# Three-Dimensional Structure of 2-Amino-3-ketobutyrate CoA Ligase from *Escherichia coli* Complexed with a PLP−Substrate Intermediate: Inferred Reaction Mechanism

Andrea Schmidt, J. Sivaraman, Yunge Li, Robert Larocque, João A. R. G. Barbosa, Christopher Smith, Allan Matte, Joseph D. Schrag, and Miroslaw Cygler*

*Biotechnology Research Institute and Montréal Joint Centre for Structural Biology, National Research Council of Canada, 6100 Royalmount Avenue, Montréal, Québec H4P 2R2, Canada*

*Received September 19, 2000; Revised Manuscript Received February 23, 2001*

ABSTRACT: 2-Amino-3-ketobutyrate CoA ligase (KBL, EC 2.3.1.29) is a pyridoxal phosphate (PLP) dependent enzyme, which catalyzes the second reaction step on the main metabolic degradation pathway for threonine. It acts in concert with threonine dehydrogenase and converts 2-amino-3-ketobutyrate, the product of threonine dehydrogenation by the latter enzyme, with the participation of cofactor CoA, to glycine and acetyl-CoA. The enzyme has been well conserved during evolution, with 54% amino acid sequence identity between the *Escherichia coli* and human enzymes. We present the three-dimensional structure of *E. coli* KBL determined at 2.0 Å resolution. KBL belongs to the α family of PLP-dependent enzymes, for which the prototypic member is aspartate aminotransferase. Its closest structural homologue is *E. coli* 8-amino-7-oxononanoate synthase. Like many other members of the α family, the functional form of KBL is a dimer, and one such dimer is found in the asymmetric unit in the crystal. There are two active sites per dimer, located at the dimer interface. Both monomers contribute side chains to each active/substrate binding site. Electron density maps indicated the presence in the crystal of the Schiff base intermediate of 2-amino-3-ketobutyrate and PLP, an external aldimine, which remained bound to KBL throughout the protein purification procedure. The observed interactions between the aldimine and the side chains in the substrate binding site explain the specificity for the substrate and provide the basis for a detailed proposal of the reaction mechanism of KBL. A putative binding site of the CoA cofactor was assigned, and implications for the cooperation with threonine dehydrogenase were considered.

Threonine degradation occurs via three pathways; the main pathway involves a two-step process that converts threonine to glycine (*1*−*3*). This pathway (Figure 1) is common to prokaryotic and eukaryotic cells and involves two enzymes: L-threonine dehydrogenase [TDH,[1] EC 1.1.1.103 (*4*−*6*)] and 2-amino-3-ketobutyrate CoA ligase (KBL, EC 2.3.1.29) (*2*, *3*, *7*−*9*). In the first step, which requires the cofactor NAD, TDH converts L-threonine into L-2-amino-3-ketobutyrate. This intermediate then serves as a substrate for KBL, which, in the presence of CoA, catalyzes the conversion of 2-amino-3-ketobutyrate to glycine and acetyl-CoA. Glycine can be utilized as an educt for serine synthesis, while acetyl-CoA is utilized in a large variety of metabolic reactions [e.g., EcoCyc (*10*)]. The reactions catalyzed by TDH and KBL are coupled, both in vivo and in vitro (*9*), and the two enzymes have been shown to form a complex, most likely containing one tetramer of TDH and two dimers of KBL (*7*, *8*, *11*). This concerted mode of action is consistent with the highly reactive nature of the 2-amino-3-ketobutyrate intermediate, which in aqueous solution undergoes spontaneous decarboxylation to yield aminoacetone and $CO_2$. The equilibrium between threonine, glycine, 2-amino-3-ketobutyrate, and aminoacetone is determined by the intracellular concentrations of threonine, glycine, NAD, and CoA (*12*).

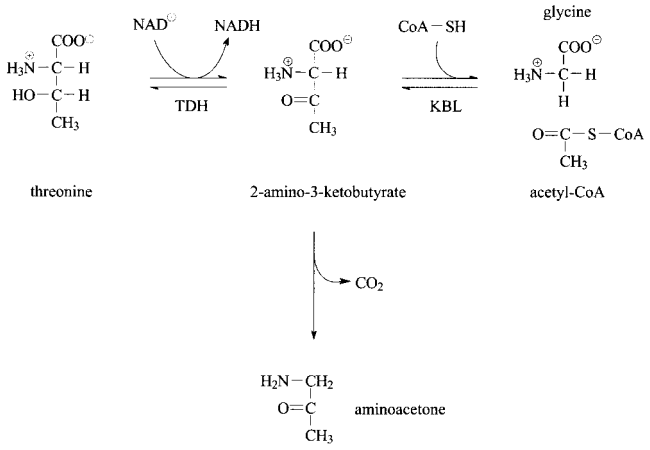


FIGURE 1: Threonine degradation (or synthesis) pathway via threonine dehydrogenase (TDH) and 2-amino-3-ketobutyrate CoA ligase (KBL). Without KBL, the main product of threonine oxidation would be aminoacetone. The equilibrium between all of these compounds is determined mainly by the relative amounts of threonine, NAD, glycine, and CoA. Figure produced in ChemWindows 5.0 (SoftShell Int. Ltd.)/CorelDraw 7.

* Corresponding author. Tel: (514) 496-6321. Fax: (514) 496-5143. E-mail: mirek.cygler@bri.nrc.ca.

[1] Abbreviations: CoA, coenzyme A; NAD, nicotinamide adenine dinucleotide; KBL, 2-amino-3-ketobutyrate CoA ligase; PLP, pyridoxal 5′-phosphate; TDH, L-threonine dehydrogenase; AONS, 8-amino-7-oxononanoate synthase; AON, 8-amino-7-oxononanoate.

10.1021/bi002204y CCC: $20.00 
Published on Web 04/05/2001

2-Amino-3-ketobutyrate CoA ligases (also known as glycine *C*-acetyltransferase or AKB ligase) from both prokaryotes and eukaryotes have been extensively studied (*11*, *13*−*18*). The first enzyme was cloned from *Escherichia coli* (*19*), and cloning of a human enzyme has been described recently (*14*). The sequence identity between the *E. coli* and human enzymes is 54% and indicates strong conservation of KBL during evolution. In addition to coenzyme A, KBL also requires pyridoxal phosphate (PLP) for activity and, therefore, belongs to the class of PLP-dependent enzymes.

The PLP-dependent enzymes form one of the most versatile protein families. The range of reactions catalyzed by these enzymes spans a broad spectrum and includes transaminations, $\beta$-eliminations, racemizations, replacement reactions, and a variety of other reactions (*20*−*23*). With few exceptions, these enzymes are involved in amino acid metabolism (*20*, *23*) and play a vital role as a link between carbon and nitrogen metabolism (*24*).

Several classification schemes of PLP-dependent enzymes have been introduced over the years (*21*, *23*, *25*, *26*). Their classification was far from simple due to generally low sequence homology between these enzymes, even between enzymes sharing the same fold. Initially, the classification was based on sequence alignments and the reaction specificity, i.e. which carbon atom of the quinonoid intermediate, $\alpha$, $\beta$, or $\gamma$, was the target for the catalytic activity (*26*). More recently, the PLP-dependent enzymes have been divided into four fold-type groups on the basis of their three-dimensional structures (*21*, *27*). Sequence analysis using the Family Profile Analysis algorithm (*28*), capable of identifying very distant relationships between protein families, reconciled the sequence-based and structure-based classifications and provided grouping of enzymes in agreement with structural characteristics of 23 proteins with known three-dimensional structures (*23*). The four independent families are named the $\alpha$ family (aspartate aminotransferase family), the $\beta$ family (tryptophan synthase family), the D-alanine aminotransferase family, and the alanine racemase family. While there is some correlation between the family class and the $\alpha$, $\beta$, or $\gamma$ carbon reactive center, there is no one-to-one correspondence between the two classification schemes (*21*, *29*). For example, the $\alpha$ family contains enzymes with either $\alpha$ or $\gamma$ carbon reactivity.

During the course of the reaction the PLP cofactor forms a covalent bond to the enzyme, a Schiff base linkage between the aldehyde carbonyl of PLP and the $\epsilon$-amino group of a lysine (*20*, *21*, *24*, *30*). During catalysis, the lysine side chain of this internal aldimine is replaced by the incoming amino acid substrate to form the external aldimine. The general chemical steps in the PLP-catalyzed reactions are well established (*20*). The differences in the type of reaction catalyzed by various enzymes reflect local differences in the environment of PLP and the specific substrate.

In this study we have investigated the three-dimensional structure of 2-amino-3-ketobutyrate CoA ligase from *E. coli*, a member of the $\alpha$ family of PLP-dependent enzymes. Although neither PLP nor the substrate was added to the protein during purification or crystallization, the crystals contained KBL complexed with pyridoxal phosphate in the form of its external aldimine with 2-amino-3-ketobutyrate. This fortuitously trapped reaction intermediate allowed us to propose the molecular details of the catalytic mechanism. The structure of KBL is similar to that of *E. coli* 8-amino-7-oxononanoate synthase (*29*), with which it shares 35% amino acid identity.

## EXPERIMENTAL PROCEDURES

*Cloning and Expression of the E. coli kbl Gene.* The full-length kbl gene was amplified by PCR with recombinant Taq DNA polymerase (Amersham-Pharmacia) using genomic DNA from the strain MC1061 as a template. Primers were obtained from Hukabel Scientifique Ltée. (Montréal, Canada). The forward primer ORL287 5′-TCTGGAGAATCG**CATATG**CGTGGAG-3′ and reverse primer ORL288 5′-CGCTTTCATCTC**GGATCC**TCAGGCG-3′ incorporating *Nde*I and *Bam*HI restriction sites, respectively (in bold), were utilized for directional cloning into a derivative of pGEX-4T1 vector (Amersham-Pharmacia) in order to obtain an in-frame fusion with *Schistosoma japonicum* glutathione *S*-transferase (GST) and a thrombin cleavage site. This derivative cloning vector contained an insertion, which introduced a *Nde*I site just upstream of the *Bam*HI normally present in pGEX-4T1. Transformants were screened by PCR for the presence of the correct insert, and positives were then restricted with *Hinc*II for confirmation. The plasmid was then transformed into the protein expression strain DL41 from Stratagene (La Jolla, CA).

An overnight culture of the transformed *E. coli* DL41 strain was inoculated in defined LeMaster medium (*31*) containing 100 $\mu$g/mL ampicillin and grown at 37 °C. This culture was diluted 10-fold in the same medium and grown under the same conditions for 2 h, after which time it was transferred to 22 °C and 100 $\mu$M IPTG (Sigma, St. Louis, MO) was added. The induced culture was grown for 20 h before being harvested by centrifugation (15 min, 4 °C, 4000*g*).

*Purification and Crystallization.* Cells were resuspended in 40 mL of lysis buffer [50 mM Tris-HCl, pH 7.5, 0.4 M NaCl, 1 mM EDTA, 1% (w/v) Triton X-100, 5% (w/v) glycerol, 20 mM DTT, and 1 tablet of complete protease inhibitor cocktail from Roche Molecular Biology, Montreal, Quebec, Canada]. After sonication (sonicator W-375, Heat Systems-Ultrasonics Inc., 6 × 30 s at 50% power with 30 s between cycles) the lysate was cleared by ultracentrifugation at 150000*g* at 4 °C for 45 min.

The protein supernatant was passed through a DEAE-Sepharose column (Pharmacia) and the flow-through subsequently loaded on a 5 mL glutathione−Sepharose 4B column (Pharmacia). After two wash steps, first with 4 bed volumes of buffer A [50 mM Tris-HCl, pH 7.5, 1 M NaCl, 5% (w/v) glycerol, 1 mM EDTA, and 1% (w/v) Triton X-100) and then with 3 bed volumes of buffer B [50 mM Tris-HCl, pH 7.5, 0.2 M NaCl, and 5% (w/v) glycerol], the bound GST fusion protein was cleaved on the beads by addition of ~100 $\mu$g of $\alpha$-thrombin (Haematologic Technologies Inc.) and incubation for 75 min at room temperature. KBL was then eluted from the column in ~3 bed volumes of buffer B, and the protein was concentrated by ultrafiltration to ~8 mg/mL in 20 mM Tris-HCl, pH 7.5, 0.2 M NaCl, and 1 mM EDTA. The protein was assessed for purity by SDS−PAGE and for solution behavior by dynamic light scattering (DLS; DynaPro-801, Protein Solutions, Charlottesville, VA).

Table 1: Data Collection Statistics for KBL MAD Data Sets[a]

| | $\lambda_1$ | $\lambda_2$ | $\lambda_3$ |
|---|---|---|---|
| wavelength (Å) | 0.961123 | 0.97955 | 0.979305 |
| resolution (Å) | 100−2.0 (2.07−2.0) | 100−2.0 (2.07−2.0) | 100−2.0 (2.07−2.0) |
| $R_{sym}$ | 0.040 (0.093) | 0.075 (0.163) | 0.054 (0.135) |
| $\langle I/\sigma \rangle$ | 20.5 (11.6) | 11.5 (12.0) | 16.9 (8.8) |
| completeness | 99.6 (99.0) | 99.9 (99.8) | 99.7 (98.6) |
| redundancy | 3.5 | 3.5 | 3.5 |
| no. of reflections | 351450 | 354763 | 348424 |
| unique reflections | 97421 | 97732 | 97661 |
| unique reflections merged, $F(+),F(-)$ | | 50707[b] | |
| $R$-factor[c] (no. of reflections) | | 0.151 (45591) | |
| $R$-free[d] (no. of reflections) | | 0.206 (5094) | |

[a] Values for the last shell are given in parentheses. [b] Refinement performed using merged data collected at $\lambda = 0.97955$ Å. [c] $R$-factor $= \sum|F_o - F_c|/\sum F_o$. [d] $R$-free $= R$-factor, but for a random set of 10% of the unique reflections.

Crystals were grown using the hanging drop vapor diffusion method. The droplets contained 2 $\mu$L of protein solution and 4 $\mu$L of reservoir solution [2.2 M $(NH_4)_2SO_4$, 0.1 M sodium/potassium phosphate, pH 6.4, 3.4% (w/v) glycerol, and 3 mM DTT]. Orthorhombic crystals appeared within a few days at 18 °C. The best crystals with cell dimensions $a = 63.94$ Å, $b = 98.66$ Å, and $c = 118.70$ Å and space group $P2_12_12_1$ were obtained by macroseeding. The crystals contain two molecules per asymmetric unit related by a noncrystallographic 2-fold axis parallel to the crystallographic *b*-axis. Prior to data collection, a crystal was transferred into a cryoprotectant solution consisting of 2.2 M $(NH_4)_2SO_4$, 0.1 M sodium/potassium phosphate, pH 6.4, and 23% (w/v) glycerol for 10−20 s. It was then mounted in a Rayon cryoloop (Hampton Research) and flash-frozen in liquid nitrogen at 100 K.

*Data Collection, MAD Phasing, and Model Building.* Data were collected using a Quantum-4 CCD detector (Area Detector Systems Corp.) at the X8C beamline at the National Synchrotron Light Source, Brookhaven National Laboratory, Upton, NY. Diffraction images were collected with 1.5° oscillations and an exposure time of 45 s.

Multiwavelength anomalous dispersion (MAD) data sets were collected on a Se-Met labeled crystal at three different wavelengths: $\lambda_1$ (peak) = 0.979305 Å, $\lambda_2$ (inflection point) = 0.97955 Å, and $\lambda_3$ (high energy remote) = 0.961123 Å at 100 K to 2.0 Å resolution (Table 1). Data processing and scaling were performed with DENZO and SCALEPACK (*32*). All of the 24 expected selenium sites per dimer were identified by SOLVE (*33*) with an overall figure of merit (FOM) of 0.87 in the resolution range 12−2.0 Å. A map calculated after solvent flattening and skeletonization by DM (*34*) was of excellent quality with an improved overall FOM of 0.93. The bones [O (*35*)] clearly showed the secondary structure for both molecules of the dimer in the asymmetric unit. A self-rotation function was calculated using program polarrfn of the CCP4 suite, version 4.0 (*36*), to determine the orientation of the expected 2-fold axis. In addition, a noncrystallographic symmetry search using the selenium positions was carried out using the CCP4 suite.

Starting from the initial MAD-phased map, automated model building was attempted by using warpNtrace (*37*) and refinement as implemented in the ARP-wAR*P* 5.0 program package (*38*). Using warpNtrace, a set of four independent free atom models was built into the initial electron density map and subjected to repeated cycles of refinement with refmac (CCP4) and automatic chain tracing. The protein size was set to 4000 atoms, the resolution range to 12.0−2.0 Å, the refinement cycle number to 75, and the number of molecules in the asymmetric unit to 2. Remaining parameters required by the program were set to default values. After each rebuilding round, consisting of 75 cycles, the resulting partial model was checked visually using the program O, and solvent atoms were removed. The resulting model was used as the starting point for the next run. Selenium atoms found in SOLVE were included in the refinement. When ~200 of the expected 802 residues were traced, phase restraints were applied and rebuilding and tracing continued as outlined above. When the protein chain was 95% complete, side chains were inserted by the side_trace routine of the ARP package using default settings. The model was then adjusted manually using the program O, including corrections of side chain positions and building of missing residues from the N-termini and small loop regions into the electron density. When the protein model was complete, solvent molecules were introduced into the structure using ARP. For the protein, the restraints weight was set to 0.6, phase restraints were omitted. and the refinement cycle number was decreased to 20. No noncrystallographic symmetry restraints were applied throughout the whole refinement process. The progress of model building and refinement was monitored with the *R*-factor, the *R*-free, the connectivity index (during building), the correlation of $F_o$ to $F_c$, and the root-mean-square deviation from mean density. In the region near the putative active site there was residual electron density, which was clearly identified as pyridoxal phosphate forming a Schiff base with the substrate, 2-amino-3-ketobutyrate (Figure 2). In addition, four glycerol molecules were modeled into the appropriate electron density at the surface of the protein. The external aldimine and the glycerol molecules were included in the final steps of the refinement. The crystallographic refinement converged to an *R*-factor of 0.151 (*R*-free of 0.206) and a correlation coefficient ($F_oF_c$) of 0.95 for 50708 reflections (no $\sigma$ cutoff) and 7258 atoms. The structure was validated with PROCHECK (*39*). The final coordinates have been deposited in the Protein Databank with access code 1FC4.

## RESULTS AND DISCUSSION

KBL was expressed at a high level, in the range of 20 mg/L cell culture. The purified protein migrated as a single, 43 kDa band on an SDS−PAGE gel. Dynamic light scattering measurements corresponded to a monodisperse solution with an estimated molecular weight for the protein of ~80 kDa, indicating that KBL exists in solution as a dimer.

*Quality of the KBL Model.* Both monomers (residues Gly-Ser-His-Met1...Ala398) are well defined in the electron density map, including the three additional residues at the N-terminus (Gly-Ser-His) that remain after cleavage with thrombin. Subtle differences between monomers do occur in some short loop regions and side chain positions. The refinement converged to a rather low *R*-factor of 0.15, indicating overall a well-defined structure. The geometry of

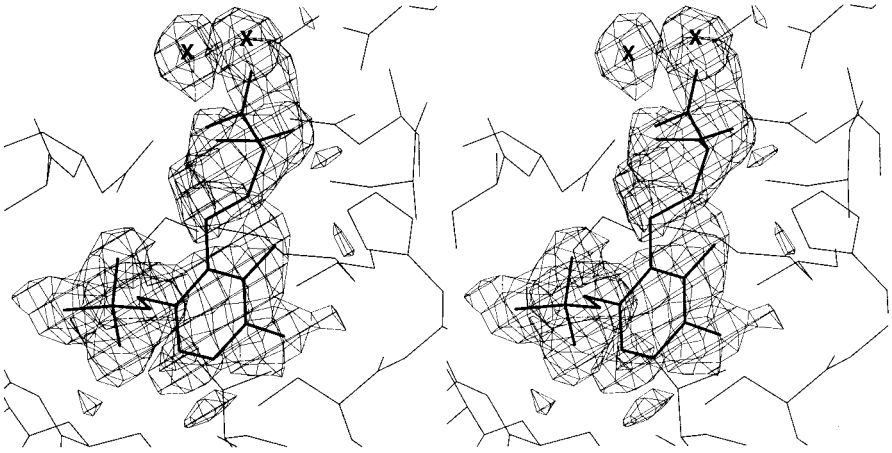

FIGURE 2: Electron density in the simulated annealing omit map for the external aldimine contoured at the 3.0$\sigma$ level. Residues within 4 Å of the aldimine were removed from calculations. Two water molecules are marked with an x. Figure created in O and Oplot (*35*).

the model is very good, with 88.5% of the non-glycine or non-proline residues present in the most favored regions of the Ramachandran plot. Only two residues, the two active site lysines, which are very well defined in terms of electron density, are found in generously allowed regions. This is also observed in other PLP-dependent enzymes, e.g., the structure of bacterial tryptophanase [tryptophan indole-lyase (*40*)]. The external aldimine is for the most part well defined in the electron density (Figure 2) and corresponds to the substrate trapped as a Schiff base with the pyridoxal phosphate cofactor.

*Structure of the KBL Monomer*. KBL belongs to the α family of PLP-dependent enzymes (*23*), for which aspartate aminotransferase is the prototypic enzyme. In topological terms KBL belongs to the α/$\beta$ protein class. The molecule can be divided into three segments: the N-terminal arm, the small domain, and the large PLP-binding domain (Figure 3). A topology diagram is shown in Figure 4.

The N-terminal arm (residues 2−19) is formed by a single helix, which extends away from rest of the monomer and primarily contributes to dimer formation. The small domain is made of two segments assembled from the N-terminal (residues 22−57) and C-terminal (residues 295−398) sections of the primary sequence. This domain consists of two small, nearly perpendicular $\beta$-sheets and three α-helices, which cover the external, solvent-directed surface of the sheets. The PLP-binding domain (residues 58−294) has as its main feature a large, seven-stranded twisted $\beta$-sheet in which all strands but the second one are parallel (Figures 3 and 4). This $\beta$-sheet is found in all PLP-binding proteins belonging to the α family (*21*). In KBL there are three short, extended stretches (strands) adding to the large sheet on the side involved in dimer contacts. Three α-helices cover the concave side of the large sheet and are approximately parallel to the neighboring $\beta$-strands. The convex side of the $\beta$-sheet is covered by six α-helices, one of which is only one turn long. The loops between secondary structure elements are generally very short.

The two domains are assembled at an angle (Figure 3) and are supported by two consecutive, long α-helices (residues 280−294 and 296−314). This arrangement creates a deep cleft between the domains. The concave side of the $\beta$-sheet of the PLP-binding domain faces the side of the $\beta$-sheet of the small domain free of α-helices. The bottom

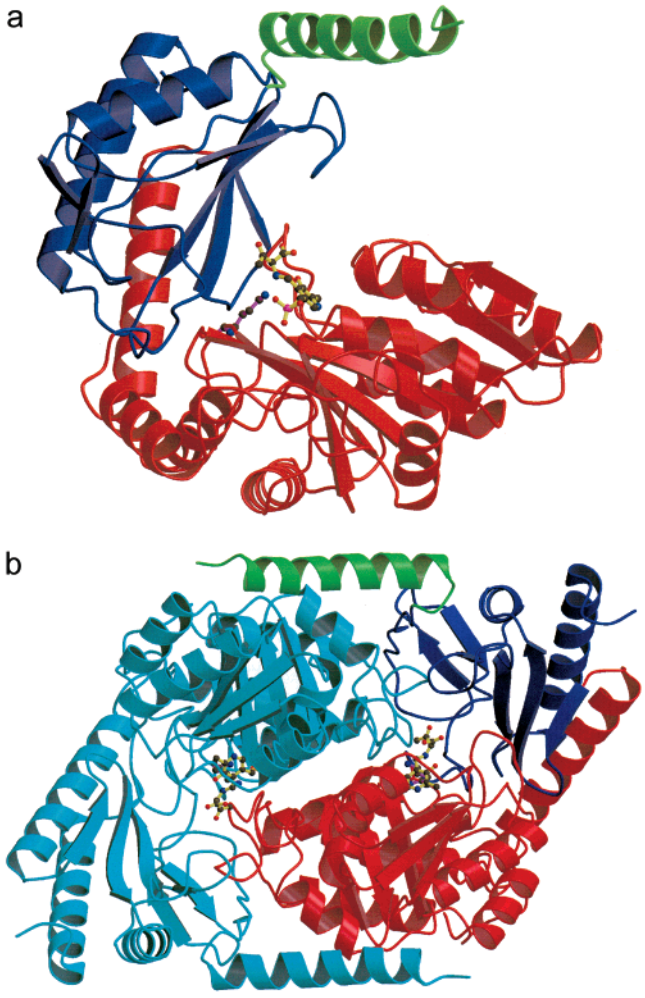


FIGURE 3: Ribbons representation of KBL. (a) A monomer with the three domains shown in different colors: the N-terminal arm is in green, the small domain is in blue, and the PLP binding domain is in red. The external aldimine and Lys244 side chain are shown in ball-and-stick representation. The view is from the side occupied by the second monomer. (b) The dimer, with the color scheme as above for the monomer. The second monomer is in light blue. The view is along the 2-fold axis, and the aldimines are at the bottom of the funnel. Figure produced with MOLSCRIPT (*54*) and Raster3D (*55*).

of the cleft is formed by residues from the C-terminal ends of the $\beta$-strands of the PLP-binding domain on one side and

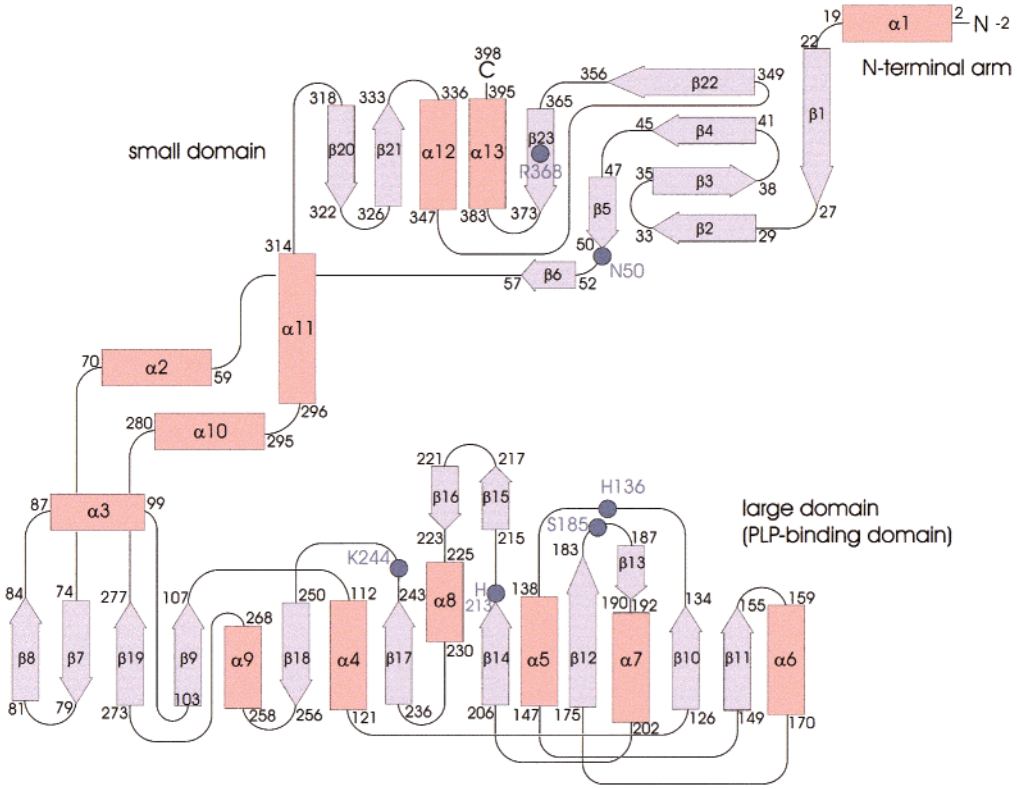


FIGURE 4: Topology diagram for KBL, drawn with CorelDraw 7. It shows the domain structure of the protein. Residues marked by circles are located in the putative active site, in close vicinity to the pyridoxal phosphate moiety.

by the side chains from the $\beta$-sheets of the small domain on the other side. A seven residue long loop from the PLP-binding domain that holds Lys244, the crucial residue that forms a Schiff base with the PLP, protrudes into this cleft (Figure 3).

*Structure of the KBL Dimer*. The two molecules of KBL form a tight dimer. An extensive contact surface is provided by the concave sides of the $\beta$-sheets of the two PLP-binding domains which face each other. The small domain contacts the large domain of the other molecule along the edge of the large $\beta$-sheet. Finally, the N-terminal arm wraps the second molecule along the N-terminal ends of the large $\beta$-sheet. Approximately one-fourth, or 4260 Å$^2$, of each monomer's surface is buried upon association, based on a probe radius of 1.4 Å. The dimer is very compact: as each monomer has 16 226 Å$^2$ accessible surface, the dimer has only 24 355 Å$^2$.

The active site contains residues from both monomers and is fully formed only when the dimer is assembled. The PLP-binding cleft within a monomer is complemented by residues from the second monomer and transformed into a cavity to which access is along a deep, narrowing down funnel-like opening. Side chains from both monomers contact the PLP moiety; however, the putative catalytic residues come from only one monomer. The substrate binding cavity has an amphipathic character. Residues on one side of the pyrimidine ring are polar, acidic, and basic, including Lys244. The other side of the cavity has a hydrophobic character, dominated by several phenylalanines (Figure 5). Access to the substrate binding cavity, as observed in this structure, is too narrow to admit either the PLP or the substrate, implying that a conformational rearrangement within the dimer has to occur for the complete catalytic cycle to take place. This could be of either a local nature, with movement of several loops, or, more likely, a movement of the entire small domain (*20*). Small domain movement upon substrate binding was observed in a structurally related enzyme, 8-amino-7-oxononanoate synthase (*41*), where the active site cleft of a monomer becomes narrower in the presence of the external aldimine. The large buried surface of the KBL dimer makes a dynamic exchange of protein molecules between dimers and monomers rather unlikely. Such equilibrium has been observed, however, when the association of the molecules into a dimer occurs with a relatively small buried surface area, as in the case of ornithine decarboxylase, a PLP-dependent enzyme with only 655 Å$^2$ of buried surface upon association (*42*). The large buried surface area of KBL indicates that the enzyme might need other interactions, like the binding of CoA and/or complex formation with TDH, to trigger conformational changes leading to substrate binding and product release.

*Cofactor Binding.* Although neither PLP nor 2-amino-3-ketobutyrate was added to the protein during purification or crystallization, significant residual electron density was observed in the active site cavity of KBL. A molecule of pyridoxal phosphate fitted well to part of this density, with the strongest feature corresponding to the phosphate group and somewhat lower density for the pyridine ring. The remaining electron density corresponded well to the expected shape of 2-amino-3-ketobutyrate in the form of an external aldimine with PLP (Figure 2). Thus, it appears that we have found in the crystal a trapped intermediate of the reaction catalyzed by KBL. This external aldimine has remained bound within the active site throughout purification and

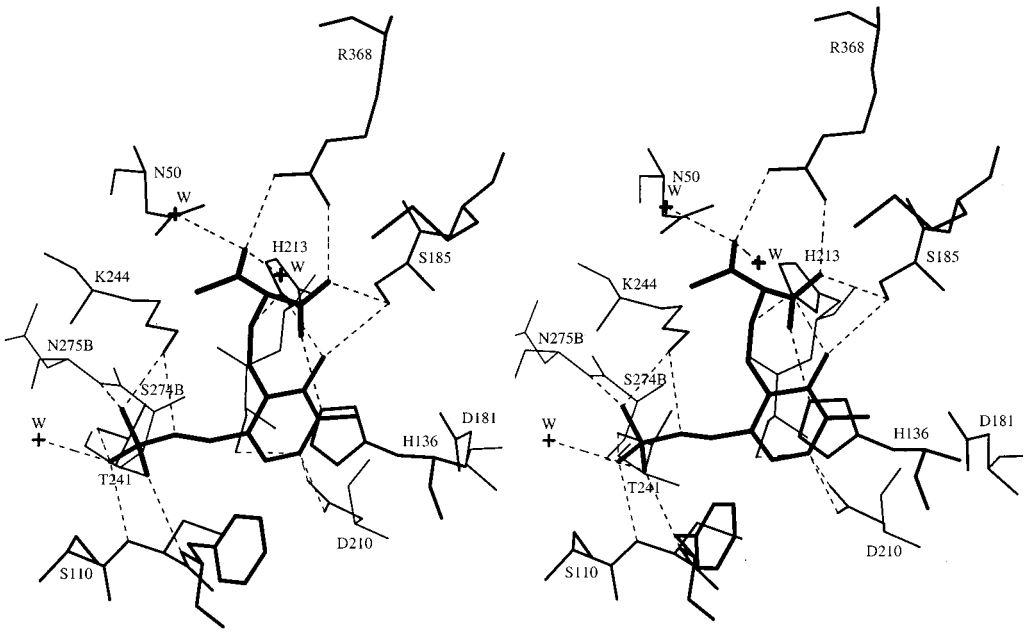


FIGURE 5: Interactions between the external aldimine and the protein. The aldimine is shown in thick lines, residues from the monomer forming the binding site are in medium thickness, and residues from the second monomer that cover the binding site are in thin lines (marked with the letter B after the residue number). Hydrogen bonds are shown as dashed lines, and the positions of water molecules are marked with a +. Lys244 is in a very good position to interact with atoms in the imine link. Figure prepared with MOLSCRIPT.

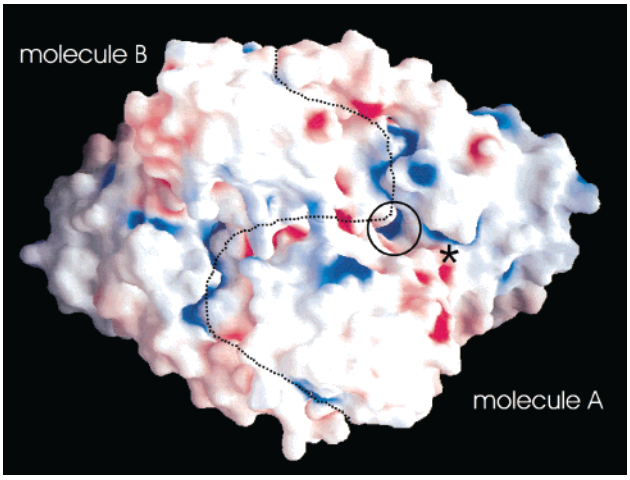


FIGURE 6: Surface representation of the KBL dimer. It shows accumulation of basic charge in the contact zone between the two monomers. The dotted line indicates the dimer interface, and the circle denotes the access funnel to the active site. The asterisk shows the putative CoA binding site. Figure produced with GRASP (*56*)/CorelDraw 7.

crystallization of the enzyme. A key factor contributing to trapping of this intermediate is most likely the absence of coenzyme A, which is necessary for the subsequent reaction step. As the KBL protein is highly overexpressed in *E. coli*, there might be insufficient coenzyme A for the reaction to proceed. The external aldimine formed in the first step of the overall reaction is tightly bound in KBL's substrate binding cavity and is almost completely buried within the dimer interface (Figure 6).

External aldimines have been observed in the crystal structures of other PLP-dependent enzymes, for example, in human ornithine aminotransferase complexed with inhibitors (*43*) and in 8-amino-7-oxononanoate synthase (AONS) complexed with PLP and its substrate (*41*). However, in these cases PLP and/or the substrate had to be introduced deliberately by soaking or cocrystallization and corresponded to the product rather than substrate aldimine.

Several residues in the vicinity of the external aldimine have already been proven to play a role in catalysis, including Lys244 (*18*) and Arg368 (*11*), and in the crystal structure they were indeed found to interact closely with the substrate and cofactor (Figure 5). It is clear from the electron density that there is no covalent linkage between the PLP cofactor and the Lys244 amino group. While Lys244 remains in proximity to the C4′ and the imine nitrogen atom of the PLP, its N$\epsilon$ atom contributes to the stabilization of the phosphate group by forming hydrogen bonds with two of its oxygens. The phosphate is held in place by additional hydrogen bonds involving all of its oxygens and the NH groups of Cys111, Phe112, and Asn275, O$\gamma$ of Ser274, and O$\gamma$1 of Thr241. None of these groups compensates the negative charge of the phosphate, as it is expected that the active site Lys244 is in a deprotonated state in order to carry out its function as a general base and proton abstractor (*44*−*49*). The pyridoxal aromatic ring is held in position by a hydrogen bond from the side chain of Asp210 to the positively charged ring nitrogen (Figure 5), a weak hydrogen bond between the O$\gamma$ of Ser185 and O3′, and stacking against the ring of His136. The imine nitrogen forms a weak hydrogen bond to N$\epsilon$2 of His213, which in turn is hydrogen bonded from N$\delta$1 to the Tyr52 hydroxyl group. From the geometry of the hydrogen-bonding pattern, His213 appears to be in a protonated form. NMR studies on KBL from beef liver mitochondria indicated protonation of this residue (*15*). This is also indirectly supported by the fact that in tryptophanase (*40*) and tyrosine phenol-lyase (*50*) the structural equivalent of this histidine is an arginine, a residue normally carrying a positive charge (Figure 5).

The 2-amino-3-ketobutyrate moiety is found in a conformation close to cis with regard to the C4′−N bond. This differs from the conformation found in other external aldimines, e.g., in AONS (*41*), where the conformation along

this bond is trans. This end of the external aldimine is involved in an extensive network of hydrogen bonds, which includes several water molecules (Figure 5). Arg368, a residue conserved throughout the protein class, interacts with the keto oxygen and with one of the carboxyl oxygens. The same carboxyl oxygen is also hydrogen bonded to Ser185 and to a water molecule. The other carboxyl oxygen is hydrogen bonded to His136, which in turn forms a hydrogen bond to Ser138, and stacks against the pyridine ring of PLP. In the observed conformation the bond to be modified in the next step of the reaction (between Cα and C of the keto group) is nearly in the plane of the pyrimidine ring, while the bond to the acidic group is nearly perpendicular to this plane. Such orientation is opposite to the accepted model of the reaction (*20*, *51*). At the same time, it is the acidic rather than the keto group that is pointing toward the entrance to the funnel leading outside. All this suggests that the 2-amino-3-ketobutyrate as found in this structure may not be in the productive state. The transformation to a productive state would probably require a rotation around the N(imino)−Cα bond by ∼180°. Figure 2 shows that the electron density for the pyrimidine ring is weaker than that for the phosphate group on one end and the ketobutyrate on the other end. This likely indicates that the pyrimidine ring in the center is more mobile than the ends of the aldimine molecule but may also reflect the fact that the two orientations of the ketobutyrate group require somewhat different position of the ring and that we indeed observe a mixture of orientations.

Among sequences of KBL proteins from various species, the residues that make contact with the external aldimine, Lys244, Arg368, Asp210, His213, and His136, are conserved. Extending the comparison to the α family of PLP-dependent enzymes (Figure 5) shows that only Lys244, Arg368, and Asp210 are strictly conserved. His136 seems to play a role only for positioning of the pyridoxal phosphate's aromatic ring, as it is replaced by phenylalanine or tyrosine in most other sequences (Figure 5). His213 is mostly replaced by arginine and in one case by glutamine. As can be expected, those amino acids interacting with the cofactor and the carbonyl group of the amino acid substrate are better conserved than those interacting with the variable side chain of the amino acid substrate.

The molecule found in the active site is a reaction intermediate, which is normally immediately attacked by CoA and cleaved into an acetyl group, which remains attached to the CoA molecule, and a glycine molecule, which remains attached to the PLP. In fact, the funnel leading to the active site cavity is just wide enough to admit the pantothenate/β-mercaptoethylamine arm of a CoA molecule. As well, the local surface shows a cluster of basic residues that could act as a binding site for the phosphate groups. Near this putative phosphate binding site we found a hydrophobic patch that could accommodate the adenine moiety. All of these residues lie in a cleft forming an extension of the funnel and thus may represent the CoA-binding site. Interestingly, there is residual electron density in this area that according to its shape does not correspond to water molecules but is too weak to allow definite assignment as CoA. Nevertheless, modeling a CoA molecule into the cleft showed that it would fit nicely into its assigned binding site (result not shown).

We have subsequently collected data from a crystal soaked in 10 mM CoA for 2 days (2.6 Å resolution) and from KBL crystallized in the presence of CoA (3.2 Å resolution). In both cases we have found in a difference map that the strongest features were at the position occupied by the phosphate group and the Cα atom of the aldimine. These peaks were not connected even at a low contour level. Since ammonium sulfate was present at a concentration of 2.2 M in the crystallization solution, we interpreted the first peak as a sulfate ion and the second as a potential sulfate ion (many potential hydrogen-bonding partners and Arg368). No density corresponding to CoA could be identified. This lack of density may be interpreted as indirect evidence that CoA allowed the reaction to be completed, liberating products and cofactor from the active site.

*Comparison with Related Enzymes.* The three-dimensional structure of KBL shares similarity with other enzymes of the α family that in most cases is restricted to the PLP-binding domain, e.g., 8-amino-7-oxononanoate synthase (*29*), several aminotransferases (*52*), tyrosine phenol-lyase (*50*), tryptophan indole-lyase (*40*), and ornithine transaminase (*53*). All of these enzymes form dimers or tetramers (tyrosine phenol-lyase) in solution and in the crystalline state.

The sequence homology of KBL to other PLP-dependent proteins with known structures is rather low. The best homology in the PDB is found with 8-amino-7-oxononanoate synthase (AONS, PDB code 1BS0), which has 33% sequence identity with KBL. Since a detailed comparison of AONS with other PLP-dependent enzymes has been described, we will limit this discussion mostly to the comparison of KBL with AONS. The latter is involved in biotin biosynthesis and catalyzes the stereospecific, decarboxylative condensation of L-alanine with pimeloyl-CoA (*29*, *41*). It also shows a three-domain structure (a superposition with KBL is shown in Figure 7) with an N-terminal arm consisting of a single helix, a small domain, and a PLP-binding domain (*29*). The superposition of the PLP-binding domains of KBL and AONS (PDB code 1BS0, empty substrate binding site) results in a root-mean-square (rms) deviation of 1.24 Å for 211 Cα atom pairs; the small domains superimpose with an rms deviation of 1.19 Å for 91 Cα atom pairs. Superposition of entire monomers gives an rms deviation of 1.41 Å for 296 Cα atom pairs, indicating small differences in the arrangement of the two domains between KBL and AONS. The dimeric arrangement of these two molecules is also similar. Finally, superposition of dimers of KBL and AONS leads to an rms deviation of 1.42 Å for 584 Cα atom pairs, identical to that for the monomers. The active site architectures of KBL and AONS are also similar. AONS transfers an amino acid moiety to pimeloyl-CoA and shows a similar funnel-shaped access to its catalytic center, which admits the pantothene arm of CoA.

Recently, the structures of the AONS complexes with cofactor and with its external aldimine (PDB code 1DJ9) have been published (*41*), and the comparison with KBL shows that the contacts made by the PLP are very well conserved, including the hydrogen bonds between the AONS active site Lys236 and the PLP phosphate group (Figure 7). However, the directions of the amines attached to PLP are different. The long arm of AON extends into a pocket formed within AONS, which in KBL is occupied by several side chains. The AON conformation along the N(imino)−Cα

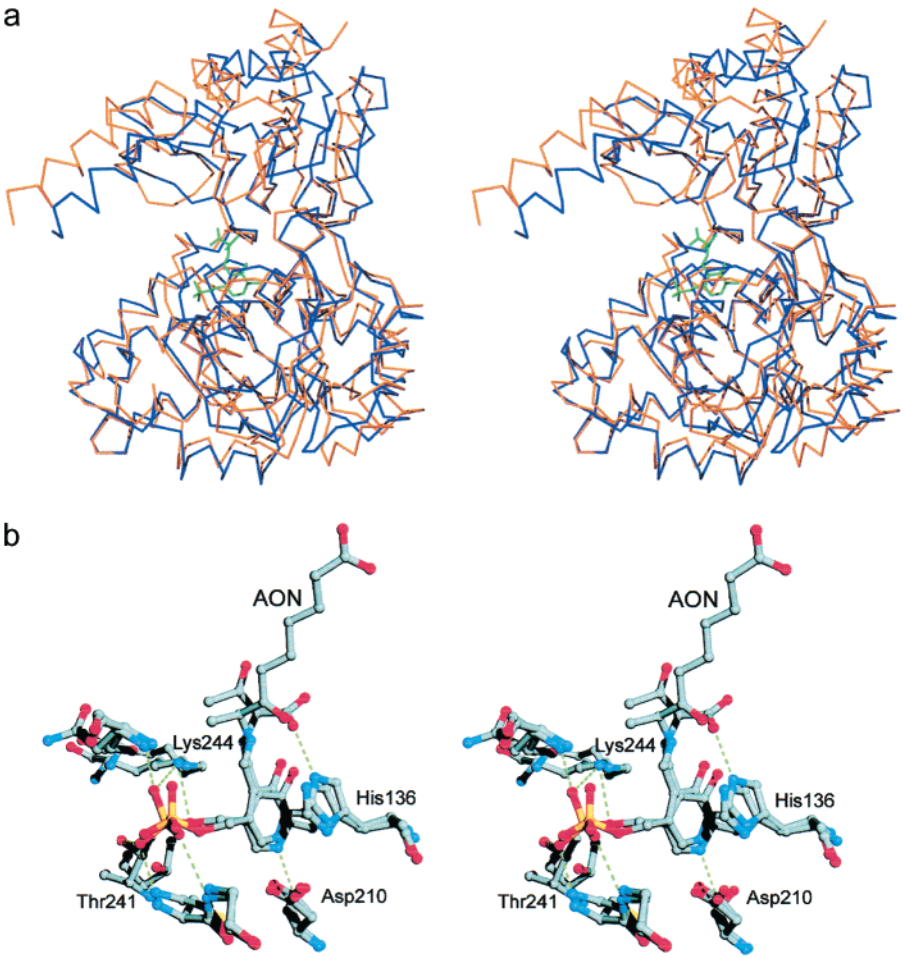


FIGURE 7: (a) Stereoview of the superposition of Cα traces of KBL (gold) and AONS (no PLP, PDB code 1BS0, blue). The superposition is based on the PLP domains only and shows a difference in the orientation of the small domain. The aldimine bound to KBL is shown in green. (b) Close-up on the external aldimines in KBL and AONS (PDB code 1DJ9) after superposition of their PLP-binding domains. Hydrogen bonds and residue numbers are shown for KBL only. Figures created with sPDBviewer (*57*) and POV-ray version 3.1g for Windows (http://www.povray.org/).

bond is trans, and the imino nitrogen is in the plane of the pyrimidine ring of PLP. In KBL aldimine the corresponding torsion angle is ∼−140°, with the N(imino) deviating from the piperidine plane. However, the methyl group and the carbonyl oxygen in AON are in the same locations as the methyl from the keto group and one of the carboxylic oxygens of ketobutyrate in KBL complex (Figure 7). It is also interesting to note that, while the individual domains overlap rather well, their relative disposition is somewhat different in these enzymes. AONS without PLP shows a more open conformation than either AONS or KBL with bound external aldimine (Figure 7). Domain movements in KBL are therefore expected upon substrate binding (*41*).

Alignments of KBL sequences (result not shown) show blocks of very well conserved residues. The level of sequence identity between *E. coli* and human KBLs makes the present structure a very good model for the human enzyme. This is the first structure of a 2-amino-3-ketobutyrate CoA ligase, highly representative for all KBL enzymes.

*Active Site Residues and Their Involvement in Catalysis.* On the basis of the orientation of the cofactor−substrate adduct and the contacts with the protein, we were able to assign catalytic functions to the residues situated in the active site according to the reaction mechanism shown in Figure 8. The entire reaction sequence follows an ordered bi-bi mechanism. The reaction mechanism consists of three steps: the formation of the external aldimine (Figure 8, panel 1), the acetylation of CoA (Figure 8, panel 2), and the product release and restoration of the protein to its original state (Figure 8, panel 3).

Generally, steps 1 and 3 (the transaldimination steps) require an acid catalyst. In KBL, this function is most likely performed by His213, which appears to be in a protonated state. The incoming substrate first replaces Lys244, assisted by His213, forming the external aldimine seen in the crystal structure (Figure 8, panel 1). Amino acids such as Arg368, His136, and Asp210 are important for correct positioning of the reactants. In step 2, CoA binds to the enzyme and removes the acetyl group, with Lys244 abstracting the excess proton and hence functioning as the required base (Figure 8, panel 2). In this process, the formation of a quinoid intermediate is a strong driving force. Acetyl-CoA then dissociates from the protein and Lys244 reprotonates the quinoid intermediate to form the glycine external aldimine (Figure 8, panel 3). Subsequently, in a reverse step of step 1, Lys244 replaces the glycine moiety, which is finally released as the product. The KBL structure shows the state just before the entry of coenzyme A, which removes the acetyl group. As CoA cannot be found in the crystal structure, we are not certain of the specific interactions between CoA

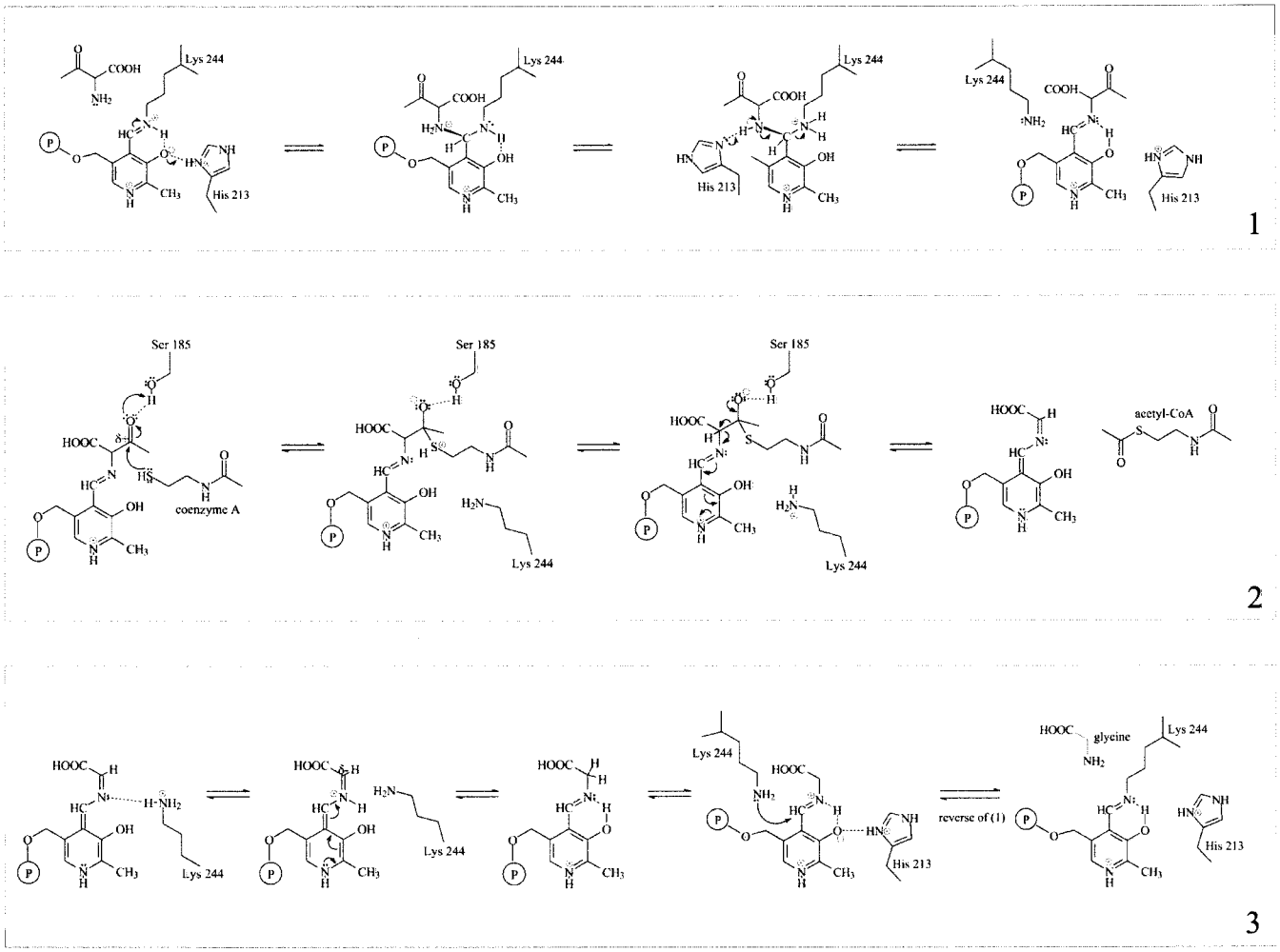


FIGURE 8: Proposed reaction mechanism for KBL. The first step (panel 1) has been established before. Its product is the external aldimine as seen in the KBL structure. Panel 2 shows the incoming of CoA and the release of acetyl-CoA. The second product of this step is the quinoid intermediate. Panel 3 shows how glycine is released and the protein restored to its original state. Figure produced in ChemWindows 5.0/CorelDraw 7.

and the protein. It is likely that Ser185 assists in the nucleophilic attack of the sulfhydryl group of CoA.

In the studies carried out on the mitochondrial 2-amino-3-ketobutyrate CoA ligase, it was postulated that His213 assists the incoming acetyl-CoA in the reverse reaction in the same way as it assists the incoming amino acid substrate during formation of the external aldimine (Figure 8, panel 1). His213 should therefore also be crucial for the formation and release of acetyl-CoA in the forward reaction of *E. coli* KBL. However, we suggest that Ser185 takes over this role, as the ketone carbonyl group of the 2-amino-3-ketobutyrate moiety may form a close H-bonding contact with this residue and hence points away from the OH group of PLP and from His213.

There is experimental evidence that KBL and threonine dehydrogenase (TDH, EC 1.1.1.103), the two enzymes catalyzing two consecutive reactions in threonine degradation, form a complex. This has been determined by gel filtration and fluorescence experiments, and the measured stoichiometry indicates that two dimers of KBL are associated with one tetramer of TDH (*7*). The structure of KBL with the two active sites on the same side of the dimer suggests that KBL binds to TDH with this face opposite the TDH active sites, which would facilitate channeling of the unstable 2-amino-3-ketobutyrate.

## ACKNOWLEDGMENT

We thank Dr. Stéphane Raymond for the establishment and maintenance of our local projects database, Véronique Sauvé for her collaboration in the project, and Dr. Denis Banville for his help in establishing cloning protocols.

## REFERENCES

1. Bird, M. I., and Nunn, P. B. (1983) *Biochem. J. 214*, 687−694.
2. Bell, S. C., and Turner, J. M. (1976) *Biochem. J. 156*, 449−458.
3. Komatsubara, S., Murata, K., Kisumi, M., and Chibata, I. (1978) *J. Bacteriol. 135*, 318−323.
4. Newman, E. B., Kapoor, V., and Potter, R. (1976) *J. Bacteriol. 126*, 1245−1249.
5. Boylan, S. A., and Dekker, E. E. (1981) *J. Biol. Chem. 256*, 1809−1815.
6. Epperly, B. R., and Dekker, E. E. (1991) *J. Biol. Chem. 266*, 6086−6092.
7. Tressel, T., Thompson, R., Zieske, L. R., Menendez, M. I., and Davis, L. (1986) *J. Biol. Chem. 261*, 16428−16437.
8. Bird, M. I., Nunn, P. B., and Lord, L. A. (1984) *Biochim. Biophys. Acta 802*, 229−236.

9. Marcus, J. P., and Dekker, E. E. (1993) *J. Bacteriol. 175*, 6505−6511.
10. Karp, P. D., Riley, M., Saier, M., Paulsen, I. T., Paley, S. M., and Pellegrini-Toole, A. (2000) *Nucleic Acids Res. 28*, 56−59.
11. Mukherjee, J. J., and Dekker, E. E. (1992) *Arch. Biochem. Biophys. 299*, 147−153.
12. Bird, M. I., and Nunn, P. B. (1979) *Biochem. Soc. Trans. 7*, 1276−1277.
13. Mukherjee, J. J., and Dekker, E. E. (1987) *J. Biol. Chem. 262*, 14441−14447.
14. Edgar, A. J., and Polak, J. M. (2000) *Eur. J. Biochem. 267*, 1805−1812.
15. Tong, H., and Davis, L. (1995) *Biochemistry 34*, 3362−3367.
16. Tong, H., and Davis, L. (1994) *J. Biol. Chem. 269*, 4057−4064.
17. Marcus, J. P., and Dekker, E. E. (1993) *Biochem. Biophys. Res. Commun. 190*, 1066−1072.
18. Mukherjee, J. J., and Dekker, E. E. (1990) *Biochim. Biophys. Acta 1037*, 24−29.
19. Aronson, B. D., Ravnikar, P. D., and Somerville, R. L. (1988) *Nucleic Acids Res. 16*, 3586.
20. Hayashi, H. (1995) *J. Biochem. (Tokyo) 118*, 463−473.
21. Jansonius, J. N. (1998) *Curr. Opin. Struct. Biol. 8*, 759−769.
22. John, R. A. (1995) *Biochim. Biophys. Acta 1248*, 81−96.
23. Mehta, P. K., and Christen, P. (2000) *Adv. Enzymol. Relat. Areas Mol. Biol. 74*, 129−184.
24. Vederas, J. C., and Floss, H. G. (1980) *Acc. Chem. Res. 13*, 455−463.
25. Mehta, P. K., Hale, T. I., and Christen, P. (1993) *Eur. J. Biochem. 214*, 549−561.
26. Alexander, F. W., Sandmeier, E., Mehta, P. K., and Christen, P. (1994) *Eur. J. Biochem. 219*, 953−960.
27. Grishin, N. V., Phillips, M. A., and Goldsmith, E. J. (1995) *Protein Sci. 4*, 1291−1304.
28. Mehta, P. K., Argos, P., Barbour, A. D., and Christen, P. (1999) *Proteins 35*, 387−400.
29. Alexeev, D., Alexeeva, M., Baxter, R. L., Campopiano, D. J., Webster, S. P., and Sawyer, L. (1998) *J. Mol. Biol. 284*, 401−419.
30. Morino, Y., and Nagashima, F. (1984) *Methods Enzymol. 106*, 116−137.
31. Hendrickson, W. A., Horton, J. R., and LeMaster, D. M. (1990) *EMBO J. 9*, 1665−1672.
32. Otwinowski, Z., and Minor, W. (1997) *Methods Enzymol. 276*, 307−326.
33. Terwilliger, T. C., and Berendzen, J. (1999) *Acta Crystallogr., Sect. D: Biol. Crystallogr. 55* (Part 4), 849−861.
34. Cowtan, K. D., and Zhang, K. Y. (1999) *Prog. Biophys. Mol. Biol. 72*, 245−270.
35. Jones, T. A., Zhou, J. Y., Cowan, S. W., and Kjeldgaard, M. (1991) *Acta Crystallogr. A47*, 110−119.
36. Collaborative Computational Project Number 4. The CCP4 suite: programs for protein crystallography (1994) *Acta Crystallogr. D50*, 760−763.
37. Perrakis, A., Morris, R., and Lamzin, V. S. (1999) *Nat. Struct. Biol. 6*, 458−463.
38. Lamzin, V. S., and Wilson, K. S. (1993) *Acta Crystallogr. D49*, 129−149.
39. Laskowski, R. A., MacArthur, M. W., Moss, D. S., and Thornton, J. M. (1993) *J. Appl. Crystallogr. 26*, 283−291.
40. Isupov, M. N., Antson, A. A., Dodson, E. J., Dodson, G. G., Dementieva, I. S., Zakomirdina, L. N., Wilson, K. S., Dauter, Z., Lebedev, A. A., and Harutyunyan, E. H. (1998) *J. Mol. Biol. 276*, 603−623.
41. Webster, S. P., Alexeev, D., Campopiano, D. J., Watt, R. M., Alexeeva, M., Sawyer, L., and Baxter, R. L. (2000) *Biochemistry 39*, 516−528.
42. Coleman, C. S., Stanley, B. A., Viswanath, R., and Pegg, A. E. (1994) *J. Biol. Chem. 269*, 3155−3158.
43. Shah, S. A., Shen, B. W., and Brunger, A. T. (1997) *Structure 5*, 1067−1075.
44. Toney, M. D., and Kirsch, J. F. (1993) *Biochemistry 32*, 1471−1479.
45. Toney, M. D., and Kirsch, J. F. (1989) *Science 243*, 1485−1488.
46. Gloss, L. M., and Kirsch, J. F. (1995) *Biochemistry 34*, 3990−3998.
47. Lu, Z., Nagata, S., McPhie, P., and Miles, E. W. (1993) *J. Biol. Chem. 268*, 8727−8734.
48. Watanabe, A., Kurokawa, Y., Yoshimura, T., Kurihara, T., Soda, K., Esaki, N., and Watababe, A. (1999) *J. Biol. Chem. 274*, 4189−4194.
49. Nishimura, K., Tanizawa, K., Yoshimura, T., Esaki, N., Futaki, S., Manning, J. M., and Soda, K. (1991) *Biochemistry 30*, 4072−4077.
50. Antson, A. A., Demidkina, T. V., Gollnick, P., Dauter, Z., von Tersch, R. L., Long, J., Berezhnoy, S. N., Phillips, R. S., Harutyunyan, E. H., and Wilson, K. S. (1993) *Biochemistry 32*, 4195−4206.
51. Dunathan, H. C. (1966) *Proc. Natl. Acad. Sci. U.S.A. 55*, 712−716.
52. McPhalen, C. A., Vincent, M. G., and Jansonius, J. N. (1992) *J. Mol. Biol. 225*, 495−517.
53. Shen, B. W., Hennig, M., Hohenester, E., Jansonius, J. N., and Schirmer, T. (1998) *J. Mol. Biol. 277*, 81−102.
54. Kraulis, P. J. (1991) *J. Appl. Crystallogr. 24*, 946−950.
55. Merritt, E. A., and Bacon, D. J. (1997) *Methods Enzymol. 277*, 505−524.
56. Nicholls, A., Sharp, K. A., and Honig, B. (1991) *Proteins: Struct., Funct., Genet. 11*, 281−296.
57. Guex, N., and Peitsch, M. C. (1997) *Electrophoresis 18*, 2714−2723.


BI002204Y