1993) Tetrahedron Lett. 34, 4769–4772
5. Kruse, C. G. (1992) Chirality in Industry, John Wiley, Chichester, UK
6. Johnson, D. V., and Griengl, H. (1997) Tetrahedron 53, 617–624
7. Johnson, D. V., and Griengl, H. (1998) Adv. Biochem. Eng. Biotechnol. 63, 31–55
8. Effenberger, F. (1994) Angew. Chemie Int. Ed. Engl. 33, 1555–1564
9. Johnson, D. V., Zabelinskaja-Mackova, A. A., and Griengl, H. (2000) Curr. Opin. Chem. Biol. 4, 103–109
10. Hanefeld, U., Stranzl, G., Straathof, A. J., Heijnen, J. J., Bergmann, A., Mittelbach, R., Glatter, O., and Kratky, C. (2001) Biochim. Biophys. Acta 1544, 133–142
11. Klemptier, N., Pichler, U., and Griengl, H. (1995) Tetrahedron: Asymmetry 6, 845–848
12. Wagner, U. G., Hasslacher, M., Griengl, H., Schwab, H., and Kratky, C. (1996) Structure 4, 811–822
13. Gruber, K., Gugganig, M., Wagner, U. G., and Kratky, C. (1999) Biol. Chem. 380, 993–1000
14. Zuegg, J., Gruber, K., Gugganig, M., Wagner, U. G., and Kratky, C. (1999) Protein Sci. 8, 1990–2000
15. Kratky, C., Gartler, G., Krammer, B., Schwab, H., and Kratky, C. (2004) J. Biol. Chem. 279, 20501–20510
16. Gartler, G., Kratky, C., and Gruber, K. (2007) J. Biotechnol. 129, 87–97
17. Hasslacher, M., Schall, M., Hayn, M., Griengl, H., Kohlwein, S. D., and Schwab, H. (1996) J. Biol. Chem. 271, 5884–5891
18. Wagner, U. G., Schall, M., Hasslacher, M., Hayn, M., Griengl, H., Schwab, H., and Kratky, C. (1996) Acta Crystallogr. Sect. D 52, 591–593
19. Orwinowski, Z., and Minor, W. (1997) Methods Enzymol. 276, 307–326
20. CCP4 (1994) Acta Crystallogr. Sect. D 50, 760–763
21. Murshudov, G. N., Vagin, A. A., and Dodson, E. J. (1997) Acta Crystallogr. Sect. D 53, 240–255
22. Lamzin, V. S., and Wilson, K. S. (1997) Methods Enzymol. 277, 269–305
23. Jones, T. A., Zou, J. Y., Cowan, S., and Kjeldgaard, M. (1991) Acta Crystallogr. Sect. A 47, 110–119
24. Sheldrick, G. M. (1997) SHELXL-97, University of Gottingen, Gottingen Germany
25. Vaguine, A. A., Richelle, J., and Wodak, S. J. (1999) Acta Crystallogr. Sect. D 55, 191–205
26. Schmidt, M. W., Baldridge, K. K., Boatz, J. A., Elbert, S. T., Gordon, M. S., Jensen, J. H., Koseki, S., Matsunaga, N., Nguyen, K. A., Su, S., Windus, T. L., Dupuis, M., and Montgomery, J. A., Jr. (1993) J. Comput. Chem. 14, 1347–1363
27. Engh, R. A., and Huber, R. (1991) Acta Crystallogr. Sect. A 47, 392–400
28. Hasslacher, M., Kratky, C., Griengl, H., Schwab, H., and Kohlwein, S. D. (1997) Proteins Struct. Funct. Genet. 27, 438–449
29. Stranzi, G. R., Gruber, K., Steinkellner, G., Zangger, K., Schwab, H., and Kratky, C. (2004) J. Biol. Chem. 279, 3699–3707
30. Bauer, M., Griengl, H., and Steiner, W. (1999) Biotechnol. Bioeng. 62, 20–29
31. Kraulis, P. J. (1991) J. Appl. Crystallogr. 24, 946–950
32. Merritt, E. A., and Murphy, M. E. (1994) Acta Crystallogr. Sect. D 50, 869–873
33. Nicholls, A. J. (1993) GRASP, Columbia University, New York