Solution Structure and Dynamics of the Lipoic Acid-bearing Domain of Human Mitochondrial Branched-chain α-Keto Acid Dehydrogenase Complex*

Received for publication, November 15, 2001, and in revised form, January 22, 2002
Published, JBC Papers in Press, February 11, 2002, DOI 10.1074/jbc.M110952200

Chi-Fon Chang‡, Hui-Ting Chou§, Jacinta L. Chuang¶, David T. Chuang¶, and Tai-huang Huang§‡

From the ‡Institute of Biomedical Sciences, Academia Sinica, Taipei, Taiwan 11529, Republic of China, the §Department of Physics, National Taiwan University, Taipei, Taiwan 10673, Republic of China, and the ¶Department of Biochemistry, University of Texas Southwestern Medical Center, Dallas, Texas 75339

The lipoyl-bearing domain (LBD) of the transacylase (E2) subunit of the branched-chain α-keto acid dehydrogenase complex plays a central role in substrate channeling in this mitochondrial multienzyme complex. We have employed multidimensional heteronuclear NMR techniques to determine the structure and dynamics of the LBD of the human branched-chain α-keto acid dehydrogenase complex (hBLBD). Similar to LBD from other members of the α-keto acid dehydrogenase family, the solution structure of hBLBD is a flattened β-barrel formed by two four-stranded antiparallel β-sheets. The lipoyl Lys residue resides at the tip of a β-hairpin comprising a sharp type I β-turn and the two connecting β-strands 4 and 5. A prominent V-shaped groove formed by a surface loop, L1, connecting β1 and β2-strands and the lipoyl lysine β-hairpin constitutes the functional pocket. We further applied reduced spectral density functions formalism to extract dynamic information of hBLBD from 15N-1H, 15N-13C, and (1H-15N) nuclear Overhauser effect data obtained at 600 MHz. The results showed that residues surrounding the lipoyl lysine region comprising the L1 loop and the Lys β-turn are highly flexible, whereas β-sheet S1 appears to display a slow conformational exchange process.

The mammalian mitochondrial branched-chain α-keto acid dehydrogenase (BCKD) complex catalyzes the oxidative decarboxylation of branched-chain α-keto acids derived from leucine, isoleucine, and valine to give rise to branched-chain acyl-CoAs (1). The reaction products are indirectly channeled into the Krebs cycle or linked to lipid and cholesterol biosynthesis. In patients with inherited maple syrup urine disease, the activity of the BCKD complex is deficient, which results in the accumulation of branched-chain α-keto acids. This metabolic block has severe clinical consequences including often fatal ketoacidosis, neurological derangement, and mental retardation in survivors (2).

The mammalian BCKD complex is a member of the highly conserved α-keto acid dehydrogenase complexes consisting of the pyruvate dehydrogenase complex (PDC), the α-ketoglutarate dehydrogenase complex (KGDC), and the BCKD complex with similar structure and function (3). The mammalian BCKD complex consists of three catalytic components as follows: a heterotetrameric (α2β2) branched-chain α-keto acid decarboxylase or E1, a homo-24-meric dihydrolipoyltransacetylase or E2, and a homodimeric dihydrolipoamide dehydrogenase or E3. E1 and E2 components are specific for the BCKD complex, whereas the E3 component is common among the three α-keto acid dehydrogenase complexes. In addition, the mammalian BCKD complex contains two regulatory enzymes, the BCKD kinase and the BCKD phosphatase that regulate activity of the BCKD complex through phosphorylation (inactivation)/dephosphorylation (activation) cycles (4). The BCKD complex is organized around the cubic E2 core, to which 12 copies of E1, and unspecified copies of E3, the BCKD kinase, and the BCKD phosphatase are attached through ionic interactions (5). The molecular mass of the BCKD multienzyme complex is estimated to be 4 × 10^6 daltons.

The E2 chain of the human BCKD complex consists of three independently folded domains, i.e. the amino-terminal lipoyl-bearing domain (hBLBD), the interim E1/E3-binding domain, and the carboxyl-terminal inner core domain, that are linked through flexible linkers between domains. The inner core domain confers the subunit-binding site for forming the 24-meric structure. The E2 inner core structure from Azobacter vinelandii PDC shows that the 24 E2 subunits are assembled into 8 trimers with each occupying the corners of a hollow cube (6). The active site resides on the interface formed by two neighboring subunits in a trimer. The CoA substrate must enter the 29-Å-long active site channel from the inside of the cube, whereas the lipoamide substrate enters from the outside. The chain folds of the E1/E3-binding domains of the Escherichia coli and Bacillus stearothermophilus KGDCs were determined by NMR spectroscopy (7, 8). This remarkably small domain of 35 amino acid residues is made up of two parallel helices, a 5-residue helix-like turn and an irregular loop. In the E1/E3-
binding domain of the bacterial PDC, one of the two helices binds one molecule of E1 and the other binds E3 but not simultaneously (9). The structures of these domains from the related PDC and KGDC complexes serve as bases for modeling interactions of the E2 component with other components of α-keto acid dehydrogenase complexes (9–11).

The LBD of the α-keto acid dehydrogenase complex mediates the transfer of the acyl moiety from the E1 to the E2 active site. After acyl transfer, the reduced lipoic acid moiety on LBD is re-oxidized by E3 with NAD⁺ as the ultimate electron acceptor. Thus, during the oxidative decarboxylation of α-keto acids, LBD plays a central role in substrate channeling by visiting the three active sites in cognate E1, E2, and E3 (12). Moreover, in mammalian PDC, the inner LBD mediates the feedback regulation of pyruvate dehydrogenase kinase activity by NAD⁺/NADH and CoA-acetyl-CoA. LBD exerts this role by changing its oxidation, reduction, and acetylation state of the lipoyl group (13, 14). The solution structures of LBDS from E. coli PDC (15) and KGDC (16) have also been solved by NMR spectroscopy, along with the structures of LBDS of PDC and KGDC from other species (17–20). The overall structures of LBDS are highly conserved. In each case, the LBD consists of a flattened β-barrel made up of two 4-stranded β-sheets. The lipoic lysine residue is located at the tip of a protruding β-turn in one sheet, with the amino and carboxyl termini adjacent to each other in the other sheet. A prominent surface loop, which is juxtaposed with the amino and carboxyl termini adjacent to each other in

Experimental Procedures

Sample Preparation—hBLBD protein was purified from the BL21(DE3) strain of E. coli containing a plasmid that carried the gene of hBLBD under the control of T7 polymerase. To facilitate protein expression and purification, the recombinant lipoic acid-bearing domain was constructed to contain a total of 93 amino acids including an extra methionine residue at the amino terminus, a leucine, a glutamic acid, and six histidine residues at the carboxyl terminus. However, the wild type protein sequence is used to number the amino acid sequence described in this paper. The uniform [15N- and/or [13C]-labeled proteins were obtained by growing E. coli cells in M9 medium supplemented with [15]NH₄Cl and/or [U-13C]glucose. The protein was purified by nickel-nitritotriacetic acid affinity chromatography, and the purity of the protein was found to be better than 95%, based on the Coomassie Blue-stained gel. NMR samples contain ~2–3 mM protein in 50 mM phosphate buffer, pH 7.5, 100 mM NaCl, 0.02% (w/v) NaN₃, and 10% (v/v) D₂O. To prepare D₂O samples, concentrated protein solutions were lyophilized and redissolved in D₂O.

NMR Experiments and Resonance Assignments—NMR experiments were carried out at 910 K on Bruker AVANCE-600 and AMX-500 NMR spectrometers at Karlsruhe, Germany, both equipped with field gradient accessories. Experimental parameter settings were similar to that described previously (26). [15]N-T₁, [13C]-T₂, and steady-state heteronuclear [1H-15N]-NOE (where T₁ and T₂ are, spin-lattice relaxation time (or longitudinal relaxation time) and spin-spin relaxation time (or transverse relaxation time), respectively) data were obtained at 310 K using standard pulse sequences as described previously (27–29). Water suppression was achieved using the WATERGATE sequence (29) with the States-time proportional phase increment quadrature detection mode in the indirect dimension (31). Two sets of T₁ and T₂ experiments, and three pairs of the steady-state heteronuclear NOE experiments were collected at 600 MHz (14.1 T). Saturation was achieved by applying a power of 31880 Hz (14.1 T) at a time interval of 2 ms. All data were transferred to SGI workstations (Silicon Graphics) for processing and further analysis using Bruker XWINNMR and AURELIA (Auto-mated Resonance Line Assignment) software packages (Karlsruhe, Germany). The proton chemical shifts were referenced to 2.20 dimethyl-2-silapentane-5-sulfonate at 0 ppm. The [15]N and [13C] chemical shift assignments were calculated using the program XEASY (30) from the [15]N/H and 0.251449590 for [15]C/H (33).

Assignment of the main chain H₆, N₃, C₇, C₈, and C₉ chemical shifts were based on CBCA(NH)/CBCA(CO)/NH, HNCO, and HN(CA)/CO experiments.

Assignment of side chain resonances were achieved from analysis of [13]N-labeled TOCSY-HSQC and [13]C-labeled-edited HCCH-TOCSY, HBHAN, and HBHA/CO/NH spectra. The stereospecific assignments were obtained according to the method of Wagner et al. (34).

Assignment of β-protons was obtained from 3J coupling constants estimated from HNHB and HACAHB experiments and relative intensities of the NOEs from the NH-C₂-H and C₃-H-C₄-H protons in NOESY spectra. Assignments of the valine methyl resonances were obtained from the 3J coupling and the NOE cross-peak intensities between the NH protons and NH protons (35). Stereo-specific assignments of the NH group of Asp and Gin residues were made on the basis of the chemical shift difference between the two protons, with the more downfield resonance assigned to the proton and the upfield resonance assigned to the cis proton. H/D exchange rates of the amide protons were determined by recording a series of HSQC spectra in several time intervals after they had been redissolved in D₂O.

Experimental Restraints—NOE distance restraints were derived from three-dimensional [15]N-labeled NOESY-HSQC (mixing time 50 and 100 ms) and [13]C-edited NOESY-HSQC (100 ms) experiments. The NOE spectra were peak-picked and quantified using the automatic picking routine of AURELIA (36). NOESY cross-peaks were assigned both manually and automatically by ARIA (37) using peak lists from NOESY raw data and the chemical shift table for hBLBD. ARIA consisted of a series of routines that partially assigned and calibrated NOE data, performed violation analysis, and merged assigned data from different sources. These assignments were also reconfirmed manually, and the distance constraints (range from 1.8 to 6.0 Å) set by ARIA were used for final structure calculations. Backbone torsion angle, total of 59 φ, were obtained using the [3J(CIH)] coupling constants derived from the HHNCA experiment (36) and the empirical Karplus equation (38). In addition, 23 β torsion restraints predicted from TALOS (39) were found to be consistent with NOE patterns and were included in the final structure calculation. Hydrogen bond constraints were included in calculations only if the corresponding amide protons were found to exchange slowly, and if the β-sheet inter-strand NH-NH, NH-C₂-H₁, and C₃-H₁, C₄-H₁ NOE cross-peaks were observed (34). Two restraints were assigned for each hydrogen bond, 2.5 to 3.3 Å for the N-O distance and 1.5 to 2.3 Å for the H₂O distance.

Structure Calculations and Analysis—All structural calculations were done with a simulated annealing method using PARALLHGD forcefield, starting from random polypeptide chains, aided by the software ARIA interfaced with CNS (40). The parameter setting was similar to that described by Nilges et al. (41). The initial temperature for torsion dynamics was 2000 K with a 50 K decrease per cycle of dynamics. The temperature after the first cooling step was 1000 K, and the final temperature after second cooling step was set to 50 K. Nine iterations (50 structures for the ninth iteration and 20 structures for other iterations) were performed for the last ARIA run, from which 941 unambiguous NOEs and 33 ambiguous NOEs were manually reconfirmed. A total of 974 distance constraints, 164 dihedral angles, and 42 hydrogen bonds were used to generate the final 100 structures (Table I). The best 20 lowest energy structures were further analyzed.

The coordinates of the various LBDS were taken from the Protein Data Bank maintained by the Research Consortium for Structural Biology. The specific files are as follows (all Protein Data Bank codes: 1lac, B. stearothermophilus PDC (17); 1lyu, L. lactis (18); 1pmr, PDC (19); 1fyc, PDC (20); 1kgdc, human PDC (21); 1kgd, mouse PDC (22); 1kgd, human PDC (23); 1kdc, mouse PDC (24); 1kdc, mouse PDC (25); 1kdc, mouse PDC (26); 1kdc, mouse PDC (27); 1kdc, mouse PDC (28); 1kdc, mouse PDC (29); 1kdc, mouse PDC (30); 1kdc, mouse PDC (31); 1kdc, mouse PDC (32)). All these data were analyzed with the programs InsightII (MSI), GRASP (42), MOLMOL (43), and PROCHECK-NMR (44) on SGI computers.
were calculated from peak heights of the resonances in a series of $^1$H-$^{15}$N HSQC spectra. The longitudinal and transverse relaxation times $T_1$ and $T_2$ were obtained by a two-parameter nonlinear optimization of a single exponential function using Marquardt-based nonlinear least squares curve fitting algorithm (SigmaPlot, Jandel Scientific). The uncertainties in the measured peak heights were set equal to the root mean square base-line noise in the spectra. The uncertainties in $R_1$ and $R_2$ (where $R_1$ and $R_2$ are longitudinal relaxation rate constant and transverse relaxation rate constant, respectively) were obtained from curve fitting with SigmaPlot, and the errors in NOE were the standard deviation of three pairs of repeated NOE experiments.

\[ R_1 = \left( 1/T_1 \right) = (d/4)[J(\omega_N - \omega_H) + 3J(\omega_H)] + 6J(\omega_N + \omega_H) + c'd'(\omega_H) \]  
\[ \quad + 6J(\omega_N + \omega_H) + c'd'(\omega_H) \] (Eq. 1)

\[ R_2 = \left( 1/T_2 \right) = (d/4)[J(\omega_H - \omega_N) + 3J(\omega_H) + 6J(\omega_N)] + 6J(\omega_N - \omega_H) + c'd'(\omega_H) \] (Eq. 2)

\[ \text{NOE} = 1 + (d/4)(\gamma_H/\gamma_N)[6J(\omega_H + \omega_N) - J(\omega_N - \omega_H)] \] (Eq. 3)

where $R_1$ ($T_1$) and $R_2$ ($T_2$) are the longitudinal and transverse relaxation rates (times), respectively; $d = (\mu_0/\gamma_H\gamma_NT^2\pi)^2$; $\gamma_N$ is the nuclear gyric ratio of $^1$H and $^{15}$N, respectively; $\omega_N$ and $\omega_H$ are the Larmor frequencies of $^1$H and $^{15}$N, respectively; $\Delta\omega = \omega_1 - \omega_0$ is the $^{15}$N chemical shift anisotropic of polypeptide chains (46). Dynamic parameters are contained in the spectral density function, $J(\omega)$, at five frequencies: $\omega_0$, $\omega_1$, $\omega_2$, $\omega_3$, and $\omega_4$. The exact expression of $J(\omega)$ depends on the motional model used (47–53). Complete determination of the five spectral density functions requires determination of at least five independent relaxation parameters (54, 55) that are not easy to obtain. Fortunately, it was shown experimentally that the spectral density function is a slow varying function at high frequency between $\omega_1 - \omega_0$ and $\omega_4 + \omega_0$. With the additional assumption that $J(\omega) = \lambda/\lambda_0^2 + \lambda_0^2$ where the first and second terms represent contributions to $J(\omega)$ from the overall rotation and internal dynamics, respectively, and substituting a single value of $J(\omega_0)$ for a linear combinations of $J(\omega_0)$, $J(\omega_1)$, and $J(\omega_2)$ in Equations 1–3 gives rise to the following Equations 4–7 (56).

\[ J(0) = (6R_2 - 3R_1 - 2.72\sigma_{NH})[3d^2 + 4c^2] \] (Eq. 4)

\[ J(\omega_0) = [4R_1 - 5\sigma_{NH}][3d^2 + 4c^2] \] (Eq. 5)

\[ J(0.87\omega_0) = 4\sigma_{NH}/(d^2) \] (Eq. 6)

\[ \sigma_{NH} = R_1(\text{NOE}) - 1\gamma_H/\gamma_N \] (Eq. 7)

Thus, for each $^{15}$N site the three spectral density functions, $J(0)$, $J(\omega_0)$, and $J(0.87\omega_0)$, can be determined completely from the three experimental data, $T_1$, $T_2$, and NOE. The $J(0.87\omega_0)$ is essentially the same as the $\text{NOE}(0)$ of Ledevre et al. (57) and $J(\omega_0)$ of Ishima and Nagayama (58). For a rigid globule molecule the overall rotational correlation time $\tau_5$ can be obtained from spectral density functions as described previously (59) and is given by Equation 8.

\[ \tau_5 = \omega_0^{-1}(J(0) - J(\omega_0))/J(\omega_0) \] (Eq. 8)

RESULTS

Resonance Assignments and Folding Integrity—The two-dimensional $^{1}$H-$^{15}$N-HSQC spectrum and assignments of the amide resonances are displayed in Fig. 1. Automated resonance assignment schemes, based on CBCANH, CBCA(CO)NH, HNCO, and HN(CA)CO experiments, were employed to facilitate the assignment of the main chain $H^N$, $H^N$, $C^N$, $C^O$, and $C'$ resonances (60, 61). Under the experimental conditions (pH 7.5 at 310 K), the Trp23 indole $H^1$ and the five pairs of side chain NH$_2$ for Asn62 and all four glutamine residues (Gln2, Gln6, Gln33, and Gln41) were observed. However, no signals were detected for the backbone amides of Met$^{-1}$, Gly1, Gln2, Glu12, Gly14, Arg16, Asp43, Asp84, and the six His tag residues in the carboxyl terminus, suggesting that these $H^N$ nuclei are exposed and undergo fast solvent exchange. Assignments of side chain resonances were achieved by analysis of $^{1}$H-$^{15}$N-edited TOCSY-HSQC and $^{1}$H-$^{15}$C-edited HCCH-TOSY, HBXAH, and HBX(CO/NH) spectra. Nearly all $^1$H, $^{15}$N, and $^{13}$C resonances were assignments for more than 80% of the spin systems, and the chemical shift table was available as entry 5078 in Bio Magnetic Resonance Bank (www.bmrb.wisc.edu). The integrity of the structural fold of hbLBD was verified by comparing the $^{1}$H-$^{15}$N-HSQC spectra of hbLBD with that of a longer construct, the di-domain. The di-domain consists of the amino-terminal hbLBD domain (amino acids 1–84), the linker region (amino acids 85–103), and the downstream peripheral component-binding domain (amino acids 105–152). Essentially all resonances from the hbLBD are present in the di-domain spectrum.

Secondary Structure Determination—Fig. 2 summarizes the NMR parameters employed for identifying the secondary structures of hbLBD (62). These parameters include the consensus chemical shift index of $H^N$, $C^N$, $C^O$, and $C'$ resonances (CSI) (33), $H^N$-$H'$ coupling constants ($J_{NH}$), the medium range backbone NOEs (d$_{NH}$ (i, i + 1), d$_{NH}$ (i, i + 1), d$_{NH}$ (i, i + 1), and H/D exchange rates. Based on the consensus CSI plot, large $J_{NH}$ values (>8 Hz), the presence of medium and long range NOEs, the following eight $\beta$-strands were assigned: $\beta$1(Val1-Lys7), $\beta$2(Val8-Trp21), $\beta$3(Asp8-Ser23), $\beta$4(Leu27-Ser42), $\beta$5(Asp46-Thr59), $\beta$6(Asp54-Leu60), $\beta$7(Asp66-Ala69), and $\beta$8(Pro73-Thr79). From the observation of inter-strand $H^N$-$H'$, $H^N$-$H'$, and $H^N$-$H''$ NOEs these eight $\beta$-strands were found to form two
anti-parallel /H9252-sheets, with the first sheet, S1, formed by /H92521, /H92523, /H92526, and /H92528 and the second /H9252-sheet, S2, composed of /H92522, /H92524, /H92525, and /H92527. Apparently 2-strand is twisted as it uses the same side to bond to both 4- and 7-strands. This may explain the presence of a gap in middle of this region in the CSI plot for this strand.

The Tertiary Fold—The superposition of the resultant family of 20 best structures of the hbLBD is shown in Fig. 3. The coordinates have been deposited in Protein Data Bank (accession numbers: average structure, 1k8m (Protein Data Bank) and RCSB014696 (RCSB); ensemble of 20 best structures, 1k8o (Protein Data Bank) and RCSB014698 (RCSB)). Table I summarizes the structural statistics. These structures have average root mean square deviations of 0.88 Å for all backbone atoms and 0.39 Å for backbone atoms in the well defined secondary structure regions. The quality of these structures is good, as judged by the low root mean square deviation values for bond lengths (0.00207 Å), bond angles (0.40°), and impropers (0.32°) from the ideal geometry. Furthermore, 98.9% of the backbone torsion angles for non-glycine and non-proline residues fall in the most favorable regions, and all residues are in the generally allowed regions in the Ramachandran plot, as calculated with the program PROCHECK_NMR (44) (data not shown).

Fig. 4 shows the ribbon representation of the tertiary folding of hbLBD. Similar to other LBDs, the main feature of the structure is the presence of two /H9252-sheets, folded into the shape of a flattened /H9252-barrel. The order of /H9252-strands in /H9252-sheet S1 is /H92521-8-6-3 and in /H9252-sheet S2 is 5-4-2-7. The lipoxygenation site residue, Lys44, is located at the tip of a type I sharp turn of a /H9252-hairpin formed by 4- and 5-strands on the S2-sheet. The amino and carboxyl termini are close to each other in space and are located on the far end of S1-sheet. Another prominent feature of the structure is the presence of a V-shaped groove between the Lys44 /H9252-hairpin and L1 loop connecting strands 1 and 2. The molecular packing in the middle of this groove is rather sparse, and a thin crevice of ~11 Å long is visible in this region in the space-filling model of hbLBD (structure not shown). There appears to be little interaction between L1 and the Lys44 /H9252-hairpin except the charge interaction between the side chains of Lys44, Asp43, and Arg46. The interior core of the /H9252-barrel is packed with hydrophobic residues, including Phe6,
Reduced Spectral Density Analysis—Shown on the left panels of Fig. 5, a–c, are the relaxation data obtained at 600 MHz, 310 K. The average values of the relaxation parameters are as follows: $R_1 = 2.15 \pm 0.09 \text{ s}^{-1}$; $R_2 = 7.70 \pm 0.20 \text{ s}^{-1}$; and NOE = 0.715 ± 0.016. As expected, $R_1$ remains relative constant over the entire sequence, whereas variation in $R_2$ is the largest. Variation in $R_2$ occurs primarily at the beginning of L1 loop, the Lys 44- turn, and the carboxyl-terminal one-third of the LBD sequence. Variation in NOE appears to be limited to three regions, around L1, the Lys 44- turn, and the carboxyl terminus. To gain more insight into the dynamic behavior of the protein, we have calculated the reduced spectral density functions from Equations 4–7. Reduced spectral density functions reflect the degree of motion in particular frequency regions. Thus, higher value of $J$ indicates the presence of motion at the frequency MHz region. The results are shown on the right panels of Fig. 5, d–f. $J(0)$, $J(N)$, and $J(H)$ represent reduced spectral density function at 0, 60, and 522 MHz, respectively.

**TABLE I**

Summary of structural constraints and structural statistics

| Description                                      | Value |
|--------------------------------------------------|-------|
| All NOE distance restraints*                     | 974   |
| Unambiguous total                                | 94    |
| Ambiguous                                        | 33    |
| Total dihedral angles                            | 164   |
| $\phi$ (degree)                                  | 82    |
| $\Psi$ (degree)$^b$                              | 70    |
| $\chi^1$ (degree)$^b$                            | 12    |
| Hydrogen bonds                                   | 42    |
| RMSD from geometrical analysis                   | 0.0021 ± 0.0001 |
| Bond angles (degree)                             | 0.400 ± 0.002 |
| Improper torsions (degree)                       | 0.317 ± 0.002 |
| Atomic RMSD (secondary region, pairwise)         | 0.39 ± 0.09 |
| Backbone (N, C$_a$, C')                          | 0.91 ± 0.15 |
| All heavy atoms                                  |       |
| Ramachandran statistics                          |       |
| Most favorable region (%)                        | 80.4  |
| Additional allowed region (%)                    | 18.5  |
| Generally allowed region (%)                     | 1.1   |

* Final NOE restraints defined by the program ARIA.

$^b$ Dihedral angles predicted from the program TALOS.

$^c$ Residues 3–7, 18–23, 29–32, 37–42, 46–50, 54–60, 66–68, and 73–79.

**FIG. 4.** The ribbon representation of the structure of hbbLBD with lowest energy and no NOE violation. $\beta$-Sheet S1 is shown in cyan ($\beta1$, $\beta3$, $\beta8$, and $\beta9$) and $\beta$-sheet S2 is colored red ($\beta2$, $\beta4$, $\beta5$, and $\beta7$). The side chain backbone of the lipoyl lysine residue, Lys$^{44}$, is shown in green. Each strand is numbered. L1 denotes the loop connecting strands $\beta1$ and $\beta2$, and the amino and carboxyl termini are labeled.

**FIG. 5.** Variation with sequence of $R_1$ (a), $R_2$ (b), NOE (c), $J(0)$ (d), $J(N)$ (e), and $J(H)$ (f). The relaxation parameters were obtained at 310 K at pH 7.5. The reduced spectral density function, $J(\omega)$, were calculated from the relaxation parameters from Equations 4–7. $J(0)$, $J(N)$, and $J(H)$ represent reduced spectral density functions at 0, 60, and 522 MHz, respectively. Shaded areas highlight the $\beta$-strand regions.
Fig. 6. Superposition of the worm representations and charge distributions of the structures of various LBDs. a and b are two different views of the structure determined in the present study. c–h are structures taken from Protein Data Bank files as described in the text. The worm representation and the charge distribution superimposed in a given figure are shown in the same orientation. Orientations of structures shown in (b–h) are adjusted so that the lipoyl lysine β-hairpins are roughly in the same orientation, looking down the crevice between the lipoyl lysine β-turn and L1. The lipoyl lysines are colored in purple. These figures were generated with the program GRASP (42). The electrostatic potential is calculated from default charge table (full. crg) and colored from red (negative charge) to blue (positive charge). Abbreviations used (Protein Data Bank file names) are: hhLBD (1k8m), LBD from human branched-chain KDC; akLBD (1ghj), LBD from A. vinelandii KGDC; apLBD (1iyu), LBD from A. vinelandii PDC; bpLBD (1lac), LBD from B. stearothermophilus PDC; ekLBD (1pmr), LBD from E. coli KGDC; epLBD (1qjo), LBD from E. coli PDC; and hpLBD (1fyc), the inner lipoyl domain of human PDC.
prominent long loop, L1, connecting strands identical in all LBD structures. Juxtaposed to this hairpin is a Glu13, Gly14, and Arg16 reside in the L1 loop, and Asp43 is in the residues and are expected to be mobile. On the other hand, several segments clearly exhibit motion that deviates from the rigid body motion, as evident from their marked differences in the spectral density functions. In particular, the L1 loop, the Lys44 terminus have much larger J(II), indicative of the presence of picosecond fast motion in these segments. The presence of slow motions in β1-, β7-, and β8-strands can be delineated from their relatively larger values of J(0). It should be noticed that the amide protons of residues Gly1, Gly2, Glu13, Gly14, Arg16, Asp43, and Asp84 cannot be observed. These residues are likely to exhibit dynamic behavior. Gly1, Gly2, and Asp84 are terminal residues and are expected to be mobile. On the other hand, Glu13, Gly14, and Arg16 reside in the L1 loop, and Asp43 is in the Lys44 β-turn. The inability to observe the amide resonances of these residues suggests the presence of motion in the backbone of these residues.

**DISCUSSION**

**Structural Comparison with Other Lipoyl Domain**—In this paper we report the first LBD structure from any BCKD complex. To fulfill the unique biological function of channeling the acyl group, LBD must interact with the three components of the dehydrogenase complex and the regulatory kinase. Interaction of LBD with the E1 component and the kinase is species-specific, thus it is of interest to compare the structural differences of this domain from various species. The sequence homology/identity to the various LBDs are as follows: *B. stearothermophilus* PDC, 58.43%; *A. vinelandii* PDC, 47.33%; *A. vinelandii* KGDC, 42.28%; *E. coli* PDC, 48.37%; and *E. coli* KGDC, 48.28%. A comparison of structures of these LBDs clearly shows the striking similarity of the folding topology and tertiary structure (Fig. 6). These LBD structures invariably possess eight β-strands in similar locations and fold into two anti-parallel β-sheets, each of which contains three major strands and one minor strand. In all cases, the lipoylation site (Lys44) is located at the tip of a type I β-hairpin formed by strands β4 and β5. The conformation of this hairpin is nearly identical in all LBD structures. Juxtaposed to this hairpin is a prominent long loop, L1, connecting strands β1 and β2. This loop is exposed to the surface and was implicated to be a major determinant for the interaction of lipoyl domain with the E1 component (22). Site-directed mutagenesis studies further showed that a structural motif rather than a sequence motif flanking the lipoylation site is responsible for the recognition of this site by the lipoylation enzymes (63). Subsequent studies showed that the interactions between LBD and E1 are more complex and are not attributable to a single determinant on the LBD domain (64–66). In fact alanine substitutions of residues corresponding to Ile11, Gly12, and Gly14 in L1 loop and Thr20 and Glu39 in the far end of the hbLBD structure were shown to have marked reduction in the E1 reaction rate (64). Thus, the interaction site between E1 and LBD appears to encompass surfaces that extend from one end to the other in both structures.

A closer inspection of the seven structures further shows that despite the gross similarities the overall shapes of these structures are visibly different. Of particular importance is the difference in shape near the surface loop and the β-hairpin region. As shown in the worm representations in Fig. 6 when looking down this crevice with the lipoyl lysine hairpin oriented in the same orientation, the orientation and the shapes of the surface loops vary significantly in different LBD structures. The differences are also reflected on the surface charge distributions, which were generated with the GRASP program (42). In the figure, blue patches are the loci of the positively charged residues (lysines and arginines); red patches are the loci of negatively charged residues (aspartic acids and glutamic acids), and white areas are hydrophobic. A prominent feature of the hbLBD is the presence of a large hydrophobic surface at the surface loop loci. Apparently the charged groups Asp10, Glu13, and Glu17 are pointing outward to create a large hydrophobic groove in the center with the conserved residues Asp53 and Lys44 shown as red and blue patches, respectively. The size of this hydrophobic patch differs in different LBDs. Another prominent difference in these structures is the distribution of the charged groups. We propose that the relative shape and orientation of the L1 loop together with the charge distribution near the L1 loop are the main determinants in substrate specificity and species variation in LBDs. Whether and how this hydrophobic patch and the surrounding charged groups interact with the cognate E1, E2, and E3 components remain to be elucidated. Nevertheless, the above observation is in accord with the notion that recognition of the lipoyl domain by its cognate E1 components is complex and cannot be attributed to any particular determinant on the domain (21, 66). It is important to point out that this region is also the least defined region (Fig. 3) due to the scarcity of NOE available and the fact that the molecular packing in the middle of the clamp is rather sparse.
Structural Basis of Mutations with Pathological Consequences—A total of 63 mutations have been described to date in the BCKD complex genes of maple syrup urine disease patients, under the classification of molecular phenotype. Majority of the mutations reported to date affect E1α (type Iα) or E2 (type II) loci (2). Among these mutations, there are six type II missense mutations in hbLBD (E27del, I37M, P73L, P73R, D76Y, and I77T) which affect stability of the E2 subunit and without affect on lipoic acid incorporation. The hbLBD structure shows that the two residues, Ile37 and Ile77, are buried inside the hydrophobic core and are likely to contribute significantly to the stability of the domain. Therefore, the pathological consequence of these two mutations could potentially be due to the affect on stability and folding of the molecule. The other three mutated residues are surface residues on sheet S2. Pro73 is the first residue in strand 8. Due to the unique backbone structure in proline, a mutation in this residue can have significant effects on the overall folding. Glu27 is located in the loop between strands β2 and β3. An in-frame deletion in the residue results in the absence of E2 protein (data not shown), suggesting this residue also has profound effects on folding of the E2 chain. This residue is physically not far from the important surface loop in the catalytic site. It is conceivable that this residue might also be involved in protein-protein interaction with the cognate components. Asp76 resides in the middle of strand 8, and its side chain is also on the surface. Whether the pathological effect of this residue is due to its involvement in recognition with other components of the BCKD complex or due to perturbation on the protein folding remains to be elucidated. Interestingly, with the exception of hydrophobic core residue Ile37 all other missense mutations occur in the S2-sheet, suggesting that β-sheet S2 surface is likely to play an important role in the biological function of LBD.

Dynamics of hbLBD—Proteins are known to be dynamic, and they sample numerous conformational states (67, 68). The concept of a free energy landscape suggests that even the so-called native state of a protein is actually a statistical ensemble of different conformational substrates with similar free energies (69). The degree of ease with which these individual conformers cross the barriers separating them is directly linked to the degree of binding specificity of the protein as well as the degree of flexibility of the protein (70). The role of active site flexibility in enzyme catalysis has been recently emphasized (71, 72). Enzymes enhance biochemical reactions typically by 10^{-10}–10^{-12}fold (73). The source of this enormous rate enhancement is still intensely debated, and many theories explaining catalytic power have been proposed. Each catalytic event requires a minimum of three or more steps, and transition among these states occurs within a few milliseconds. These intramolecular conformational fluctuations, occurring on a wide range of time scales, are suitable for NMR studies (74–77).

Reduced spectral density analysis of the relaxation data of hbLBD (Fig. 5) reveals that most of the backbone amide protons cannot be observed, i.e. Glu13, Gly14, Arg16, and Asp43, were assigned the same values of the preceding residues. The distinct features of these structures are as follows: (i) residues in the lipoyl lysine-containing region, comprising the L1 loop and the β-turn, have the largest J(H) values; (ii) slow motion with large J(0) appears to be localized in and near the strands of S1 β-sheets, i.e. β1, β6, β7, and β8. Fast motions in picosecond time regime are often associated with small order parameter and the presence of flexibility of a region. Slow motions in millisecond and microsecond regimes, on the other hand, are attributed to conformational or chemical exchange. Thus, the lipoyl Lys44 site of hbLBD is the most flexible region. In contrast, the S1 β-sheet, although rather rigid in the fast time scale, shows a slow conformational exchange motion. There is no indication of the presence of motion in the S2 β-sheet. The two strands forming the β-hairpin, i.e. β4 and β5, are also rather rigid. Perham and co-workers (21) have also measured the 15N NMR relaxation parameters of LBD from E. coli PDC. The results are in general agreement with our present study. The residues in and surrounding the lipoyl lysine β-turn are highly mobile and become less flexible after lipoylation of the lysine residue. However, there are two major differences in the two LBDs. First, the L1 loop in the hbLBD as reported in this study appears to be much more mobile than the corresponding loop in E. coli LBD. Only residue Asp13 in the surface loop of E. coli LBD has NOE much lower than residues in the bulk of the structure, by contrast in hbLBD residues from Ser9 to Arg16 are highly mobile. Second, no motion for β-sheet S1 was detected in E. coli LBD even though a slow conformational exchange motion can be clearly delineated in hbLBD.

The presence of high flexibility in the lipoyl Lys44 site suggests that this lipoyc acid-bearing site is not in a preformed configuration. Binding of the flexible lipoylation site region to its cognate enzymes results in conformational rearrangement of interacting residues so as to promote optimal binding required for catalysis, similar to that observed in E. coli thioesterase I (27) and dihydrofolate reductase (72). Because LBD must interact with four different proteins (E1, E2, E3, and the dehydrogenase kinase) the ability of LBD to adapt to specific conformations is essential for these interactions. The structures of BCKD kinase (25) and pyruvate dehydrogenase kinase isozyme 2, both from Rattus norvegicus have been solved recently (78). Both structures reveal the presence of a rather large cavity between the carboxyl- and amino-terminal domains. In the pyruvate dehydrogenase kinase isozyme 2 structure, it was suggested that the binding of LBD to this cleft likely positions the lipoyc acid ligand close to the nucleotide-binding pocket to allow the stimulation of kinase activity by the bound lipoyc acid (78). Furthermore, only a few kinase molecules are associated with the E2 core of the PDC, whereas these kinase molecules must phosphorylate many more E1 molecules bound to the same E2 scaffold (79). To accomplish this task, it has been suggested that the kinase moves over the surface of the E2 core of PDC in a “hand over hand” motion in which the kinase alternately binds and releases from LBDs while remaining bound to the complex (80). The presence of high flexibility in the lipoyl Lys44 region facilitates this type of motion and lends an experimental support to this hypothesis.

Acknowledgments—We thank Dr. M. Nilges for providing the ARIA program and Dr. Chien Ho and Dr. K. V. Chary for providing the computer-assisted resonance assignment programs.
Solution Structure and Dynamics of the Lipoic Acid-bearing Domain of Human Mitochondrial Branched-chain α-Keto Acid Dehydrogenase Complex

Chi-Fon Chang, Hui-Ting Chou, Jacinta L. Chuang, David T. Chuang and Tai-huang Huang

J. Biol. Chem. 2002, 277:15865-15873.
doi: 10.1074/jbc.M110952200 originally published online February 11, 2002

Access the most updated version of this article at doi: 10.1074/jbc.M110952200

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 76 references, 9 of which can be accessed free at
http://www.jbc.org/content/277/18/15865.full.html#ref-list-1