Molecular structure of leucine aminopeptidase at 2.7-Å resolution
(x-ray crystallography/zinc enzyme/exopeptidase/bestatin)

Stephen K. Burley*†, Peter R. David*, Allen Taylor‡, and William N. Lipscomb*

*Gibbs Chemical Laboratory, Harvard University, Cambridge, MA 02138; †Department of Medicine, Brigham and Women’s Hospital, Boston, MA 02115; and ‡U.S. Department of Agriculture Human Nutrition Research Center on Aging, Tufts University, Boston, MA 02111

Contributed by William N. Lipscomb, June 26, 1990

ABSTRACT The three-dimensional structure of bovine lens leucine aminopeptidase (EC 3.4.11.1) complexed with bestatin, a slow-binding inhibitor, has been solved to 3.0-Å resolution by the multiple isomorphous replacement method with phase combination and density modification. In addition, the structure of the isomorphous native enzyme has been refined at 2.7-Å resolution, and the current crystallographic R factor is 0.169 for a model that includes the two zinc ions and all 487 amino acid residues comprising the asymmetric unit. The enzyme is physiologically active as a hexamer, which has 32 symmetry and is triangular in shape with a triangle edge length of 115 Å and maximal thickness of 90 Å. The monomers are crystallographically equivalent and each is folded into two unequal α/β domains connected by an α-helix to give a comma-like shape with approximate dimensions of 90 × 55 × 55 Å³. The secondary structural composition is 40% α-helix and 19% β-strand. The N-terminal domain (160 amino acids) mediates trimer–trimer interactions and does not appear to participate directly in catalysis. The C-terminal domain (327 amino acids) is responsible for catalysis and binds the two zinc ions, which are 2.88 Å apart. The pair of metal ions is located near the edge of an eight-stranded, saddle-shaped β-sheet. One zinc ion is coordinated by carboxylate oxygen atoms of Asp-255, Asp-332, and Glu-334 and the carbonyl oxygen of Asp-332. The other zinc ion is coordinated by the carboxylate oxygen atoms of Asp-255, Asp-273, and Glu-334. The active site also contains two positively charged residues, Lys-250 and Arg-336. The six active sites are themselves located in the interior of the hexamer, where they line a disk-shaped cavity of radius 15 Å and thickness 10 Å. Access to this cavity is provided by solvent channels that run along the twofold symmetry axes.

The aminopeptidases form a group of exopeptidases that catalyze removal of amino acids from the N terminus of a protein. These enzymes are ubiquitous in nature and are of critical biological and medical importance because of their key role in protein degradation and in the metabolism of biologically active peptides. Whereas the mechanism of action and the three-dimensional structures of carbonpeptidases and endopeptidases are known in significant detail, our understanding of the aminopeptidases is much less well developed.

Leucine aminopeptidases (LAPs; EC 3.4.11.1) are widely distributed cystosolic exopeptidases that catalyze the hydrolysis of amino acids from the N terminus of polypeptide chains (1, 2). As their name implies, the LAPs cleave leucyl substrates, although substantial rates of enzymatic cleavage are seen with most amino-terminal amino acids. Bovine lens LAP is a hexameric enzyme of molecular weight 324,000, which consists of six identical subunits (3) of molecular weight 54,000 and 12 zinc ions (4). The amino acid sequence of the protein has been determined by both protein (5) and DNA (6).

Wallner and A.T., unpublished data) sequencing methods. In addition, a potent, slow-binding inhibitor of LAP, bestatin or [(2S,3R)-3-amino-2-hydroxy-4-phenylbutanoyl]leucine, was isolated from culture filtrates of Streptomyces olivoreticuli (6) and was shown to have a $K_i = 2 \times 10^{-8}$ M for bovine lens LAP (7).

Although there is considerable interest in LAP and in other aminopeptidases, no structural data are available for any aminopeptidase at high resolution. Two electron microscopic studies of bovine lens LAP have shown that the enzyme is a hexamer with 32 symmetry (8, 9). Previous crystallographic studies have been limited to crystallization of bovine lens LAP (10), hog kidney LAP (11), and Escherichia coli methionine aminopeptidase (12) and a description of an electron-density map of bovine lens LAP determined by single isomorphous replacement and solvent flattening at 5-Å resolution (13).

We have solved the three-dimensional structure of the bestatin-inhibited enzyme by using the multiple isomorphous replacement (MIR) method at 3.0-Å resolution with phase combination and density modification and have refined the structure of the isomorphous native enzyme at 2.7-Å resolution. The active site has been identified and the environment of the two zinc ions has been characterized. Here, we report the general structural features of the enzyme; its secondary, tertiary, and quaternary structures; and the structure of the active site. A detailed analysis of the mechanism of inhibition of the enzyme by bestatin and the results of refinement of both the native and bestatin-inhibited enzymes at high resolution will be published elsewhere.

MATERIALS AND METHODS

Purification and Crystallization. LAP from bovine lens was purified according to a modification of the method of Carpenter and coworkers (14) by one of us (A.T.) and was stored in a buffer consisting of 50 mM Tris Cl (pH 7.8) and 50 μM ZnSO₄. The enzyme was inhibited by incubation with its slow-binding inhibitor bestatin at a concentration of 10 μM. Isomorphous crystals of the native or bestatin-inhibited enzyme were grown by one of us (A.T.) using the hanging-drop method, by vapor diffusion against the storage buffer with 1–2 M Li₂SO₄ as the precipitant. Under these conditions, crystals in the form of hexagonal bars with maximum diameter of about 0.25–0.3 mm could be obtained in 1–2 months. They show systematic absences of the hexagonal space group $P\bar{6}_1\bar{2}2\bar{1}$, with $a = 132$ Å and $c = 122$ Å (10). The unit cell contains two hexamers with the asymmetric unit consisting of one promoter of molecular weight 54,000, and the solvent occupies 57% of the crystal volume.

Abbreviations: LAP, leucine aminopeptidase; MIR, multiple isomorphous replacement.

The atomic coordinates and structure factors have been deposited (August 1, 1990) in the Protein Data Bank, Chemistry Department, Brookhaven National Laboratory, Upton, NY 11973 (references 1LAP, R1 LAPSF).

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked ‘advertisement’ in accordance with 18 U.S.C. §1734 solely to indicate this fact.
Table 1. Heavy-atom derivative preparation

<table>
<thead>
<tr>
<th>Compound</th>
<th>Conc.</th>
<th>pH</th>
<th>Time</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ethylmercury phosphate (Hg1)</td>
<td>125 μM</td>
<td>7.8</td>
<td>5 days</td>
</tr>
<tr>
<td>Mercuric chloride (Hg2)</td>
<td>250 μM</td>
<td>7.8</td>
<td>5 days</td>
</tr>
<tr>
<td>Iridium hexachloride (Ir1)</td>
<td>50 mM</td>
<td>7.8</td>
<td>5 days</td>
</tr>
<tr>
<td>Iridium hexachloride (Ir2)</td>
<td>250 mM</td>
<td>7.8</td>
<td>5 days</td>
</tr>
<tr>
<td>Sodium diuranate (DU1)</td>
<td>½ sat.</td>
<td>6.4</td>
<td>8 days</td>
</tr>
<tr>
<td>Sodium diuranate (DU2)</td>
<td>½ sat.</td>
<td>6.8</td>
<td>8 days</td>
</tr>
</tbody>
</table>

sat., Saturation.

Structure Determination. The three-dimensional structure determination was carried out by three of us (S.K.B., P.R.D., and W.N.L.). Heavy-atom derivatives of the inhibited enzyme were prepared by soaking crystals after equilibration of the heavy-atom solutions and the crystals by vapor diffusion (see Table 1 for details of the soaking conditions; the inhibited enzyme was chosen for the MIR study in the hopes of minimizing interactions between the active-site metals and heavy-atom reagents).

Diffraction data for the native enzyme were collected at the Resource for Crystallography located in the laboratory of N. h. Xuong (University of California, San Diego), where two multiwire proportional chambers (15) were used. Diffraction data for the bestatin-inhibited enzyme and six heavy-atom derivatives were measured to 3.0-Å resolution using a Xentronics-type area detector (16) with the Harvard data-collection software (17). Small crystals and the moderately large unit cell required the use of double Franks mirrors (18, 19). Reduction to intensities was performed using the BUDDHA package (17) and reduction to structure-factor amplitudes and local scaling was performed using the algorithm of Fox and Holmes (20).

Table 2 gives the statistics for data collection for the native enzyme, the enzyme-inhibitor complex, and its heavy-atom derivatives.

Isomorphous difference Patterson syntheses revealed a common, major heavy-atom binding site in the two mercurial derivatives, and the atomic position was refined by the method of Terwilliger and Eisenberg (21). Despite repeated measurement attempts, anomalous scattering from mercury and uranium heavy atoms was not detected by either of the two Xentronics-type area detectors available in our laboratory and supplied by Siemens Analytical X-ray. However, anomalous scattering from the ethylmercury phosphate derivative was successfully measured using the FAST detector in the laboratory of J. Pflugrath (Cold Spring Harbor Laboratory) (data not shown). Difference Fourier syntheses confirmed the location of the common, major mercury binding site and allowed identification of two independent minor mercury binding sites, one in each of the two mercury derivatives. In addition, these preliminary MIR phases allowed identification of heavy-atom binding sites in the diuranate- and iridium-soaked crystals. Further improvement in the accuracy of the phase data resulted from combining Wang’s density modification method implemented by Wolf-gang Kabsch (22, 23) with the Hendrickson-Lattman coefficients obtained from the MIR method (24). After density modification, each of the derivatives was reexamined by difference Fourier synthesis to identify yet more minor heavy-atom binding sites (see Table 2 for the results of heavy-atom refinement). The final average figure-of-merit obtained from the MIR method at 3.0-Å resolution was 0.65.

The initial electron-density maps showed some well-defined right-handed α-helices and β-strands, but the majority of the loop regions were uninterpretable. A discontinuous, unrefined polyalanine model consisting of 347 of a possible 487 amino acids was used for the first round of phase combination and density modification (25, 26). Subsequent rounds of phase combination/density modification used refined, discontinuous models obtained from manual rebuilding and chain extension using the molecular graphics package FRODO (27, 28) with an Evans & Sutherland Picture System 390, and positional refinement using XPLOR (29) with a DECstation 3100. The amino acid sequence was aligned by identifying three cysteine residues from the refined positions of the three mercury binding sites. In all, eight rounds of phase combination/model building/partial structure refinement were required to reveal the locations of all 487 amino acids and two zinc ions that comprise the asymmetric unit. This iterative process was carried out at 3.0-Å resolution to give a final R factor of 0.248 and root-mean-square deviations of 0.016 Å and 3.7° for bond lengths and angles, respectively. The active site showed additional electron density adjacent to the zinc ions that is consistent with the structure of the slow-binding inhibitor bestatin, but no attempt was made to establish the precise orientation of the bestatin at this resolution. Recently, x-ray diffraction data to 2.25-Å resolution from bovine lens LAP complexed with bestatin were obtained at the National Synchrotron Light Source in collaboration with R. M. Sweet. Refinement of these data is proceeding and the structure of the bestatin-inhibited enzyme will be published elsewhere.

Refinement of the isomorphous native structure at 2.7-Å resolution was carried out using the structure factors obtained with the Xuong-Hamlin area detector system. The entire model was examined using a series of unrefined delete maps that spanned the polypeptide chain, and minor corrections to side-chain and backbone positions were made. Further model building and positional and temperature-factor refinement reduced the R factor to 0.169, with root-mean-

Table 2. Data collection statistics

<table>
<thead>
<tr>
<th>Crystal</th>
<th>Resolution, Å</th>
<th>No. unique/no. collected</th>
<th>Rsymm*</th>
<th>ΔR †</th>
<th>Rcenric ‡</th>
<th>Phasing power §</th>
</tr>
</thead>
<tbody>
<tr>
<td>Native</td>
<td>20–2.7</td>
<td>16,334/275,554</td>
<td>0.102</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bestatin</td>
<td>20–3.0</td>
<td>11,388/57,147</td>
<td>0.057</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hg1</td>
<td>20–3.0</td>
<td>10,345/63,489</td>
<td>0.076</td>
<td>0.12</td>
<td>0.58</td>
<td>1.35</td>
</tr>
<tr>
<td>Hg2</td>
<td>20–3.0</td>
<td>10,562/43,299</td>
<td>0.068</td>
<td>0.11</td>
<td>0.55</td>
<td>1.45</td>
</tr>
<tr>
<td>Ir1</td>
<td>20–3.0</td>
<td>10,890/41,656</td>
<td>0.060</td>
<td>0.11</td>
<td>0.57</td>
<td>1.31</td>
</tr>
<tr>
<td>Ir2</td>
<td>20–3.25</td>
<td>8,620/30,301</td>
<td>0.070</td>
<td>0.15</td>
<td>0.62</td>
<td>1.13</td>
</tr>
<tr>
<td>DU1</td>
<td>20–4.0</td>
<td>4,056/12,474</td>
<td>0.076</td>
<td>0.21</td>
<td>0.66</td>
<td>1.04</td>
</tr>
<tr>
<td>DU2</td>
<td>20–4.0</td>
<td>4,282/21,806</td>
<td>0.066</td>
<td>0.19</td>
<td>0.67</td>
<td>1.24</td>
</tr>
</tbody>
</table>

*Rsymm = Σ|I–(l)|/ΣI, where l is the observed intensity and (l) is the average intensity obtained from multiple observations of symmetry-related reflections.
†ΔR = Σ|Fcalc| – |Fobs|)/Σ|Fobs|, where |Fobs| is the protein structure-factor amplitude and |Fcalc| is the heavy-atom-derivative structure-factor amplitude.
‡Rcnn = 2Σ|Fobs| – |Fcalc|)/2Σ|Fobs|, where |Fcalc| is the observed heavy-atom structure-factor amplitude, where |Fcalc| is the calculated heavy-atom structure-factor amplitude.
§Phasing power = rms (|Fobs|/E), where |Fobs| is the heavy-atom structure-factor amplitude and E is the residual lack of closure.
RESULTS AND DISCUSSION

Molecular Structure. The 487 amino acids and two zinc ions comprising the monomer are folded into two α/β-type quasispherical globular domains to give a comma-like shape, with approximate maximal dimensions of 90 × 55 × 55 Å³. Fig. 1A shows the molecule as a ribbon drawing and Fig. 1B illustrates a stereodrawing of the α-carbon backbone with the two zinc ions. The secondary structure composition of the monomer is 40% α-helix and 19% β-strand.

The N-terminal 150 residues fold to give a five-stranded β-sheet sandwiched between four α-helices. The N terminus itself occurs as the middle strand of the sheet, which has four parallel strands and one antiparallel strand. There is a long loop connecting an α-helix and the fifth β-strand comprising the sheet. This loop contains the site between residues Arg-137 and Lys-138, where trypsin cleaves the hexameric enzyme but leaves the enzyme intact and biochemically active (33). There is no evidence of proteolytic cleavage of this or any other peptide bond in the protein.

An α-helix runs from residue 151 to 170 and connects the N-terminal domain with the C-terminal domain. This region of the molecule is dominated by a central, eight-membered, saddle-shaped β-sheet, which is sandwiched between groups of α-helices and constitutes the hydrophobic core of the C-terminal domain. In addition, there is a smaller threestranded β sheet, which is located on the monomer surface and is involved in interaction with other members of the enzyme hexamer. The two zinc ions and the active site are entirely located within the C-terminal or catalytic domain. This finding is consistent with the results of photoaffinity labeling experiments, which localized the active site to the larger of the two trypsin-cleavage fragments (A.T., unpublished data).

Hexamer Structure. Fig. 2 A and B illustrate the α-carbon backbone of the LAP trimer and hexamer, respectively. The two crystallographic 32 hexamers found in the unit cell are centered at (½, ½, ½) and (½, ½, ½). When viewed down its threefold symmetry axis the hexamer is triangular in shape with a triangle edge length of 115 Å. The maximum extent of the hexamer along the threefold axis is 90 Å. The view of hexamer shown in Fig. 2B compares favorably with similar electron microscopic views (8, 9). In addition, the lower-resolution models based on the results of electron microscopy (8, 9) are also in good agreement with the distribution of protein within the hexamer. The catalytic domains are clustered around the threefold axis, and the upper and lower trimers are related to one another by a twofold rotation. The N-terminal domains extend outward from the catalytic domains and are located far from the center of the hexamer. It is remarkable that our model of the LAP hexamer does not agree with the results of Vasil’ev et al. (13), who used single isomorphous replacement and solvent flattening to calculate an electron-density map for the hexamer at 5-Å resolution. Their model of the hexamer, derived from a “solvent flattened” electron-density map (see figure 2 in ref. 13), does

FIG. 1. (A) Schematic ribbon drawing of the enzyme monomer (31, 32). α-Helices are shown as ribbon helices, β-strands are shown as ribbons with arrows, and loops are drawn as single lines. This view should be compared to that shown in the stereodrawing illustrated in B, which is labeled with appropriate landmarks. (B) Stereodrawing of the α-carbon backbone of the enzyme monomer. The α-carbon atoms are denoted by small open circles and the two zinc ions located in the active site are represented by the large black circles. The N and C termini are labeled N and C, respectively. The trypsin cleavage site is indicated with the word CUT, and another four residues are labeled for the reader’s convenience.
not show electron-density features corresponding to the N-terminal domains.

Although LAP consists of only a single polypeptide chain and has no known cooperativity, the hexameric enzyme complex illustrated in Fig. 2B bears a striking resemblance to the gross features of the structure of the multimeric regulatory enzyme aspartate transcarbamoylase (ATCase; aspartate carbamoyltransferase, EC 2.1.3.2), which is being actively studied in one of our laboratories (34). In LAP the catalytic domain is analogous to the catalytic or C chain of ATCase, and the LAP N-terminal domain is analogous to the regulatory or R chain of ATCase.

Active Site. The enzyme active site is illustrated in Fig. 3, which shows the electron density of the active-site residues and the two zinc ions. The active site is located near the edge of the central eight-stranded \( \beta \)-sheet (see Fig. 1B). The six active sites are themselves located in the interior of the hexamer, where they line a disk-shaped cavity of radius 15 Å and height 10 Å. Access to this interior cavity is via solvent channels that run along the crystallographic twofold axes, which are found in the planes \( z = 0 \) and \( z = \frac{1}{2} \).

The zinc ions are 2.88 Å apart, which is slightly more than twice their covalent radii (1.25 Å) and considerably more than twice their ionic radii (0.75 Å). As suggested by earlier biochemical studies, the metal environments are not equivalent (4). One of the zinc ions appears to be more tightly bound (site 2) than the other. The metal in site 2 is coordinated by one O\(^8\) atom of Asp-255, one O\(^8\) atom of Asp-332, the carbonyl oxygen of Asp-332, and one O\(^8\) atom of Glu-334, in an approximately tetrahedral arrangement. The other zinc ion appears less tightly bound to the protein (site 1) and is coordinated by one O\(^8\) atom of Asp-273, one O\(^8\) atom of Asp-255, and one O\(^8\) atom of Glu-334. No fourth ligand was seen in the electron-density difference maps. At this limit of resolution and stage in the refinement, there are no density features consistent with water molecules visible in the active site. Earlier metal-substitution studies of bovine lens LAP showed that replacement of zinc in either site 1 or site 2 with
magnesium, manganese, or cobalt ions affected both $k_{cat}$ and $K_m$ (14), thereby suggesting that the metal ions influence each other directly or via the protein.

This arrangement of the zinc ions has not been previously observed in x-ray crystallographic studies of zinc-containing proteases. It is, however, somewhat analogous to the arrangement of the metal ions in myohemerythrin, in which the two iron atoms are separated by about 3.23 Å and are partially coordinated by the carboxylate side chains of glutamate and aspartate residues (35). Unlike myohemerythrin, metal binding by LAP does not involve any histidine side chains.

The active site also includes two positively charged amino acid side chains (Lys-250 and Arg-336). The side chain of Lys-250 is located closest to zinc site 1. Arg-336 is nearly equidistant from the two metal-ion sites. It is tempting to speculate that one or both of these side chains may influence catalysis. However, any comments regarding the enzyme mechanism must await refinement of high-resolution structures of suitable enzyme–ligand complexes.

We thank Dr. N. h. Xuong for use of the data collection Resource for Crystallography (University of California, San Diego); Dr. J. Plouggrath for use of the FAST detector (Cold Spring Harbor Laboratory); and Dr. J. E. Gouaux for help with native data collection. We acknowledge the contributions of Drs. P. Moody and G. S. Shoham, who helped with preliminary aspects of this investigation. In addition, we thank Drs. M. Blum, H. Kei, J. Kuriyan, J. Y. Liang, G. A. Petsko, J. Plouggrath, M. Saper, and J. H. Wang for their useful discussions. S. B. thanks Drs. Eugene Braunwald and Marshall Wolf of the Brigham and Women’s Hospital for their support and encouragement. We thank the National Institutes of Health for Grant GM06920 (to W.N.L.) and the Medical Foundation, Inc., for fellowship support of S.K.B. Purification and crystal growth by A.T. were supported by United States Department of Agriculture Contract 53-K06-5-10 and grants from the Massachusetts Lions Eye Research Fund, Inc., the Research Corporation, and the Daniel and Florence Guggenheim Foundation.