Leucine, Isoleucine, Valine-binding Protein from *Escherichia coli*

STRUCTURE AT 3.0-Å RESOLUTION AND LOCATION OF THE BINDING SITE

Mark A. Saper$ and Florante A. Quiocio$^†

From the Department of Biochemistry, Rice University, Houston, Texas 77251

The structure of the leucine, isoleucine, valine-binding protein, an integral part of the high-affinity, branched-chain aliphatic amino acid transport system in *Escherichia coli*, has been solved at 3.0-Å resolution by x-ray crystallography. Five isomorphous heavy atom derivatives, including anomalous differences from a samarium derivative, were used. A model of the polypeptide chain backbone reveals two distinct, globular domains connected by three strands. Each domain consists of a β-sheet core flanked by at least two helices on either side. Difference Fourier analyses of crystals soaked in L-leucine, L-isoleucine, or L-valine have located a single amino acid-binding site in the cleft formed by the two domains. Despite the lack of significant sequence homology, the bilobate and secondary structure observed were similar to that found in the structures of L-arabinose- and D-galactose-binding proteins previously determined in our laboratory.

Periplasmic binding proteins from Gram-negative bacteria are essential components in the high affinity, osmotic shock-sensitive transport systems for carbohydrates, amino acids, and ions (1). Some of these proteins also function as receptors in chemotaxis (2). Of the five binding proteins crystallized in our laboratory (3), the remarkably similar structures of the L-arabinose-binding protein at 2.4-Å resolution (4) and the D-galactose-binding protein at 3.0-Å (5) have been previously determined. A unified understanding of binding protein function requires the comparison of these sugar-binding structures with others differing in substrate specificity.

Osmotic shock-sensitive, high affinity transport of branched-chain aliphatic amino acids by *Escherichia coli* utilizes the leucine, isoleucine, valine-binding protein (6). The purified protein binds one molecule of substrate with a $K_d$ of $2 \times 10^{-7}$ M (7, 8). The amino acid sequence (9) indicates a molecular weight of 36,770 for 344 residues.

The 3.0-Å resolution structure of the leucine, isoleucine, valine-binding protein reported here adds an amino acid receptor to the family of transport protein structures determined in our laboratory. Furthermore, difference Fourier analysis to locate the amino acid-binding site in LIV-BP

$^*$This investigation was supported by Grants GM-26485 and GM-21371 from the National Institutes of Health and Grant C-581 from the Robert A. Welch Foundation. 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.

$^†$Robert A. Welch Foundation Predoctoral Fellow.

$^\dagger$Recipient of a John S. Guggenheim Memorial Foundation Fellowship.

$^\S$The abbreviations used are: LIV-BP, leucine, isoleucine, valine-binding protein; ABP, L-arabinose-binding protein.

Materials and Methods

Materials—L-[1-14C]Leucine (69 mCi/mmol) was obtained from Amersham-Searle; polyethylene glycol 6000 from Fluka AG; guanosine HCl (Sequanal grade) from Pierce; and Tris base (Ultra Pure) from Schwarz/Mann.

Protein Purification and Crystallization—The LIV-BP used in these studies was initially isolated from *E. coli* AE84 and subsequently from *E. coli* AE840218/pOX-5, a plasmid-containing strain with the cloned L-leucine gene for LIV-BP (12). Both strains were kindly provided by Dr. Dale L. Oxender and Dr. James J. Anderson, Department of Biological Chemistry, University of Michigan. The osmotic shock procedure of Willis et al. (13), as modified by Drs. Anderson and Oxender, was used to purify the binding protein. The purified protein exhibited a $K_d$ of $8 \times 10^{-7}$ M (determined by equilibrium microdialysis with radioactive L-leucine as the substrate) and migrated as a single band in native polyacrylamide gel electrophoresis.

In order to crystallize the protein in a form originally obtained by Meador and Quiocio (14), it was necessary to remove endogenously bound amino acid (8) by a rapid, yet mild and effective procedure originally developed in this laboratory for L-arabinose- and D-galactose-binding proteins (15, 16). To monitor the removal of bound substrate from LIV-BP, 5 μl of stock [14C]leucine (100 μCi/ml) was added to 5 ml of LIV-BP (3.6 mg/ml) in 50 mM NaCl, 0.02% NaN3, 10 mM Tris-HCl, pH 7.6. The protein was dialyzed against 1 M guanidine HCl, 10 mM Tris-HCl, pH 7.6, for 24 h at 4 °C. Periodically, protein and dialysate samples were removed and counted to monitor the loss of [14C]leucine from the protein. The protein was exhaustively dialyzed against 10 mM Tris-HCl, pH 7.6, and chromatographed onto a 2 × 10 cm Bio-Gel P2 (100-200 mesh) column equilibrated with the same buffer. Less than 0.1% of [14C]leucine, initially added to the protein, remained after this treatment.

The LIV-BP processed above (10 mg/ml in 50 mM NaCl, 0.02% NaN3, 10 mM Tris-HCl, pH 7.6) was mixed with an equal volume of 20% polyethylene glycol 6000, 50 mM NaCl, 0.02% NaN3, 10 mM sodium citrate, pH 4.2 (precipitant), and equilibrated by micro-vapor diffusion against 2 ml of 18% polyethylene glycol 6000, 0.02% NaN3, 10 mM sodium citrate, pH 4.2. Small, polyhedral crystals appeared within 1 h and grew to about 0.5 mm within 1 week.

A small, well-formed crystal (<0.25 mm) could be grown larger by transferring it to a clean cover slip and rinsing it with 10% polyethylene glycol 6000, 50 mM NaCl, 0.02% NaN3, 5 mM Tris-HCl, 5 mM sodium citrate, pH 4.5, to dissolve micro-nucleation sites. After totally removing the wash solution, a drop consisting of 17 μl of processed protein and 13 μl of precipitant was layered over the crystal and equilibrated versus 18% polyethylene glycol 6000 as before. A single crystal as large as 1.5 mm routinely grew from a 30-μl drop.

The polyhedral crystals of substrate-free LIV-BP, which diffract to at least 2-Å resolution, belong to the space group P2₁2₁2₁ with one molecule/asymmetric unit (14). Unit cell dimensions are shown in Table I.

It is noteworthy that the addition of L-leucine to substrate-free LIV-BP yielded small, needle-shaped crystals under the same crystallization conditions.

Data Collection and Crystallographic Methods—Three-dimensional native and derivative diffraction data were collected on a Syntax P2₁.
diffractometer with CuKα radiation, using a seven-step, 0.2° ω-scan. The data were corrected for Lorentz and polarization effects, deterioration for data used in the structure solution are presented in Table I. The heavy atom parameter refinement and phase calculations utilized PHASEREF, a program written by M. Rossmann and modified by L. F. Ten Eyck and S. J. Remington (see Ref. 31).

RESULTS AND DISCUSSION

Harker sections from difference Patterson maps and threedimensional difference Fourier maps located the major heavy atom derivative sites for an initial isomorphous replacement refinement at 4.1-Å resolution (figure of merit = 0.76). Sites for the samarium derivative were confirmed in anomalous and combined difference Patterson maps (17). A native data scale factor of 41.5 was determined by scaling the native data of the L-arpbinose-binding protein, whose scale had been refined at 2.4-Å resolution.

Extension of the data to 3.0-Å resolution, followed by cycles of parameter refinement and phasing with the Sm anomalous and combined difference Patterson maps (17). A native data scale factor of 41.5 was determined by scaling the L-arpbinose-binding protein, whose scale had been refined at 2.4-Å resolution (4). The scale factor was later changed to 34.0, reducing the refined occupancy of the major samarium site to approximately 190%.

Table I

<table>
<thead>
<tr>
<th>Derivative (concentration, time)</th>
<th>Resolution</th>
<th>Number of crystals</th>
<th>Unit cell dimensions*</th>
<th>Number of overlaps</th>
<th>Rmerge</th>
<th>Number of independent reflections</th>
</tr>
</thead>
<tbody>
<tr>
<td>Native</td>
<td>2.8</td>
<td>11</td>
<td>30.87 (8)</td>
<td>28,201</td>
<td>0.045</td>
<td>7,868</td>
</tr>
<tr>
<td>K3Pt(NO₃)₄ (4 mm, 43 h)</td>
<td>3.5</td>
<td></td>
<td>39.87 (4)</td>
<td>2,557</td>
<td>0.045</td>
<td>4,278</td>
</tr>
<tr>
<td>K₃ReCl₆ (10 mm, 7 days)</td>
<td>3.2</td>
<td>2</td>
<td>39.90 (6)</td>
<td>1,934</td>
<td>0.025</td>
<td>5,770</td>
</tr>
<tr>
<td>Pb(NO₃)₂ (5 mm, 7 days)</td>
<td>3.0</td>
<td>4</td>
<td>39.85 (2)</td>
<td>8,444</td>
<td>0.038</td>
<td>7,441</td>
</tr>
<tr>
<td>SmCl₃ (5 mm, 7 days)</td>
<td>3.0</td>
<td>5</td>
<td>39.90 (8)</td>
<td>16,098</td>
<td>0.047̇</td>
<td>7,593</td>
</tr>
<tr>
<td>(NH₄)₂PbCl₄ (3 mm, 7 days)</td>
<td>4.1</td>
<td>1</td>
<td>39.85 (1)</td>
<td>417</td>
<td>0.022</td>
<td>2,737</td>
</tr>
<tr>
<td>L-Leucine (50 mm, 3 days)</td>
<td>3.5</td>
<td>1</td>
<td>39.72 (1)</td>
<td>1,12</td>
<td>0.039</td>
<td>4,269</td>
</tr>
<tr>
<td>L-Isoleucine (50 mm, 3 days)</td>
<td>3.5</td>
<td>1</td>
<td>40.02 (1)</td>
<td>354</td>
<td>0.028</td>
<td>4,264</td>
</tr>
<tr>
<td>L-Valine (50 mm, 9 days)</td>
<td>4.1</td>
<td>1</td>
<td>39.82 (1)</td>
<td>204</td>
<td>0.021</td>
<td>3,129</td>
</tr>
</tbody>
</table>

*Standard deviations for data from only one crystal were determined from the diffractometer orientation matrix calculation.

†Based on intensity.

‡Rmerge includes Friedel pairs.

by fitting a polyalanine structure to the 3.0-Å map using FRODO, an implementation of Dr. T. A. Jones’ program (19) for the PS 300. Fitting of the side chains for the known amino acid sequence of LIV-BP (9) is now in progress.

The LIV-BP molecule is ellipsoidal, with overall dimensions 40 × 35 × 70 Å, and consists of two globular domains connected by three polypeptide segments (Figs. 2 and 3a). Though lacking sequence similarity, each of the two domains has essentially the same secondary structure arrangement: a central, five-strand parallel β-sheet flanked by two major α-helices on either side running antiparallel to the sheet. The secondary structure is composed of 40% α-helix and 30% β-sheet arranged in the βαβ-folding units commonly seen in other proteins (20). The β-strands from the sheet in each domain run NH₂ to COOH towards the wide cleft between the domains and exhibit the characteristic left-handed propeller twist: the NH₂ domain (Fig. 2, top) sheet swisls about 80° while the COOH domain (Fig. 2, bottom) sheet twists about 170°.

The polypeptide chain begins in NH₂ domain by folding 5 β-

\[ \text{L-Valine} \]

<table>
<thead>
<tr>
<th>TABLE I</th>
</tr>
</thead>
<tbody>
<tr>
<td>Summary of LIV-BP native, derivative, and substrate complex crystal diffraction data</td>
</tr>
<tr>
<td>All crystals were equilibrated in 20% polyethylene glycol 8000, 0.02% NaN₃, 10 mM sodium acetate, pH 4.5, with the indicated compound.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Derivative (concentration, time)</th>
<th>Resolution</th>
<th>Number of crystals</th>
<th>Unit cell dimensions*</th>
<th>Number of overlaps</th>
<th>Rmerge</th>
<th>Number of independent reflections</th>
</tr>
</thead>
<tbody>
<tr>
<td>Native</td>
<td>2.8</td>
<td>11</td>
<td>30.87 (8)</td>
<td>28,201</td>
<td>0.045</td>
<td>7,868</td>
</tr>
<tr>
<td>K₃Pt(NO₃)₄ (4 mm, 43 h)</td>
<td>3.5</td>
<td></td>
<td>39.87 (4)</td>
<td>2,557</td>
<td>0.045</td>
<td>4,278</td>
</tr>
<tr>
<td>K₃ReCl₆ (10 mm, 7 days)</td>
<td>3.2</td>
<td>2</td>
<td>39.90 (6)</td>
<td>1,934</td>
<td>0.025</td>
<td>5,770</td>
</tr>
<tr>
<td>Pb(NO₃)₂ (5 mm, 7 days)</td>
<td>3.0</td>
<td>4</td>
<td>39.85 (2)</td>
<td>8,444</td>
<td>0.038</td>
<td>7,441</td>
</tr>
<tr>
<td>SmCl₃ (5 mm, 7 days)</td>
<td>3.0</td>
<td>5</td>
<td>39.90 (8)</td>
<td>16,098</td>
<td>0.047̇</td>
<td>7,593</td>
</tr>
<tr>
<td>(NH₄)₂PbCl₄ (3 mm, 7 days)</td>
<td>4.1</td>
<td>1</td>
<td>39.85 (1)</td>
<td>417</td>
<td>0.022</td>
<td>2,737</td>
</tr>
<tr>
<td>L-Leucine (50 mm, 3 days)</td>
<td>3.5</td>
<td>1</td>
<td>39.72 (1)</td>
<td>1,12</td>
<td>0.039</td>
<td>4,269</td>
</tr>
<tr>
<td>L-Isoleucine (50 mm, 3 days)</td>
<td>3.5</td>
<td>1</td>
<td>40.02 (1)</td>
<td>354</td>
<td>0.028</td>
<td>4,264</td>
</tr>
<tr>
<td>L-Valine (50 mm, 9 days)</td>
<td>4.1</td>
<td>1</td>
<td>39.82 (1)</td>
<td>204</td>
<td>0.021</td>
<td>3,129</td>
</tr>
</tbody>
</table>

*Standard deviations for data from only one crystal were determined from the diffractometer orientation matrix calculation.

†Based on intensity.

‡Rmerge includes Friedel pairs.

\[ \text{L-Isoleucine} \]

\[ \text{L-Valine} \]

\[ \text{PHASEREF, a program written by M. Rossmann and modified by L. F. Ten Eyck and S. J. Remington (see Ref. 31).} \]

\[ \text{Evans generation} \]

\[ \text{Evans generation} \]

\[ \text{Evans generation} \]

\[ \text{Evans generation} \]

\[ \text{Evans generation} \]
3.0-Å Structure of Leucine, Isoleucine, Valine-binding Protein

FIG. 1. Composite from the electron density map of the leucine, isoleucine, valine-binding protein at 3.0-Å resolution. The protein's two domains are in the center; the larger or NH₂-terminal domain is at the upper right and the smaller or carboxyl-terminal domain at the lower left. The composite is 112 Å wide and consists of nine 1-Å sections up z. Map boundaries in fractional coordinates are \( x = -0.70 \) to 1.20, \( y = -0.35 \) to 1.25, and \( z = 0.106 \) to 0.172.

FIG. 2. Schematic representation of the folding of the leucine, isoleucine, valine-binding protein polypeptide chain determined from the 3.0 Å structure. The view is along the crystallographic z axis, with \( x \) horizontal. Helices are represented by cylinders while \( \beta \)-strands are shown by twisted arrows. Picture produced in part by a computer program written by A. M. Lesk and K. D. Hardman (36).

strands and 4 helices. The chain then traverses to form the other domain (COOH domain) with a supersecondary fold equivalent to the NH₂ domain. The polypeptide wanders into the solvent to create a large interdomain loop with a short helix. Parts of this loop are washed out in the map, indicative of a less rigid or disordered structure. The chain then returns to the NH₂ domain with a long helix, curves to the bottom of the domain forming another short helix, and then meanders with a short two-strand, antiparallel, sheet-like structure before crossing back to the COOH-domain for the final 2 antiparallel \( \beta \)-strands.

Despite little sequence homology (21, 22), the secondary structure packing in each of the LIV-BP domains is very similar to that found in the corresponding domains of the biolate ABP (Fig. 3b) and D-galactose-binding protein. Seventy-three per cent (100 C₆,s) of the NH₂-terminal domain of ABP was found to be equivalent to the corresponding domain of LIV-BP (root mean square distance = 2.37 Å) by the structure orientation technique of Rossmann and Argos (23) using the program OVRPLAP by W. Bennett. Similarly, 112 \( \alpha \)-carbons of the COOH-terminal domain of ABP (also 75%) were superimposed onto the corresponding atoms in the COOH-terminal domain of LIV-BP with a root mean square distance of 2.81 Å. Unusually high equivalence (80%) was also obtained in a comparison between the polypeptide backbone structures of ABP and D-galactose-binding protein (5).

Since the LIV-BP crystals used in the present structural analysis are devoid of bound leucine (see above), difference Fourier analysis was employed to locate the amino acid binding site. Diffraction data were collected from crystals soaked in solutions of 50 mM L-leucine, L-isoleucine, or L-valine (Table I). Each difference map showed a significant peak (≈1.7 times the second highest peak) at the same location (Fig. 4). The proposed binding site is located in the cleft between the two domains but confined only to the wall of the NH₂-terminal domain near two loops that connect \( \beta \)-sheets with helices (Fig. 3a). A model of the appropriate amino acid substrate was fit in the difference electron density. The identification of the amino acid residues involved in binding must await completion of the model fitting.

The structure determination of the leucine, isoleucine, valine-binding protein, together with the recent elucidation of a sulfate-binding protein from *Salmonella typhimurium* at 3.0-
liganded structure, is consistent with the closing of a binding cleft by a hinge-like motion of the two domains. Only modest changes in the protein’s internal energy are necessary to open and close the cleft by a flexible hinge mechanism (26). Local conformational changes, although not precluded, are insufficient to account for the difference in radius of gyration. Furthermore, the structure of ABP, which was solved with bound L-arabinose, displays a closed cleft with the sugar extensively liganded to residues from opposite walls of both domains and completely inaccessible to the solvent (27) (see Fig. 3b). The substrate is trapped within the cleft and the lobes must separate to allow the sugar to be translocated across the membrane.

The crystal structure of substrate-free LIV-BP has a very wide binding cleft, analogous to the proposed model for sugar-free ABP in solution. Upon binding of amino acid, the cleft could conceivably close by a relative twisting of the two domains around the three-strand hinge region. Nevertheless, when amino acid is added to LIV-BP crystals the cleft does not close to bring the opposing COOH-terminal domain in contact with the substrate lodged in the NHn-domain. The complex observed in the crystal is “frozen” in the open form and likely represents an intermediate to the final, liganded form of LIV-BP. Consequently, the complex in the crystal should display an increase in the dissociation constant. This was reflected in crystal-soaking experiments in 10 mM amino acid. Though orders of magnitude greater than the dissociation constant for the complex in solution (0.2 µM) (8), this concentration failed to show saturation; further increases in difference Fourier peak heights were observed when the amino acid concentration was raised to 50 mM (30% for leucine, 60% for isoleucine, and 136% for valine; compare Figs. 4d and 4e).

Other proteins have domains connected by a flexible hinge (28–31). Citrate synthase, for example, contains a large and small domain and has been crystallized in at least two forms: the tetragonal form is open with a wide cleft whereas the monoclinic form is closed with a narrow cleft (31). The product citrate is bound in the cleft of both structures. Interestingly, the open form displays no catalytic activity in the crystals and does not bind coenzyme A. Remington et al. (31) propose a structural change from the open to the closed form upon binding oxaloacetate and discuss the substrate inhibition phenomenon in light of the protein conformational changes. Furthermore, the open form of the enzyme is postulated to be the product release form. Similarly, the structure of yeast hexokinase A complexed with glucose has a closed cleft in contrast to the open conformation of the B isozyme crystallized without glucose (28).

Narrowing of the cleft may be hindered in LIV-BP because of the close packing of the molecules in the crystal (see Fig. 1). A similar reason was cited as the probable cause for the inactivity of citrate synthase in the crystalline state (31). Consistent with this, crystals grown from native, liganded LIV-BP and from substrate-free LIV-BP preincubated with amino acid are similar, but both are morphologically different from those used in this study and may reflect a substrate-induced conformational change.

In high affinity transport, periplasmic binding proteins interact with a membrane aggregate composed of at least three protein components (32–35). Two of these components are largely confined in the cytoplasmic membrane whereas the third is probably associated peripherally (32, 33). Since it is present in only small amounts (32), the aggregate recognizes the liganded form of any one of a group of related binding proteins. Therefore, it is the binding protein that confers substrate specificity to the transport system by forming a tight complex with a specific substrate. A substrate-induced

---

*J. W. Pfugrath and F. A. Quiocho, in preparation.*
FIG. 4. Difference Fourier map composites calculated from LIV-BP crystals soaked in amino acid substrates: 50 mM L-leucine (a) and 50 mM L-isoleucine (b) at 3.5-Å resolution; 10 mM L-valine (d) and 50 mM L-valine (e) at 4.1-Å resolution. The corresponding composite from the native electron density map at 3.0-Å resolution is in c. The proposed amino acid binding site is in the center of each difference map. Note the increase in peak height when the valine concentration is increased from 10 to 50 mM (see text). Each 30-Å square composite is sectioned up the y axis at 1-Å intervals and has fractional bounds: x = 0.677 to 1.43, y = -0.07 to 0.014, and z = 0.484 to 0.744. Each difference map is contoured at 2σ with 2σ-intervals. Negative difference density is represented by thin, dashed lines.
conformational change, perhaps the closing of the binding protein cleft (see above), exposes and positions a recognition site that interacts with the membrane aggregate and initiates transport. A plausible complex formation between the membrane aggregate and the periplasmic receptor is suggested by the elongated and bilobate binding protein structures. If one assumes a “pore” in the membrane aggregate, the binding protein could span across the opening of the pore, each lobe binding to an exposed part of a transmembrane protein. In such an arrangement, the binding site would be positioned directly over the pore, allowing efficient translocation of the substrate.

Acknowledgments—We are indebted to Drs. James J. Anderson and Dale L. Oxender for continued interest and advice, and to Prof. Robert Huber and Dr. Wolfgang Steigemann for the PROTEIN crystallographic program package. J. W. Pflugrath and Drs. M. N. Vyas and N. K. Vyas provided useful technical assistance and encouraging advice.

REFERENCES
Leucine, isoleucine, valine-binding protein from Escherichia coli. Structure at 3.0-A resolution and location of the binding site.

M A Saper and F A Quiocio


Access the most updated version of this article at http://www.jbc.org/content/258/18/11057

Alerts:
- When this article is cited
- When a correction for this article is posted

Click here to choose from all of JBC's e-mail alerts

This article cites 0 references, 0 of which can be accessed free at http://www.jbc.org/content/258/18/11057.full.html#ref-list-1