A Catalytic Mechanism Revealed by the Crystal Structures of the Imidazolonepropionase from Bacillus subtilis*

Received for publication, August 11, 2006, and in revised form, September 12, 2006 Published, JBC Papers in Press, September 21, 2006, DOI 10.1074/jbc.M607703200

Yamei Yu†1, Yu-He Liang†1, Erik Brostrom‡, Jun-Min Quan‡, Santosh Panjikar‡, Yu-Hui Dong‡, and Xiao-Dong Su§1,2, 3

From the †The National Laboratory of Protein Engineering and Plant Genetic Engineering, College of Life Sciences, Peking University, Beijing 100871, China, ‡Laboratory of Chemical Genomics, Shenzhen Graduate School of Peking University, Shenzhen 513055, China, §European Molecular Biology Laboratory Hamburg Outstation, c/o Deutsches Elektronen-Synchrotron, Notkestrasse 85, D-22603 Hamburg, Germany, and ¶Beijing Synchrotron Radiation Facility, Institute of High Energy Physics, Chinese Academy of Sciences, Beijing 100049, China

Imidazolonepropionase (EC 3.5.2.7) catalyzes the third step in the universal histidine degradation pathway, hydrolyzing the carbon-nitrogen bonds in 4-imidazolone-5-propionic acid to yield N-formiminol-γ-glutamic acid. Here we report the crystal structures of the Bacillus subtilis imidazolonepropionase and its complex at 2.0-Å resolution with substrate analog imidazole-4-acetic acid sodium (I4AA). The structure of the native enzyme contains two domains, a TIM (triose-phosphate isomerase) barrel domain with two insertions and a small β-sandwich domain. The TIM barrel domain is quite similar to the members of the barrel metallo-dependent hydrolase superfamily, especially to Escherichia coli cytosine deaminase. A metal ion was found in the central cavity of the TIM barrel and was tightly coordinated to residues His-80, His-82, His-249, Asp-324, and a water molecule. X-ray fluorescence scan analysis confirmed that the bound metal ion was a zinc ion. An acetate ion, 6 Å away from the zinc ion, was also found in the potential active site. In the complex structure with I4AA, a substrate analog, I4AA replaced the acetate ion and contacted with Arg-89, Try-102, Tyr-152, His-185, and Glu-252, further defining and confirming the active site. The detailed structural studies allowed us to propose a zinc-activated nucleophilic attack mechanism for the hydrolysis reaction catalyzed by the enzyme.

The histidine degradation pathway is highly conserved from prokaryotes to eukaryotes (1, 2). In bacteria, the histidine degradation pathway is operated by the histidase or HutH, which was determined by Wu et al. (3) and deposited in the Protein Data Bank under accession number 1XFK. The imidazolonepropionase or HutU, the only enzyme left in this pathway that has no published structural information so far, although some properties of this enzyme have been studied in Salmonella typhimurium (10), Pseudomonas fluorescens ATCC 11299 (11), rat liver (12), and other organisms during 1960s and 1970s. The maximal activity of this enzyme occurred at pH 7.4 with a narrow pH optimum, and its Michaelis constant was calculated to be 7 μM (12). It was also mentioned in the literature that 1 mM EDTA could not affect the enzyme activity, whereas 0.1 mM p-chloromercuribenzoate could inhibit the activity completely, and the inhibition could be reversed 12% by adding 0.3 mM reduced glutathione (12). These observations indicated that the enzyme was not a metalloprotein but a cysteine-dependent enzyme. In the following 20 years, there was no publication on the imidazolonepropionase until 1997, when Holm and Sander suggested that this enzyme could form a TIM

* This work was supported in part by the National High Technology and Development Program of China (863, Program 2002AB711A13) and Grants 985 and 211 from the Peking University. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked “advertisement” in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

The atomic coordinates and structure factors (codes 2B80 and 2G3F) have been deposited in the Protein Data Bank, Research Collaboratory for Structural Bioinformatics, Rutgers University, New Brunswick, NJ (http://www.rcsb.org/).

1 Both authors contributed equally to this work.

2 Recipient of a National Science Fund for Distinguished Young Scholars of National Natural Science Foundation of China (30325012).

3 To whom correspondence should be addressed. Tel.: 86-10-62759743; Fax: 86-10-62765669; E-mail: su-xd@pku.edu.cn.

4 The abbreviations used are: IPA, 4-imidazolone-5-propionic acid; I4AA, imidazole-4-acetic acid; TIM, triose-phosphate isomerase; Se-Met, selenomethionine; CCD, charge-coupled device; r.m.s.d., root mean square deviation.

5 R. Wu, R. Zhang, C. Shonda, and A. Joachimski, manuscript in preparation.
Crystal Structure of B. subtilis Imidazolonepropionase

FIGURE 1. A, the chemical reaction catalyzed by the imidazolonepropionase. B, imidazole-4-acetic acid sodium salt, an analog of 4-imidazolone-5-propionic acid (the substrate of imidazolonepropionase).

barrel and belonged to a novel α/β barrel amidohydrolase superfamily based on the residue-by-residue optimal alignment and superimposition of the three-dimensional structures of urease, phosphotriesterase, and adenosine deaminase (13).

Here we present the crystal structures of an imidazolonepropionase or HutI from B. subtilis, a native enzyme and an enzyme complex with its substrate analog imidazole-4-acetic acid (Fig. 1B), each to 2.0-Å resolution. This work reports the first crystal structures of the imidazolonepropionase family, and a catalytic mechanism has been proposed for this family of enzymes based on the high-resolution structures.

EXPERIMENTAL PROCEDURES

Protein Preparation and Crystallization—Selenomethionine (Se-Met)-labeled imidazolonepropionase was prepared using the method described by Doublie (14), and both wild-type and Se-Met-substituted proteins were purified and crystallized using the same procedure reported previously (15). Sodium salt of a substrate analog, imidazole-4 acetic acid (I4AA), was purchased from Sigma (Chemical Abstract Service (CAS) Number 56368-58-2). Crystals of the enzyme in complex with I4AA were obtained by co-crystallization, adding 50 mM I4AA into the mother liquor of 20% (w/v) polyethylene glycol 4000, 0.1 M Tris-HCl, pH 7.5, 0.2 M sodium acetate, and 2% (w/v) benzamidine hydrochloride.

X-ray Data Collection and Processing—X-ray diffraction data of Se-Met enzyme crystal were collected on a MAR165 charge-coupled device detector at beamline BW7A, European Molecular Biology Laboratory Hamburg outstation around the beamline 911-3, MAX laboratory (Lund, Sweden). The wave-length ranges of the experiments were selected near the K-edge of metal ion binds to the protein, fluorescence scan analysis of the protein solution used for crystallization was performed at beamline 911-3, MAX laboratory (Lund, Sweden). The wavelength ranges of the experiments were selected near the K-edge of Zn²⁺, Mn²⁺, Fe³⁺, and Co²⁺, respectively.

RESULTS AND DISCUSSION

Structure Determination of the Native Enzyme and the Complex—The structure of B. subtilis imidazolonepropionase could be solved by the single wavelength anomalous dispersion method using Se-Met-substituted protein. The crystals formed in the space group P2₁ with β angle nearly equal to 90°, and there were two molecules in the asymmetric unit. The Se-Met crystals diffracted to better than 2.0-Å resolution, and the structure model was finally refined to a crystallographic R-value of 18.8% (Rfree = 22.1%) when data from 20.0 to 2.0-Å resolution of the peak data set were used. The stereochemistry quality of the final model is reasonable as checked by PROCHECK (28); only one residue, His-272, is in the disallowed region of the Ramachandran plot; however, this residue is well defined in the density maps and stabilized by hydrogen bonds with neighboring residues through water bridges. The final model shows res-
idues 3–415 in both monomers. The two monomers are quite similar to each other with an r.m.s.d. (root mean square deviation) of 0.292 Å with 413 Ca atoms aligned. For the complex structure, the final R(_cryst) and R(_free) values of the complex structure model were 0.185 and 0.216, respectively, and the stereochemistry quality of the final model was reasonable. The crystallographic and refinement information of the two structures were summarized in Table 1.

The Overall Structure—The *B. subtilis* imidazolonepropionase consists of 421 amino acids and exists as homodimer (Fig. 2A), consistent with the result of gel filtration experiments (data not shown). The two molecules are closely packed around the barrel far from the central cavity as shown in Fig. 2B, is composed of two separate peptide segments from both N and C termini (residues 3–70 and 365–415). This domain is connected to the TIM barrel domain by two linkers: the first linker is a β-strand (residues 71–74, E), and the second linker is an α-helix (residues 350–364, αE); an extra insertion (residues 389–398) in this domain including αF-helix interacts tightly with the neighbor subunit.

The Substrate Binding Sites and Putative Active Center—Although the previous studies reported that EDTA did not affect the activity of the imidazolonepropionase (12), implicating the enzyme was not a metalloprotein, we did observe a penta-coordinated metal ion in the density map (Fig. 3A), further confirmed by the anomalous difference Fourier map (data not shown). The metal ion is located in the central cavity of the TIM barrel domain and directly coordinated by His-80-N(2), His-82-N(2), His-249-N(2), Asp-324-O(4), and a water molecule (W1 in Fig. 3A) with distances of 2.10 Å, 2.18 Å, 2.28 Å, 2.32 Å, and 2.02 Å, respectively (marked in Fig. 3C). These residues are fully conserved across different organisms (Fig. 2C). The metal-bound water also forms a hydrogen bond with Glu-252 through a water bridge (W2 in Fig. 3A). Later, the x-ray fluorescence scan of the enzyme gave clear absorption jump only at the spectrum

### Table 1

| Parameters                                      | Native enzyme crystal | Enzyme-I4AA complex |
|-------------------------------------------------|-----------------------|---------------------|
| Wavelength                                      | 0.9792                | 1.0                 |
| Space group                                     | P2_1                  | P2_1                |
| Unit cell parameters (Å, °) a = 57.73, b = 106.34, c = 66.47, β = 89.92 | 20.20 (2.03-2.0)     | 20.20 (2.03-2.0) |
| Resolution Range (Å) 20-2.0 (2.03-2.0)         | 9.8 (25.6)            | 6.7 (31.0)          |
| Completeness (%)                                | 99.8 (99.8)           | 98.9 (97.3)         |
| R(_cryst) (%)                                    | 8.3 (3.5)             | 10.7 (2.7)          |
| R(_free) (%)                                     | 0.221                 | 0.216               |
| Average B-factors (Å^2)                         | 0.006                 | 0.005               |
| Deviation from the ideality                     | 1.3                   | 1.3                 |
| Bond lengths (Å)                                | 17.14                 | 16.01               |
| Bond angles (°)                                  | 15.50                 | 14.67               |
| Side chain                                       | 17.88                 | 16.90               |
| Solvent                                         | 23.23                 | 20.54               |
| Substrate analog                                 | 16.3                  | 16.3                |
| Ramachandran plot statistics (%)                | 88.4                  | 88.8                |
| Most favored region                             | 10.9                  | 10.7                |
| Additional allowed region                       | 0.4                   | 0.3                 |
| Generously allowed region                       | 0.4                   | 0.3                 |
| Disallowed region                               | 0.3                   | 0.3                 |

*Values in parentheses refer to the highest resolution shell.

* R(_cryst) = \( \sum |I_O - I_{calc}| / \sum I_{calc} \), where the summation is over all reflections.

* R(_free) = \( \sum |I_O - I_{calc}| / \sum I_{calc} \), where the summation is over all reflections.

* R(_free) is the R-factor for a selected subset of the reflections which are not included in refinement calculations.
over K-edge of Zn$^{2+}$ (data not shown), which unambiguously demonstrated that this metal ion was a zinc ion.

Close to the cavity center of the TIM barrel, an acetate ion was also observed (Fig. 3A); this acetate ion was most likely from the crystallization buffer containing 0.2M sodium acetate. This acetate was quite close to the zinc ion site (about 6 Å away) and formed very nice salt bridges with Arg-89-N$^{\beta1}$ (2.95 Å) and Arg-89-N$^{\beta2}$ (2.29 Å). This scenario of zinc ion plus the acetate ion around the TIM barrel cavity center immediately suggested to us that this place was the active site because the acetate binding site could well be the binding site for the carboxyl group of IPA (see Fig. 1A), the substrate of the enzyme. Furthermore, the zinc ion with the coordinated water was positioned well for catalysis.

To confirm the active site and elucidate the catalytic mechanism of the imidazolonepropionase, crystals of the enzyme in complex with its substrate analog I4AA were made, and their crystal structure was solved. The $F_o - F_c$ density omit map clearly showed that I4AA molecule was bound to the enzyme at the same place where the acetate ion was bound in the native structure (Fig. 3, B and C). The overall structure of the complex is very similar to that of the native enzyme with an r.m.s.d. of 0.13 Å among 413 Ca atoms aligned in subunit A between the two structures. The only notable differences between the two structures are mainly from residues Arg-89, Glu-252, Tyr-102, and His-272 (Fig. 3C) caused by the insertion of I4AA molecule into the active site. In the complex structure, the zinc ion and the coordinated water molecule are located exactly at the same place as in the native enzyme structure. The carboxyl group of I4AA is located very close to the acetate binding site in the native enzyme structure, with about 0.6 Å shift toward the zinc site; so is the side chain of Arg-89, which still forms nice salt bridges with the carboxyl group of I4AA. Besides the salt bridges with Arg-89, I4AA forms hydrogen bonds with Tyr-102, Tyr-152, His-185, and Glu-252; these bonds keep the I4AA molecule in a correct orientation in the central cavity. The imidazole ring of the I4AA is positioned very close to the water molecule coordinated to zinc and replaces the water molecule (see Fig. 3C) bridging zinc-bound water and Glu-252 in the native structure, indicating a possible nucleophilic attack by zinc-activated water in the hydrolysis reaction.

Among the residues involved in zinc and I4AA binding, consequently important to the enzyme function, most of them are located close to the central cavity of the TIM barrel as expected from many other TIM barrel enzymes with known structures and catalytic mechanisms. His-80, His-82, Arg-89, and Tyr-102 are from the Insertion I region; His-185, His-249, and His-272 are located at the C terminus of β-strands 3, 5, and 6, respectively; and Tyr-152, Glu-252, and Asp-324 are located in the

---

**FIGURE 2.** A, the structure of B. subtilis imidazolonepropionase dimer. B, the stereoview of the overall structure of a monomer. The zinc ion is shown in CPK mode in yellow, and the acetate is shown in stick mode in green. The illustrations were made by the program Pymol. C, structural-based sequence alignment of representative HutI (imidazolonepropionase) proteins. The primary sequences of HutI from Gram-negative bacteria P. putida, A. tumefaciens, and S. typhimurium and Gram-positive bacteria B. subtilis, Staphylococcus aureus, and eukaryote X. tropicalis, H. sapiens were aligned. The secondary structure of B. subtilis and A. tumefaciens imidazolonepropionase was placed on the top and the bottom, respectively. The alignment was generated with ClustalX (29) and drawn with ESPript (30). Highly conserved residues were marked with red background (completely identical) and blue boxes (partially identical). The red upward triangles indicate residues interact with the zinc. Black stars highlight the residues forming hydrogen bonds with the substrate analog. D, the topology diagram of the imidazolonepropionase structure generated by TopDraw (37). The secondary structure distribution of the model was given by program DSSP (31).
Crystal Structure of B. subtilis Imidazolonepropionase

FIGURE 2—continued
Crystal Structure of B. subtilis Imidazolonepropionase

**β-α** loops 2, 5, and 8, respectively (Fig. 2C). Insertion I (αA-, αB-, and αC-helices) lies as a lid on the entrance of central cavity (Fig. 2A), making the zinc and substrate binding site deeply buried in the molecule. The tight coordination and deep buried location could be the reason why the zinc ion was hard to dissociate from the enzyme with EDTA. Although Insertion I covers the central cavity, there is a narrow hydrophilic tunnel leading to the barrel core in the structure, which can be the entrance path for the substrate. Because the inserted helices connect to the TIM barrel through long loops, the entire region could be flexible and change to different conformations during hydrolysis reaction.

**Structural and Functional Homologs**—In a search for related homologs of B. subtilis imidazolonepropionase using BLAST (32), it is observed that the enzyme exists widely in various organisms. We picked HutI proteins from representative Gram-negative, Gram-positive bacteria, and eukaryotes, respectively and aligned these sequences with ClustalX and ESPript (Fig. 2C). The highest percentage of identical amino acid residues among the six proteins was 73%, which was between Xenopus tropicalis and Homo sapiens, whereas the lowest was 32% between S. typhimurium and X. tropicalis. It was noticed that the conserved regions were distributed through the whole length of the proteins except at the N and C termini. The residues suggested to be important to the substrate binding and catalytic activity were highly conserved. The most significant difference between the prokaryotic and eukaryotic sequences is that there are two gaps in the prokaryotic sequences. One gap is at the D-strand from the small β-sandwich domain, and the other small gap is at the third β-β loop in the TIM barrel domain. The two gaps are far from the active center and probably do not interfere with the activity of the enzyme.

Structural similarity search by DALI (33) revealed that the catalytic domain (TIM barrel domain) of imidazolonepropionase belonged to the α/β barrel metal-dependent hydrolase superfamily. The structure with the highest similarity to imidazolonepropionase is that of cytosine deaminase from Escherichia coli (34), with an r.m.s.d. of 2.8 Å (358 Ca atoms with 18% sequence identity). Cytosine deaminase is confirmed to be a ferrous-dependent enzyme, whereas the coordination geometry of Fe2+ is quite similar to that of imidazolonepropionase (34, 35). Another enzyme in this superfamily, adenosine deaminase, is discovered to be a zinc-binding metalloprotein; adenosine deaminase has also shown very similar coordination geometry at the metal binding site in the active center (36). Recently, crystal structure of imidazolonepropionase from Agrobacterium tumefaciens was deposited in the Protein Data Bank (Tyagi et al., 6 ID code 2GOK). A. tumefaciens imidazolonepropionase shows 41% of sequence identity to that of B. subtilis, and the similarity between the two structures is high (r.m.s.d. of 1.7 Å among 393 Ca atoms). Significant differences between the two structures are located at the loop connecting β-β sandwich domain; and at the helices of Insertion I region, the average shift of αB-helix is

\[ \text{R. Tyagi, D. Kumaran, and S, Swaminathan, manuscript in preparation.} \]

**FIGURE 3.** Zinc and substrate binding sites of the enzyme. A, the native enzyme structure, B, the complex structure. The key residues and ligands were marked and represented in stick mode (carbon, cyan; nitrogen, blue; oxygen, red; and zinc, yellow). The ΔFo – Fc maps were calculated with zinc, water bound to zinc, acetate and I4AA molecules omitted and drawn at a contour level of 3σ. C, the superimposition of the native enzyme structure (carbon atoms in yellow) and complex structure (carbon atoms in cyan). The distances marked are from the native enzyme structure, indicating the zinc coordination and the salt bridge between acetate ion and the residue Arg-89.
about 10 Å. In the model of the structure of *A. tumefaciens* imidazolonepropionase, a similar metal ion binding site was also observed; however, with a ferric iron put in the binding site.

**Catalytic Mechanism**—Based on the above structural studies and analyses, we can propose a hydrolysis mechanism for the imidazolonepropionase family as outlined in Fig. 4. In the original state in the “empty enzyme” (labeled as Enzyme in Fig. 4), the zinc-bound water forms strong hydrogen bonds with Asp-324 and with Glu-252 through another “bridging” water molecule. The enzyme adopts an open conformation in the Insertion I region, ready for accepting a substrate molecule to enter through the hydrophilic tunnel. In the next step as indicated by the arrowhead in Fig. 4, a substrate molecule IPA is bound into the active site. One of the protons of the zinc-bound water was shifted to Glu-252, and the water bridge between the zinc-bound water and Glu-252 was replaced by the substrate. The IPA molecule was stabilized and correctly oriented by a strong salt bridge with residue Arg-89 and hydrogen bond network with residues Glu-252, Ser-329, and His-185. Additionally, the effect of salt bridge on Arg-89 will induce a “closed confirmation” of Insertion I region and isolate the reaction from the “water solvent” outside. The carbonyl oxygen of the substrate will be activated by the protonated carboxyl group of Glu-252, and then the zinc-bound hydroxide attacks the carbonyl carbon of the substrate to form the tetrahedral intermediate (EI-1, as shown in Fig. 4); the amide nitrogen changes its hybridization from sp2 to sp3 during the nucleophilic attack and protonated by the proton transferred from Glu-252 (EI-2, as shown in Fig. 4), facilitating the cleavage of the N-C bond of the substrate to form the enzyme-product complex (EP, as shown in Fig. 4). Finally the reaction cycle goes back to the original state as in Enzyme with the release of product and the uptake of a new water molecule.

**Conclusions**—In summary, *B. subtilis* imidazolonepropionase represents a widely existed protein family, either viewed from the sequential homology analysis or from structural similarity analysis. The structural study and the x-ray fluorescence scan experiments reported here reveal that the imidazolonepropionase is a zinc enzyme and adopts a zinc-activated nucleophilic attacking mechanism. An insertion region on the entrance of the central cavity might be important to the enzyme activity.

**Acknowledgments**—We thank the staff of the Synchrotron Radiation Sources at Beijing Synchrotron Radiation Facilities, MAX laboratory, and Deutsches Elektronen-Synchrotron at European Molecular Biology Laboratory Hamburg outstation for their help with our data collection. We also thank Dr. Rong Gao and Prof. Yundong Wu for their helpful discussions and critical reading of the manuscript and Lanfen Li for help with some experiments performed for this study.
REFERENCES
1. Revel, H. R., and Magasanik, B. (1958) J. Biol. Chem. 233, 930–935
2. Brown, D. D., Silva, O. L., Gardiner, R. C., and Silverman, M. (1960) J. Biol. Chem. 235, 2058–2062
3. Chasin, L. A., and Magasanik, B. (1968) J. Biol. Chem. 243, 5165–5178
4. Kimhi, Y., and Magasanik, B. (1970) J. Biol. Chem. 245, 3545–3548
5. Oda, M., Sugishita, A., and Furukawa, K. (1988) J. Bacteriol. 170, 3199–3205
6. Yoshida, K., Sano, H., Seki, S., Oda, M., Fujimura, M., and Fujita, Y. (1995) Microbiology (Read.) 141, 337–343
7. Schwede, T. F., Badeker, M., Langer, M., Retey, J., and Schulz, G. E. (1999) Protein Eng. 12, 151–153
8. Schwede, T. F., Retey, J., and Schulz, G. E. (1999) Biochemistry 38, 5355–5361
9. Kessler, D., Retey, J., and Schulz, G. E. (2004) J. Mol. Biol. 342, 183–194
10. Smith, G. R., Halpern, Y. S., and Magasanik, B. (1971) J. Biol. Chem. 246, 3320–3329
11. Rao, D. R., and Greenberg, D. M. (1960) Biochim. Biophys. Acta 43, 404–418
12. Snyder, S. H., Silva, O. L., and Kies, M. W. (1961) J. Biol. Chem. 236, 2996–2998
13. Holm, L., and Sander, C. (1997) Protein Eng. 10, 72–82
14. Doublie, S. (1997) Methods Enzymol. 276, 523–530
15. Yu, Y., Li, L., Zheng, X., Liang, Y. H., and Su, X. D. (2006) Biochim. Biophys. Acta 1764, 153–156
16. Otwinowski, Z., and Minor, W. (1997) Methods Enzymol. 276, 307–326
17. Panjikar, S., Parthasarathy, V., Lamzin, V. S., Weiss, M. S., and Tucker, P. A. (2005) Acta Crystallogr. Sect. D Biol. Crystallogr. 61, 449–457
18. Cottman, K. (1994) Joint CCP4 and ESF-EACBM Newsletter on Protein Crystallography 31, 34–38
19. Sheldrick, G. M., Hauptman, H. A., Weeks, C. M., Miller, R., and Usón, I. (2001) in International Tables for Crystallography (Rossmann, M. G., and Arnold, E., eds) pp. 333–351, IUCr and Kluwer Academic Publishers, Norwell, MA
20. Schneider, T. R., and Sheldrick, G. M. (2002) Acta Crystallogr. Sect. D Biol. Crystallogr. 58, 1772–1779