Crystal structure of human branched-chain α-ketoacid dehydrogenase and the molecular basis of multienzyme complex deficiency in maple syrup urine disease

Arnthor Ævarsson[1†], Jacinta L Chuang[2], R Max Wynn[2], Stewart Turley[3], David T Chuang[2] and Wim GJ Hol[1,3*]

Background: Mutations in components of the extraordinarily large α-ketoacid dehydrogenase multienzyme complexes can lead to serious and often fatal disorders in humans, including maple syrup urine disease (MSUD). In order to obtain insight into the effect of mutations observed in MSUD patients, we determined the crystal structure of branched-chain α-ketoacid dehydrogenase (E1), the 170 kDa $\alpha_2\beta_2$ heterotetrameric E1b component of the branched-chain α-ketoacid dehydrogenase multienzyme complex.

Results: The 2.7 Å resolution crystal structure of human E1b revealed essentially the full α and β polypeptide chains of the tightly packed heterotetramer. The position of two important potassium (K^+) ions was determined. One of these ions assists a loop that is close to the cofactor to adopt the proper conformation. The second is located in the β subunit near the interface with the small C-terminal domain of the α subunit. The known MSUD mutations affect the functioning of E1b by interfering with the cofactor and K^+ sites, the packing of hydrophobic cores, and the precise arrangement of residues at or near several subunit interfaces. The Tyr→Asn mutation at position 393-α occurs very frequently in the US population of Mennonites and is located in a unique extension of the human E1b α subunit, contacting the β′ subunit.

Conclusions: Essentially all MSUD mutations in human E1b can be explained on the basis of the structure, with the severity of the mutations for the stability and function of the protein correlating well with the severity of the disease for the patients. The suggestion is made that small molecules with high affinity for human E1b might alleviate effects of some of the milder forms of MSUD.

Addresses: [1]Department of Biological Structure, University of Washington School of Medicine, Seattle, Washington 98195, USA, [2]Department of Biochemistry, University of Texas, Southwestern Medical Center, Dallas, Texas 75235, USA and [3]Howard Hughes Medical Institute, University of Washington, Seattle, WA 98195, USA.

[†]Present address: Prokaria Discovery, Keldnaholt, IS-112 Reykjavik, Iceland.

*Corresponding author.
E-mail: hol@gouda.bmsc.washington.edu

Key words: α-ketoacid dehydrogenase, α-ketoacid dehydrogenase multienzyme complex, maple syrup urine disease, Mennonite, thiamin diphosphate

Received: **31 August 1999**
Revisions requested: **4 November 1999**
Revisions received: **8 December 1999**
Accepted: **22 December 1999**

Published: **28 February 2000**

Structure 2000, **8**:277–291

0969-2126/00/$ – see front matter

Introduction

The human branched-chain α-ketoacid (2-oxo acid) dehydrogenase (BCKD) complex is a member of the mitochondrial α-ketoacid dehydrogenase complex family comprising pyruvate dehydrogenase, α-ketoglutarate dehydrogenase, and BCKD complexes [1]. The multienzyme complexes in the family range in molecular mass from 4 million to 10 million daltons and share the same basic architecture, using numerous copies of three enzymes as major building blocks, referred to as E1, E2 and E3. A dihydrolipoyl transacylase (E2) forms the core of the complex with either 24 copies in octahedral symmetry or 60 copies in icosahedral symmetry depending on the type and source of the complex [2–5]. Multiple copies of an α-ketoacid dehydrogenase (E1), a dihydrolipoamide dehydrogenase (E3), and in some systems of additional proteins, including a protein kinase and a protein phosphatase [1], are attached to the E2 core through non-covalent bonds. E2 is a dynamic multidomain protein with a large catalytic domain which forms the core of the complex, plus a smaller binding domain and one or more lipoyl domains linked to the core through flexible linkers. The E2 binding domain is in many cases responsible for binding E1 and/or E3 to the core whereas the highly mobile lipoyl domain contains a covalently bound lipoamide cofactor that is central to substrate channeling within the multienzyme complex.

The α-ketoacid dehydrogenase multienzyme complexes catalyze analogous reactions in central metabolism. An α-ketoacid (pyruvate, branched-chain α-ketoacids or α-ketoglutarate) is decarboxylated by the E1 component followed by transfer of the remaining acyl group to the lipoamide cofactor bound to E2. Next, the active site of the E2 catalytic domain (in the core of the complex) catalyzes the transfer of the acyl group from lipoamide to coenzyme A (CoA) forming acyl–CoA and leaving the lipoamide in a reduced state. Eventually, the E3 component reactivates the lipoamide by oxidation using a bound FAD cofactor and NAD^+ as external electron acceptor,

thereby completing the functional cycle of the multienzyme complex [2,6].

The BCKD complex [7] is organized around a 24-meric cubic core of the 46.7 kDa dihydrolipoyl transacylase (E2b) with multiple (~6–12) copies of the 170 kDa branched-chain α-ketoacid dehydrogenase (E1b), ~6 copies of the 110 kDa dimeric E3, and a small number of copies of the E1b-specific kinase [8,9] and the E1b phosphatase [10]. The E1 component is a thiamine diphosphate (ThDP)-dependent enzyme consisting of two 45.5 kDa α subunits and two 37.8 kDa β subunits. Hence, the BCKD complex consists of multiple copies of six different polypeptide chains. After import into the mitochondrion and processing of precursor polypeptides, these six proteins fold and assemble, presumably with the assistance of chaperonins in the mitochondrial matrix, into a multienzyme complex of 4–5 million daltons [11,12].

The particular importance of the structure determination of human E1b arises from the fact that the BCKD complex catalyzes the oxidative decarboxylation of branched-chain α-ketoacids (BCKA) derived from the branched-chain amino acids (BCAA) leucine, isoleucine and valine. In patients with autosomally inherited maple syrup urine disease (MSUD), also called branched-chain ketoaciduria, the activity of the BCKD complex is deficient [13]. This results in the accumulation of BCKA and BCAA in plasma, which has severe clinical consequences including often fatal acidosis, neurological derangement and mental retardation [14]. Variations in clinical presentations have led to the classification of MSUD into five distinct phenotypes, classic, intermediate, intermittent, thiamin-responsive and dihydrolipoamide dehydrogenase-deficient [6]. The classic form, which comprises 75% of MSUD patients, is manifested within the first two weeks of life and if untreated, lethargy, seizures, coma and death ensue. Intermediate MSUD is associated with elevated levels of BCAAs and BCKAs, with progressive mental retardation and developmental delay without a history of catastrophic illness. An intermittent form of MSUD might have normal levels of BCAAs, normal intelligence and development until a stress, for example an infection, precipitates decompensation with acidosis and neurologic symptoms, which are usually reversable with dietary treatment. Thiamine-responsive MSUD is similar to the intermediate phenotype but responds to pharmacological doses of thiamine with normalization of BCAAs. The dihydrolipoamide dehydrogenase-deficient MSUD is caused by defects in the E3 component of the BCKD complex, which is shared with the pyruvate and α-ketoglutarate dehydrogenase complexes. Patients with dihydrolipoamide–dehydrogenase deficiency have, therefore, a combined deficiency of all three enzyme complexes, and usually die in infancy with severe lactic acidosis [15].

To understand the molecular and biochemical basis of MSUD, we have undertaken the determination of the crystal structure of human E1b. In this report, we describe the three-dimensional structure of this ThDP-dependent enzyme at 2.7 Å resolution, which reveals several unique structural features that are absent in the homologous enzyme from *Pseudomonas putida* [16]. Most significantly, the structure of human E1b provides remarkable insights into the mechanisms by which MSUD mutations affect the catalysis, folding and assembly of this clinically important protein.

Results and discussion

Description of the structure

The overall heterotetrameric structure of human E1b is shown in Figure 1. To a first approximation, the tetrameric structure is a four-lobed entity with each lobe in the tetrahedral arrangement roughly corresponding to one of the four subunits. Each subunit is in extensive contact with all the other three subunits. Striking features of the subunit interactions are the crossover of the N-terminal tails of the α subunits, and the extended C-terminal domains of the α subunits 'holding onto' the β subunits. The designation of the subunits as α, α′, β and β′ has been chosen such that the α subunit and the β subunit together correspond to one polypeptide of the structurally similar but dimeric yeast transketolase [17] and, by analogy, to one subunit of the dimeric E1s of the α_2 type [18].

Each 400-residue α subunit of human E1b is composed of a large α/β domain (residues 1–355) with an extended N-terminal tail (residues 1–30) and a small helical C-terminal domain (residues 356–400). The central 14-stranded β sheet in the major domain is buried by 13 helices in variable orientations. The secondary structure elements in the α subunit comprise a total of 12 strands and 15 helices (Figures 1,2). Each β subunit of 342 residues is divided into two similar sized domains corresponding to the N-and C-terminal halves of the polypeptide, comprising residues 1–191 and 199–342, respectively. Both domains have an α–β architecture comprising a central β sheet with several helices packed against the sheet on both sides (Figures 1,2).

Structural comparison with homologous enzymes

The structure of human E1b provides an opportunity to compare this eukaryotic enzyme with its bacterial homolog in *P. putida* [16]. Superposition of the two structures gives a root mean square (rms) deviation of 1.8 Å and a sequence identity of 43% for the 658 structurally equivalent residues. The most notable differences between the structures are found at the N and C termini of the α subunits. The N-terminal tail in each case crosses over and embraces the other α subunit but does so in entirely different ways in the two proteins. The 30-residue N-terminal tails of the α subunits in the human E1b are intertwined in a 'firm handshake' whereas the corresponding tails in the *P. putida* structure

extend to opposite sides on the tetramer far from crossing each others' path.

The small C-terminal domain of the human α subunit is 16 residues longer than its counterpart in *P. putida* (Figure 2), resulting in a longer last helix and an irregular tail. This extension in the human structure is the site of an important mutation causing MSUD (Y393N-α; mutations given using single-letter amino acid code) [19] and is critical for the structural integrity of the human E1b tetramer, as will be discussed later.

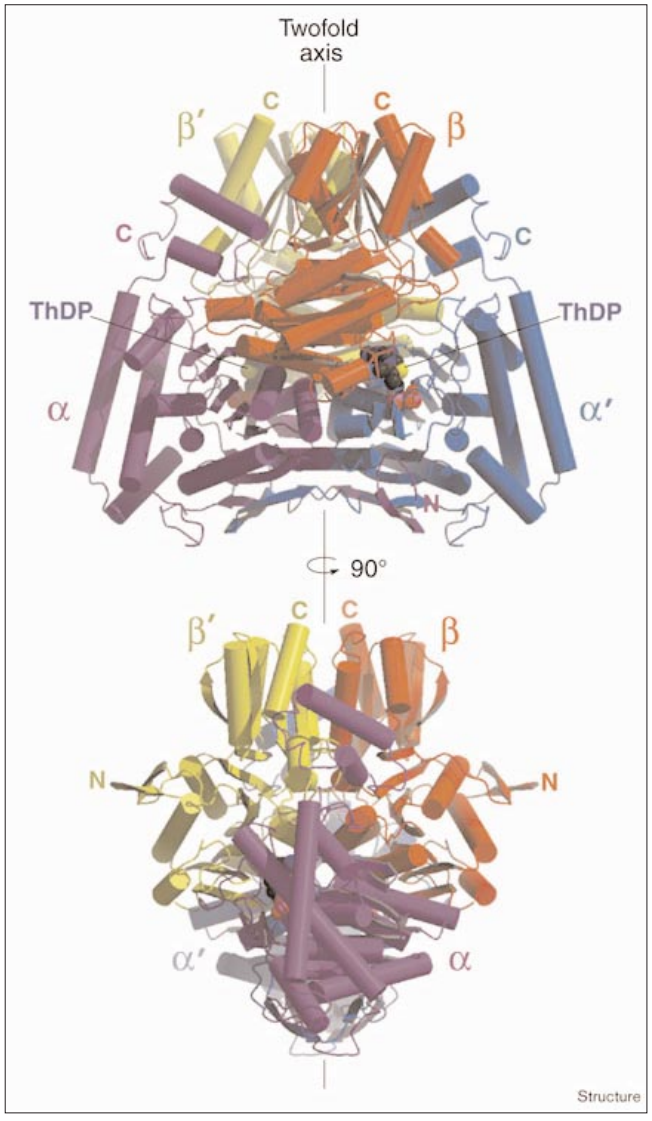
Human E1b and *P. putida* E1b share features with other proteins in the superfamily of ThDP-utilizing enzymes [16,17,20–22]. The common homologous region of these proteins includes the cores of the α subunit and the N-terminal domain of the β subunit. Comparison of the structures reveals how these common domains can have a different order within the same polypeptide [21] or be distributed over different polypeptide chains as is the case in the present structure.

Binding of the ThDP cofactor

The ThDP cofactors are bound in the two active sites of the tetramer at the α–β′ and α′–β interfaces. The sidechains of Gln112-α, Tyr113-α, Arg114-α, Arg220-α, and His291-α and the mainchain atoms of Gly194-α and Ala195-α in the α subunit are in direct contact with the oxygen atoms of the phosphate groups of the cofactor (Figure 3a). The associated magnesium ion is octahedrally coordinated between the cofactor phosphates, the carbonyl of Tyr224-α, and the sidechains of Asn222-α and Glu193-α. The sidechain of Tyr102-β′ is packed against one side of the aminopyrimidine ring of the cofactor with the sidechain of Leu164α approaching the other side of the ring, wedged in between the two ring systems of the cofactor. The sidechains of Ile226-α and Leu74-β′ also contribute to cofactor binding through hydrophobic interactions. Glu76-β′ is directly bound to the N1′ atom and the carbonyl group of Ser162-α is bound to the N4′ group on the opposite side of the aminopyrimidine ring. These two interactions, as well as the coordination of the magnesium ion and the conformation of the cofactor (having torsion angles $\Phi_T = 100°$, $\Phi_p = -71°$ in our structure), are conserved features among proteins of known structure in the ThDP superfamily and probably represent features of functional significance [23–25].

The two serine residues that are phosphorylated in human E1b are Ser292-α and Ser302-α. Inactivation of the enzyme, and consequently of the whole complex, occurs upon phosphorylation of Ser292-α at the primary site [26,27], which is located close to the active site. The primary site Ser292-α is next to the conserved His291-α, which is directly involved in binding the cofactor phosphates and this histidine residue might also play a more direct role in the catalytic mechanism [16,27]. Phosphorylation at Ser292-α most probably prevents substrate and/or cofactor binding through steric hindrance or electrostatic repulsion between the cofactor phosphates and the phosphate group introduced by phosphorylation. The function of His291-α in cofactor binding or catalysis could also be eliminated by the presence of an adjacent phosphate group or by induced conformational changes involving the loop comprising Ser292-α and His291-α. Given that this loop is also found at the opening of the active-site

Figure 1



Overall structure of the heterotetramer in two orthogonal orientations rotated about the vertical axis. The secondary structure elements are depicted as cylinders for helices and arrows for strands. The chain termini in the different subunits are labeled N and C, and the four subunits (α, purple; β, red; α′, blue; β′, yellow) are in 'canonical' colors used throughout the paper. This figure was made with the help of MOLSCRIPT [48] and the Raster3D suite [49].

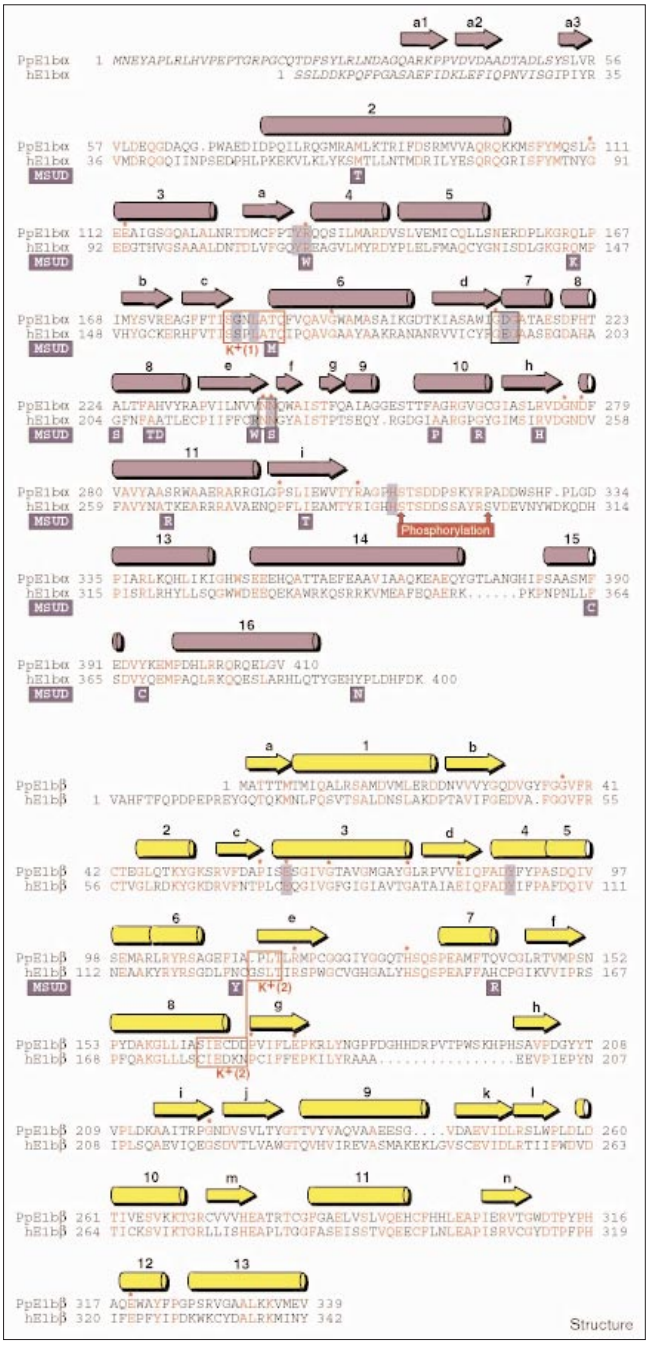


Figure 2

Annotated sequence alignment of *P. putida* and human E1b α and β subunits. The sequences are shown with amino acid residues in red for identical residues and black for non-identical residues. Residues shown in italics have a very different conformation in the two structures and are not structurally equivalent. Secondary structure elements in the human E1b structure are shown with cylinders and arrows above the sequences (purple in the α subunit and yellow in the β subunit). The numbering of secondary structure elements is consistent with that used for *P. putida* E1b [16]. Residues involved in binding the cofactor ThDP and the associated magnesium ion are indicated by shaded boxes. The sequence motif in the superfamily of ThDP-utilizing enzymes is shown using black open boxes. Regions involved in K^+ binding are shown with red open boxes and the two phosphorylation sites are indicated by red arrows. Mutations identified as causing MSUD are shown below the sequences in blue boxes.

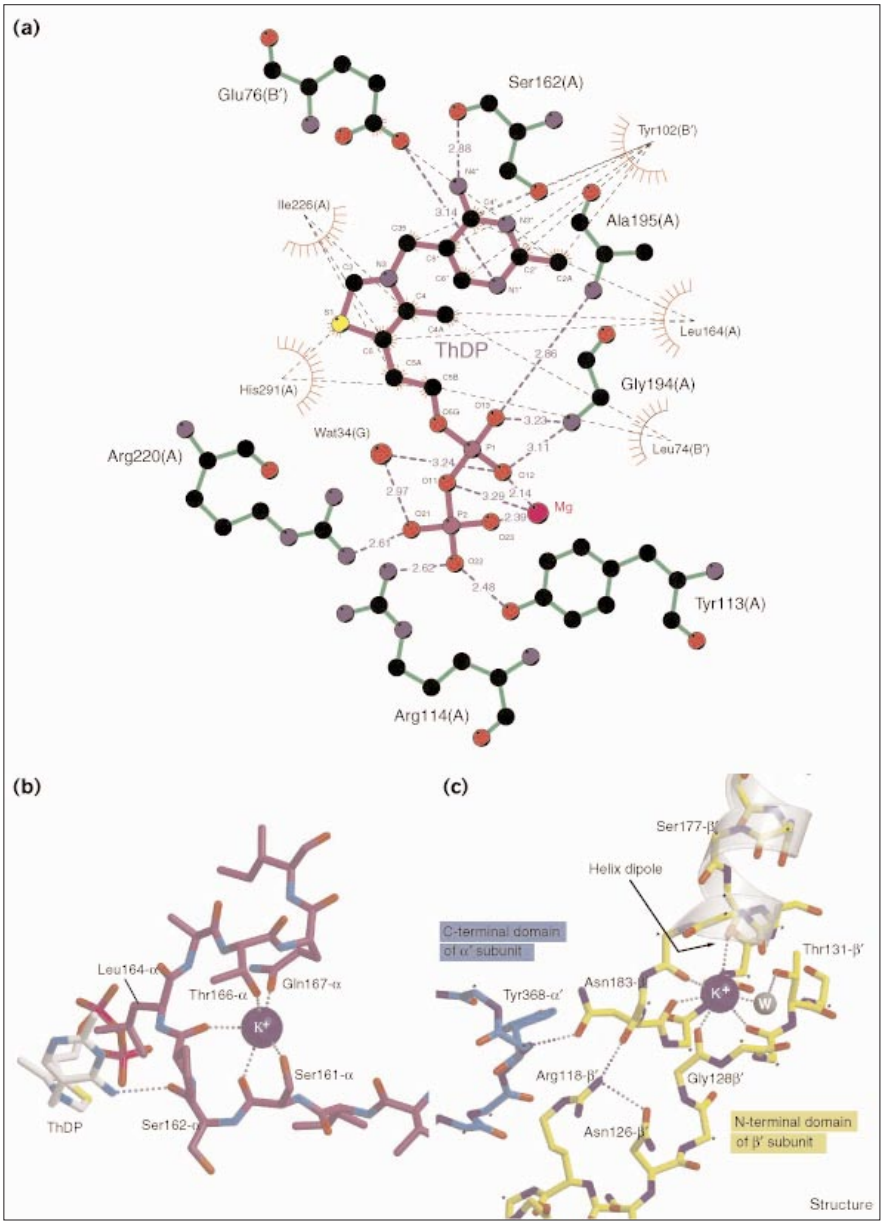
channel, it is not unlikely that the binding sites for the kinase and phosphatase and that of the E2 lipoyl domain overlap to a certain degree.

Structural potassium ions

Initial examination of difference electron-density maps showed the presence of two spherical peaks with high positive density at the 8–10σ level. These peaks were interpreted as monovalent metal ions for four reasons: firstly, the peak heights were considerably above that of all other water molecules, which rarely exceeded 5σ; secondly, the number of surrounding atoms was 5 to 6, considerably more than typically seen for water peaks; thirdly, the groups in contact with the bound entity all contained an oxygen atom; fourthly, when refined as a potassium ion the B factors obtained reasonable values. Crystallization conditions of human E1b included both 10 mM Mg^{2+} and 250 mM K^+. The latter ions are of special interest because Shimomura *et al.* [28] have shown previously that stability and activity of the branched-chain dehydrogenase multienzyme complex from rat liver is to a large extent dependent on the presence of K^+ with the E1 component being the most probable site of metal binding. It was further shown that stabilization of E1 by ThDP is dependent on the presence of K^+ and that stabilization by K^+ ions is likely to be physiologically relevant *in vivo* because the mitochondrial matrix maintains a K^+ concentration of 100 mM [28]. A similar dependence on K^+ for stability is also observed in human E1b [12]. The two positive difference map peaks in the human E1b structure were therefore interpreted as K^+, which agrees with these biochemical observations because they are bound in locations that are likely to be critical for the stability and function of the protein.

The K^+ bound at site 1 (Figure 3b) stabilizes a loop containing residues 161–166 of the α subunit. The residues in the loop are well conserved among E1b proteins including Ser161-α, Thr166-α and Gln167-α, which are directly involved in ligating the metal ion through their sidechains. Other atoms involved in coordination of the metal are mainchain carbonyl oxygens of Ser161-α and Pro163-α. The previously determined structure of *P. putida* E1b [16] probably also contains a K^+ at the corresponding position. Interestingly, the structural integrity of the K^+-binding loop 161–166 is probably very critical for the activity of the enzyme because the residues of this loop contribute directly to cofactor binding by hydrophobic interactions, through Leu164-α wedged between the two rings of the ThDP, and by hydrogen bonding between the carbonyl of Ser162-α and the 4′ amino group of the cofactor (Figure 3b). Both these interactions are a conserved

Figure 3



Cofactor and K^+-binding sites in human E1b. **(a)** Schematic representation of the cofactor-binding site. Gln112-α and protein ligands of the magnesium ion have been omitted for clarity. **(b)** Potassium site 1 (α subunit). The metal stabilizes a loop involved in cofactor binding. The metal ion is bound by two mainchain carbonyl groups and by the sidechains of Ser161-α, Thr166-α and Gln167-α. The sidechain of Leu164-α and the mainchain carbonyl group of Ser162-α make direct contact with the ThDP cofactor. **(c)** Potassium site 2 (β subunit). The metal binding at this site stabilizes regions in the β subunit at the interface with the small C-terminal domain in the α subunit. The metal is octahedrally coordinated mainly by mainchain carbonyl groups and interacts favorably with the C-terminal end of a helix dipole as indicated. Several sidechains indicated by an asterisk have been omitted for clarity. This figure was made with LIGPLOT [50], MOLSCRIPT [48] and the Raster3D suite [49].

feature among ThDP-utilizing proteins with known structure [16,17,20–22]. The binding of the K^+ ion at this position to the α chain of E1b explains why enzyme activity as well as stabilization of the protein by ThDP binding is dependent on the presence of K^+.

The second K^+-binding site is found in the β subunit at the interface with the small C-terminal domain of the α subunit (Figure 3c). The metal ion connects two regions in the β subunit with residues Gly128-β, Leu130-β, Cys178-β, Asp181-β and Asn183-β being directly involved in binding the metal. The coordination of the second K^+ is octahedral and is almost exclusively maintained by carbonyl oxygens of the protein mainchain, with only one sidechain (Thr131-β) indirectly involved in metal binding through a water molecule. An interesting feature of the metal ligation at this site is the use of carbonyl atoms from the C terminus of an α helix, utilizing the negative end of a helix dipole to interact favorably with the positively charged metal ion (Figure 3c). One of the regions stabilized by this second K^+ is directly involved in contacts with the α subunit through the sidechain of Asn183-β, which makes a hydrogen bond with the mainchain imino group of Gln369-α. The binding of K^+ at this site is likely to be important for the stability of the tetrameric E1b structure and agrees

well with the observed dependence of the enzyme activity on the concentration of potassium salts [12,28].

Subunit interactions

The interactions between the four subunits in the tetramer are extensive and consist of four different interfaces: first between the two α subunits (i.e. interface α–α′) burying 3215 Å^2 solvent-accessible surface area; second, the α–β interface burying an area of 1993 Å^2; third, the α–β′ interface burying 1890 Å^2; and fourth the β–β′ interface shielding 2239 Å^2 from the solvent. The α–α′ interaction is to a large extent formed by the intertwining N-terminal tails making the buried accessible surface area substantially more extensive than in the *P. putida* E1b structure (3215 Å^2 against 1915 Å^2). Overall, not less than 37% of the accessible surface of the α subunits and 46% of the β subunits is devoted to subunit interactions. This includes large hydrophobic patches that would be exposed prior to tetramer assembly within the mitochondrion. This is consistent with the requirement of chaperonins GroEL/GroES for the assembly of the human E1b heterotetramer in *Escherichia coli* [11] and *in vitro* [12].

Analysis of E1b mutations causing MSUD

There are currently 19 known MSUD-causing mutations that result in amino acid substitutions in the α and β subunits of human E1b (Table 1). With a few exceptions, mutations in the α subunit of E1b are often associated with the manifestation of the severe classic MSUD phenotype. The structure of human E1b can now be used to divide the group of mutations into three subgroups on the basis of the likely effect of each mutation on the structure and function of E1b. The first group includes mutations affecting cofactor binding, the second affects mainly hydrophobic cores and the third occurs at subunit interfaces (Table 1). Table 1 provides information regarding the residual BCKD complex activity, ethnic origin and the clinical phenotype of homozygous or compound heterozygous MSUD patients carrying the corresponding mutations. With some exceptions, cells from classic MSUD patients usually exhibit 0–2% of normal BCKD complex activity, whereas cells from intermediate patients show 2–30% of normal BCKD complex activity. The location of the corresponding mutations in the present structure can now be used to explain these observations.

Mutations affecting cofactor binding

The MSUD-causing mutations that can be most easily explained are substitutions of residues that are directly involved in binding the cofactor in the active site.

Asn222-α is one of the residues in the consensus sequence motif for this family of enzymes [29] and contributes to cofactor binding by ligating the associated magnesium ion (Figure 4a). Judged by the strict conservation of Asn222-α within the family, this residue has a critical function, one that can apparently not be performed by a serine residue because the N222S-α mutation causes MSUD. The explanation is probably that in wild-type human E1b, the sidechain carbonyl oxygen of Asn222-α is used for binding metal ions, whereas the hydroxyl group of the serine sidechain would be a poorer ligand to the magnesium ion. Moreover, the O^{γ} of a serine residue would not be able to extend as far towards the magnesium ion as the $O^{\delta 1}$ of the aspargine, which is another factor likely to decrease the affinity for magnesium at this site. This is reflected by a 10^3-fold increase in K_M for ThDP of Asn222S-α E1b compared with the wild-type enzyme (data not shown). The asparagine to serine mutation at position 222-α causes a severe, 'classic', case of MSUD.

The severe R114W-α and R220W-α MSUD mutations probably have quite similar effects on the cofactor-binding site. Both substitutions introduce a tryptophan residue in the place of an arginine residue directly involved in binding the cofactor through ionic interactions with its phosphate groups (Figure 4a). Besides loosing that capability, the bulky tryptophan residue probably prevents the protein from adopting a proper active-site conformation and could even prevent cofactor binding by steric hindrance. Given the major effects expected on ThDP binding, and possibly substrate binding due to these mutations, it is not surprising that this causes the classic severe type of MSUD.

One of the more intriguing mutations is T166M-α because it replaces a residue involved in ligation of K^+ at site 1. The probable abolishment of K^+ binding at this site by the T166M-α mutation is most probably caused by destabilization of the loop forming the K^+-binding site and preventing this loop from adopting the precise conformation required for cofactor binding that involves Ser162-α and Leu164-α at the tip of the loop (Figure 3b). The result is a classic severe type of MSUD.

Mutations affecting hydrophobic cores

A part of the hydrophobic core in the main domain of the α subunit is formed by the sidechains from helices 3 and 11 and strands a, d, e, h and i. Two mutations in the α subunit seem to disrupt this hydrophobic core (Figure 4b). Clearly, the large and charged sidechain of arginine introduced by the T265R-α mutation in helix 11 cannot be accommodated within this core in the current conformation. The site of the much less dramatic I281T-α mutation in strand i is very close to Thr265-α. The I281T-α substitution might seem subtle but a polar sidechain might not be easily tolerated in this exclusively hydrophobic environment. In the family of heterotetrameric E1 proteins (about 50 known sequences so far), only leucine, methionine, isoleucine and valine are found at this position. Patients heterozygous for I281T-α/Q145K-α show an intermediate clinical phenotype and the corresponding cells show low residual enzyme activity (Table 1). The

Table 1

Missense mutations in E1b subunits and clinical phenotypes of MSUD patients.

Amino acid substitution	Enzyme activity* (% normal)	Clinical phenotype	Ethnic origin	Reference
Group 1 — mutations affecting cofactor binding				
R114W-α	Unknown	Classic	Japanese	[52]
T166M-α	0	Classic	Thai	(JLC and DTC, unpublished observations)
R220W-α	0	Classic	Caucasian	[30]
N222S-α	1.3†	Classic	Caucasian	[53]
Group 2 — mutations affecting hydrophobic cores				
M64T-α	6§	Intermediate	Caucasian	(JLC and DTC, unpublished observations)
G204S-α	0	Classic	Caucasian	[30]
A208T-α	0	Classic	Japanese	[52]
T265R-α	0	Classic	Canadian	[30]
I281T-α	1.4‡	Intermediate	Japanese	[52]
Group 3 — mutations affecting subunit association				
N126Y-β	<1§	Classic	Unknown	[54]
Q145K-α	1.4‡	Intermediate	Japanese	[52]
H156R-β	0	Classic	Japanese	[52]
A209D-α	27‡	Intermediate	Caucasian	(JLC and DTC, unpublished observations)
A240P-α	0	Classic	Caucasian	[30]
G245R-α	4†	Intermediate	Hispanic-Mexican	[55]
R252H-α	27‡	Intermediate	Caucasian	(JLC and DTC, unpublished observations)
F364C-α	0	Classic	Hispanic-Mexican	[53]
Y368C-α	0	Intermediate	Caucasian	[55]
Y393N-α	0	Classic	Mennonite	[42] [31]

*Unless otherwise indicated, BCKD activity was measured using fibroblasts or lymphoblasts derived from MSUD patients and normal subjects. Intact-cell assays were carried out by incubating harvested cells with α-keto [1–^{14}C] isovalerate under isotonic conditions, and $^{14}CO_2$ released was measured [43]. The residual BCKD activity is expressed as a percentage of wild type. †E1b activity in the recombinant protein carrying the G245R-α or N222S-α mutation was measured by radiochemical E1b assays using α-keto [1–^{14}C] isovalerate as a substrate and 2,6-dichlorophenol indophenol as the electron acceptor [43]. The results were compared to that obtained with the wild-type E1b protein. ‡The residual BCKD activity was measured with intact MSUD cells that are compound-heterozygous for I281T-α/Q145K-α or A209D-α/R252H-α. Thus, the residual activity is expressed by one or both alleles in each patient. The other allele in the patient carrying the M64T-α mutation has not been identified. §The other allele in the patient carrying this mutation is a null-allele.

relatively mild form of the disease in this case corresponds well with the relatively small predicted effect of the mutations on protein structure and stability.

Two mutations that also seem to prevent the formation of an optimal hydrophobic core are G204S-α and A208T-α (not shown). In both cases, the mutation introduces a larger polar sidechain in an environment in which a smaller and non-polar residue is required.

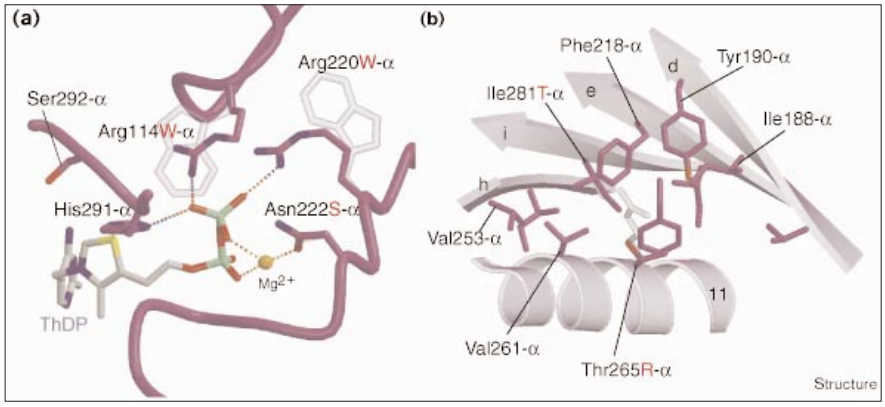
The M64T-α mutation in helix 2 appears to reveal a remarkably critical spot in a hydrophobic core in the α subunit at a location where four helices come together through hydrophobic interactions (not shown). The conservation of this methionine residue within the family of heterotetrameric E1b proteins suggests that it cannot be easily replaced by the polar and smaller sidechain of a threonine residue. Residual multienzyme complex activity in cells from patients carrying this mutation is significant and gives rise to an intermediate clinical phenotype (Table 1), in agreement with the expected limited effect of the methionine to threonine substitution at position 64-α on E1b stability.

Mutations affecting subunit interactions

The extensive subunit interactions in the heterotetrameric structure are undoubtedly crucial for maintaining the catalytically active structure of the enzyme, because the cofactor is bound at an α–β′ interface where residues from both subunits form the active site. There are a number of MSUD mutations that originate from their detrimental effect on subunit–subunit interaction surfaces (Table 1). Interestingly, all four different interfaces, α–α′, α–β, α–β′ and β–β′, located at entirely different positions throughout the heterotetramer, are effected by this class of mutations, as will be discussed later.

A209D-α, A240P-α, G245R-α and R252H-α all cluster at or close to the interface between the α subunits. This α–α′ interface in the interior of the tetramer is to a large extent formed by helix 10 in the two α subunits, which associate around the twofold axis of the tetramer. The small glycine residue at position 245-α is critical in order to allow the close interaction of these helices (Figure 5a). A substitution of glycine for arginine at this position by the G245R-α mutation would clearly prevent a tight association of the two α subunits in the manner observed in

Figure 4



MSUD mutations affecting cofactor binding and the hydrophobic core in human E1b. **(a)** The R114W-α, R220W-α and N222S-α mutations each affect residues that are directly involved in binding the cofactor phosphates and the associated magnesium ion. Asn222-α is a conserved residue in the superfamily of ThDP-utilizing enzymes and is a part of its sequence fingerprint motif ($GDG(X)_{22-30}NN$). Also shown is the primary phosphorylation site at Ser292-α [26]. **(b)** The sites of the T265R-α and I281T-α mutations appear to occur in the same hydrophobic cluster of the α subunit. These mutations are likely to prevent optimal packing in the hydrophobic core. This figure was made using MOLSCRIPT [48] and the Raster3D suite [49].

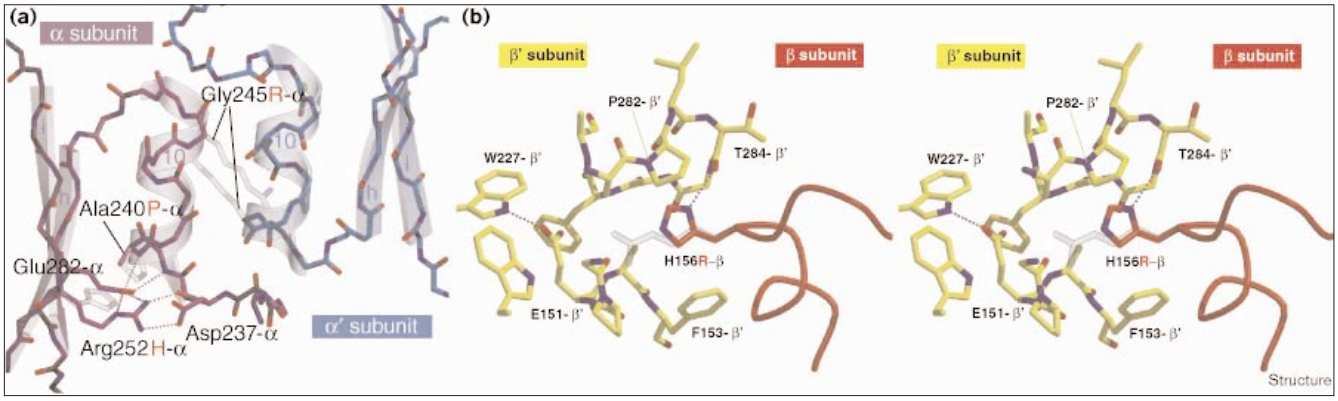
the crystal structure. It is rather surprising therefore that observations indicate that the mutant protein may nevertheless assemble in an imperfect way to produce a partially active enzyme [30]. E1b activity of a recombinant protein with the glycine to arginine mutation at position 245-α was measured to be approximately 4% compared with wild-type, and patients carrying this mutation show an intermediate clinical phenotype (Table 1).

On the basis of our structure, the two MSUD mutations A240P-α and R252H-α would both be predicted to disrupt a network of interactions formed between charged groups at the N-terminal end of helix 10 (Figure 5a). In the wild-type enzyme Arg252-α makes a salt bridge with the negatively charged sidechain of Asp237-α as well as with the sidechain of Glu282-α, which in turn interacts with the positive end of the dipole of helix 10 through a direct hydrogen bond with the amino group of Ala240-α. This extensive network of salt bridges and hydrogen bonds is probably critical for the precise positioning of helix 10 and consequently for the formation of the α–α′ interface. The R252H-α MSUD mutation introduces a shorter histidine sidechain that might not be able to maintain optimal bonds with both sidechains of Asp237-α and Glu282-α at the same time (Figure 5a). The A240P-α mutation, on the other hand, eliminates the free amino group at the beginning of helix 10 and interrupts the hydrogen bond between Glu282-α and the imino group of Ala240-α occurring in the wild-type enzyme. This mutation might also prevent formation of the bond between Glu282-α and the mainchain amino group at position Ile239-α by steric hindrance. The A240P-α mutation gives rise to a classic clinical phenotype with no residual multienzyme complex activity. In contrast, R252H-α together with A209D-α (also found at the α–α′ interface, not shown) are the least severe mutations known to cause MSUD. These two mutations occur in compound-heterozygous patients that display almost 30% residual multienzyme complex activity (Table 1).

His156-β is probably critical for the close interaction between the two β subunits. Like a 'knob in a hole', the sidechain of His156-β in one β subunit is tucked into an interface pocket formed by the other β subunit, packing against the sidechains of Phe153-β′ and Pro282-β′ (Figure 5b). In addition, His156-β makes a hydrogen bond with the mainchain carbonyl of Thr284-β′. Although the arginine sidechain, introduced by the H156R-β mutation at this position, could potentially have favorable interactions with Glu151-β′ at the bottom of the pocket this might lead to loss of a favorable interaction between Glu151-β′ and the sidechain of Trp227-β′ (Figure 5b). In addition, a specific hydrogen bond between the sidechain of His156-β and the mainchain carbonyl group of Thr284-β would be lost by the substitution. The interface pocket accommodating the His156-β sidechain might therefore not be able to accommodate the larger sidechain of an arginine residue in a satisfactory way and the histidine to arginine mutation at this position thereby prevents association of the β subunits and assembly of an active tetramer, causing a classic form of MSUD.

Another interface residue altered in some MSUD patients is Asn126-β, which appears to be involved in forming a network of hydrogen bonds that stabilizes the conformation of the polypeptide in two regions in the β subunit, around positions Asn126-β and Lys182-β (Figure 3c). These regions are critical for K^+ binding at site 2 (see above) and might also be important for subunit association. The sidechain of Asn126-β forms hydrogen bonds with the imino group of Gly128-β and with the sidechain of Arg118-β. The MSUD substitution N126Y-β would dramatically affect this region because of the larger size of the newly introduced sidechain. The favorable interaction between Arg118-β and the mainchain carbonyl of Lys182-β is likely to be disrupted with further serious consequences because the mainchain carbonyl groups of residues 181-β, 183-β as well as 128-β ligate the K^+ at site

Figure 5



MSUD mutations affecting subunit interactions. For clarity, chains are shown as polyglycine plus selected sidechains only. **(a)** Mutations close to the interface between the α subunits of the E1b tetramer. Mutations A240P-α and R252H-α are likely to prevent formation of the network of bonds between charged groups at the N-terminal end of helix 10. These interactions might be important for the precise positioning of the helix and hence for formation of the interface between the α subunits. The G245R-α mutation would introduce a large sidechain at a position where a small sidechain is required in order to allow the close approach of the interface helices α_{10} and α_{10}'. **(b)** Stereoview of the site of the H156R-β mutation at the interface between the β subunits. The precise fit of the histidine sidechain in the interface pocket formed by the opposite β subunit cannot be obtained with the larger sidechain of an arginine residue introduced by the MSUD mutation. This figure was made using MOLSCRIPT [48] and the Raster3D suite [49].

2. Interfering with this K^+ site has yet further implications because the sidechain of Asn183-β is in contact with the mainchain of the α subunit at position 369 adjacent to Tyr368-α (which is the site of another MSUD mutation; see below) at the α–β interface. The structural effects of the asparagine to tyrosine substitution at position 126-β correlate well with the observation that this mutation causes severe MSUD.

The 'Mennonite region'

There are three MSUD missense mutations that map onto the extended small C-terminal domain protruding from the bulk of the α subunit (Figure 6a). This domain has a different size in different E1 heterotetramers (Figure 2) but seems to function as support for the two β subunits extending the interaction between the α and β subunits. This small domain of each α subunit is in contact with both β subunits (Figure 6). The interface contacts are to a large extent hydrophobic in nature with a notable feature being an array of aromatic sidechains from the α subunit side packed against an array of aromatics from the β′ subunit (Figure 7). All three MSUD mutations in this part of the α subunit occur in the 'aromatic α array' and substitute an aromatic residue for a smaller and non-aromatic residue. This includes the frequent tyrosine to asparagine mutation at position 393-α found in the US Mennonite population [31,32], in which the prevalence of this mutation is 1 in 176 live births as a result of consanguinity [19,33].

Tyr393-α is packed in between two aromatic residues, Trp330-β′ and Phe324-β′, of the β′ subunit and forms two hydrogen bonds via its hydroxyl group with Asp328-β′ as well as His385-α (Figure 6b). The tyrosine to asparagine mutation at position 393-α introduces a much smaller and polar asparagine sidechain that would be unable to compensate for the extensive contact made by the native tyrosine sidechain.

The sidechain of Tyr368-α, another member of the aromatic α array affected by MSUD mutations, fits into a surface pocket on the β subunit (Figure 6a) formed by the sidechains of residues Asn183-β, Lys161-β and Pro259-β and the mainchain atoms of Gly159-β and Ile160-β. The hydroxyl group is then able to act as both hydrogen-bond acceptor and donor making hydrogen bonds to a mainchain carbonyl group (160-β) and a mainchain imino group (Gly159-β). These specific interactions would not be formed by the smaller cysteine sidechain resulting from the tyrosine to cysteine mutation at position 368-α.

Phe364-α, the third member of the aromatic array changed into a non-aromatic in certain MSUD patients, makes extensive hydrophobic interactions with the sidechains of Leu363-α, Leu375-α, Gln379-α and Leu382-α, which contribute to the compact hydrophobic core within the small C-terminal domain of the α subunit. In addition, these residues are also part of the interface to the β subunit where the sidechains of Leu362-α, Leu363-α and Phe364-α are tightly packed against the sidechain of Tyr313-β′ (Figure 6a). The phenylalanine to cysteine MSUD mutation at position 364-α is likely to affect α–β′ subunit interactions by causing defects in the subunit interface but might have more profound consequences for the folding of the small C-terminal domain itself.

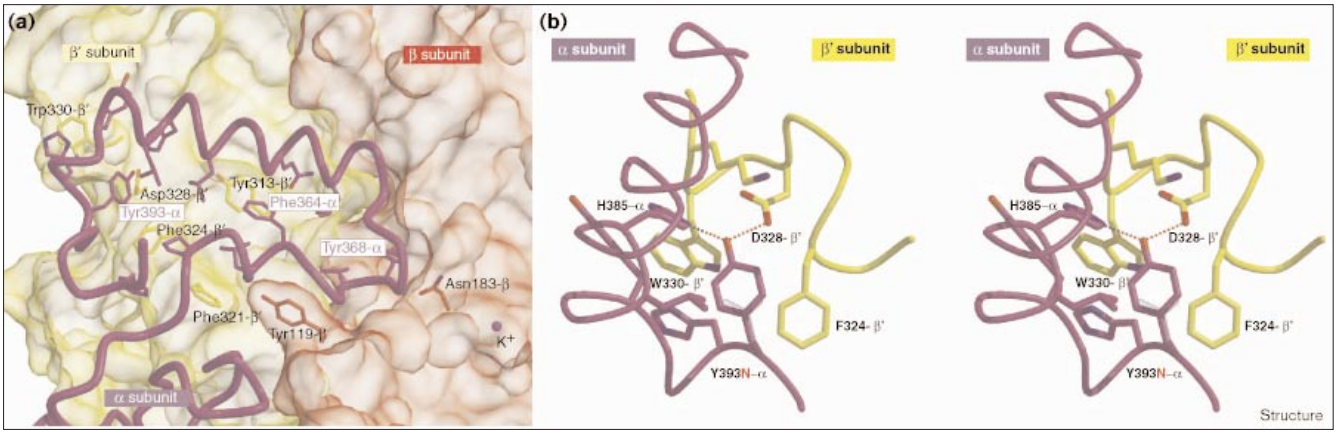
The aromatic to non-aromatic mutations in the small C-terminal domain of the α subunit are of particular interest because they have a dramatic effect on the structural integrity of the protein as a whole by impairing the process of tetramer assembly. We have shown previously that the folding and assembly of mammalian E1b heterotetramers is entirely dependent upon chaperonins GroEL/GroES either in *E. coli* [11] or *in vitro* [12]. Moreover, the chaperonin-mediated $\alpha_2\beta_2$ assembly proceeds through an αβ heterodimeric intermediate [30], which can be isolated as a relatively stable species from *E. coli* under certain growth conditions [12]. Our recent pulse-chase data shows that Y393N-α, F364C-α and Y368C-α mutations markedly reduce the assembly rate of the normal β subunit with the mutant α [30]. On the other hand, expression studies in *E. coli* indicate that in the presence of over-producing chaperonins GroEL/GroES, Y393N-α and F364C-α subunits are capable of assembly with the normal β subunit into stable heterodimers, but the latter are unable to convert to active heterotetramers [30]. Under similar conditions, the wild-type α subunit assembles with β subunits into active heterotetramers. As for the Y368C-α mutation, the mutant α subunit and the normal β appear to be capable of forming both heterotetramers and heterodimers [30].

From the structural analysis combined with the biochemical studies emerges a picture of a probable assembly pathway of the human E1b tetramer *in vivo*. The proposed assembly proceeds from single subunits, see Figure 7, via heterodimers αβ and α′β′ to the final heterotetramer. The Y393N-α mutation apparently blocks this assembly midway because it leads exclusively to formation of heterodimers. A single heterodimer of this type would correspond to association of the α subunit with the β subunit and not the β′ subunit because the α–β′ interaction is impaired by the Y393N-α mutation. In effect, the Y393N-α mutation blocks the final step of the assembly by preventing the interaction between the two heterodimers (Figure 7). The resulting heterodimers would be inactive because the active sites in which the cofactors are bound are formed in the last step of the pathway through assembly of the αβ and α′β′ dimers.

The F364C-α mutation can be predicted to have similar effects as the Y393N-α mutation has on tetramer assembly because it affects the same α–β′ and α′–β interfaces. This is also supported by the results of the chaperonin-assisted assembly studies [30].

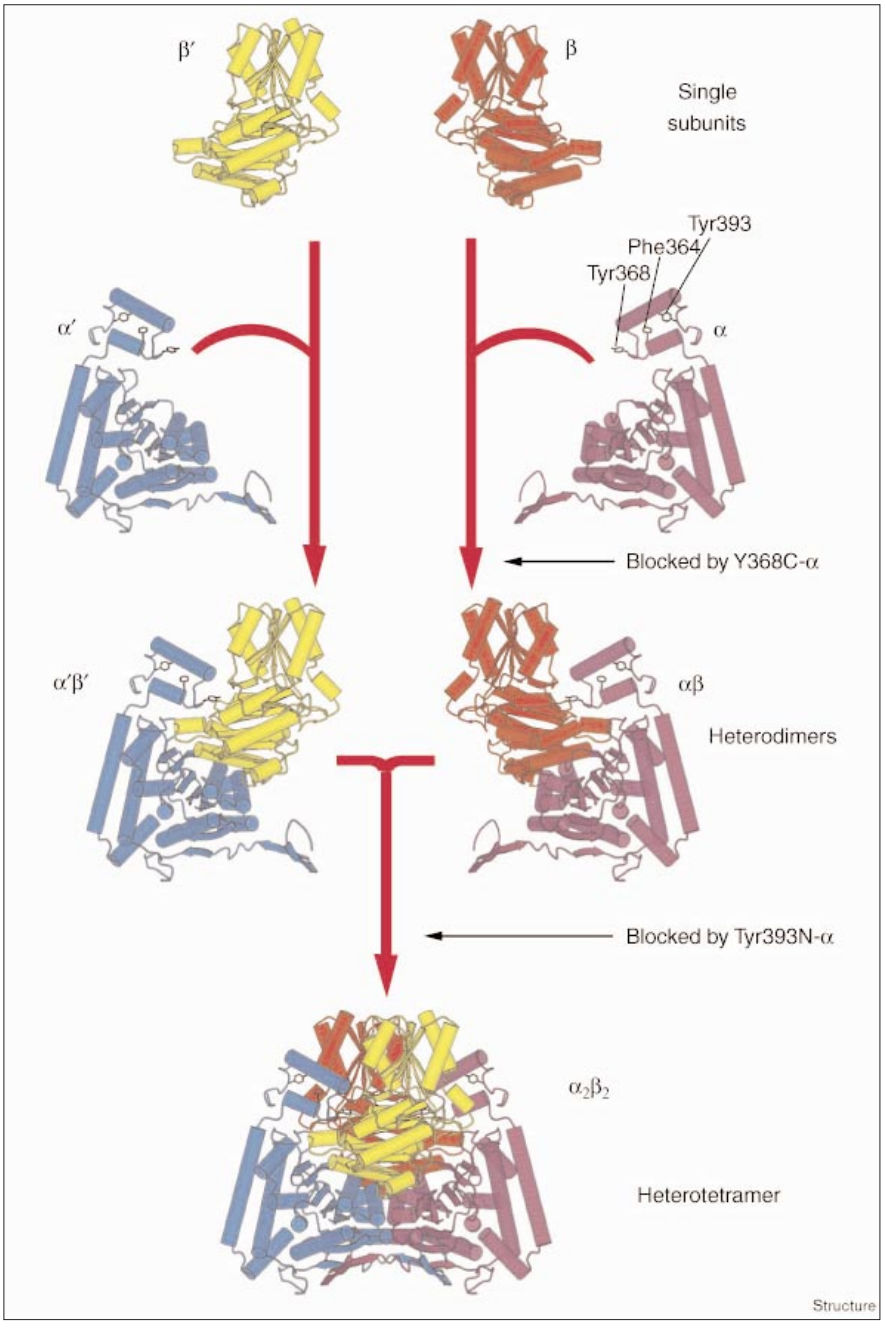
The Y368C-α mutation is distinctly different from the F364C-α and Y393N-α mutations because it affects different interfaces (α–β and α′–β′) and would seem to impair the first step in the assembly pathway in Figure 7. It also appears to be less severe, however, because it does not completely prevent assembly of the tetramer.

Figure 6



Mutations affecting subunit interactions between α and β subunits. **(a)** Mutations in the small C-terminal domain of the α subunit at the interface with the two β subunits. The C-terminal domain is shown with a coil representation of the backbone with selected sidechains. The β subunits are shown as semi-transparent surfaces, in yellow for the β′ subunit and in red for the β subunit. Some sidechains in the β subunits that are forming the interface to the α subunit are shown beneath the surface. Note that the α subunit makes contacts with both β subunits and also note the prevalence of aromatic residues at the subunit interface. The sites of the three MSUD mutations in this region of the α subunit, F364C-α, Y368C-α and Y393N-α, are highlighted. All three mutations are substitutions of aromatic residues involved in the interface to a β subunit, Tyr368-α is in contact with the β subunit (α–β interface) whereas Phe364-α and Tyr393-α contact the β′ subunit (α–β′ interface). **(b)** Close up stereoview of Tyr393-α (the site of the 'Mennonite' mutation) and its vicinity at the interface between the small C-terminal domain of the α subunits and the β′ subunit. Tyr393-α is packed in between the sidechains of Phe324-β′ and Trp330-β′ with its hydroxyl group making hydrogen bonds with the sidechains of His385-α and Asp328-β′. The Y393N-α mutation is the most frequently observed mutation that causes MSUD [31,42]. This figure was created using GRASP [51], MOLSCRIPT [48] and the Raster3D suite [49].

Figure 7

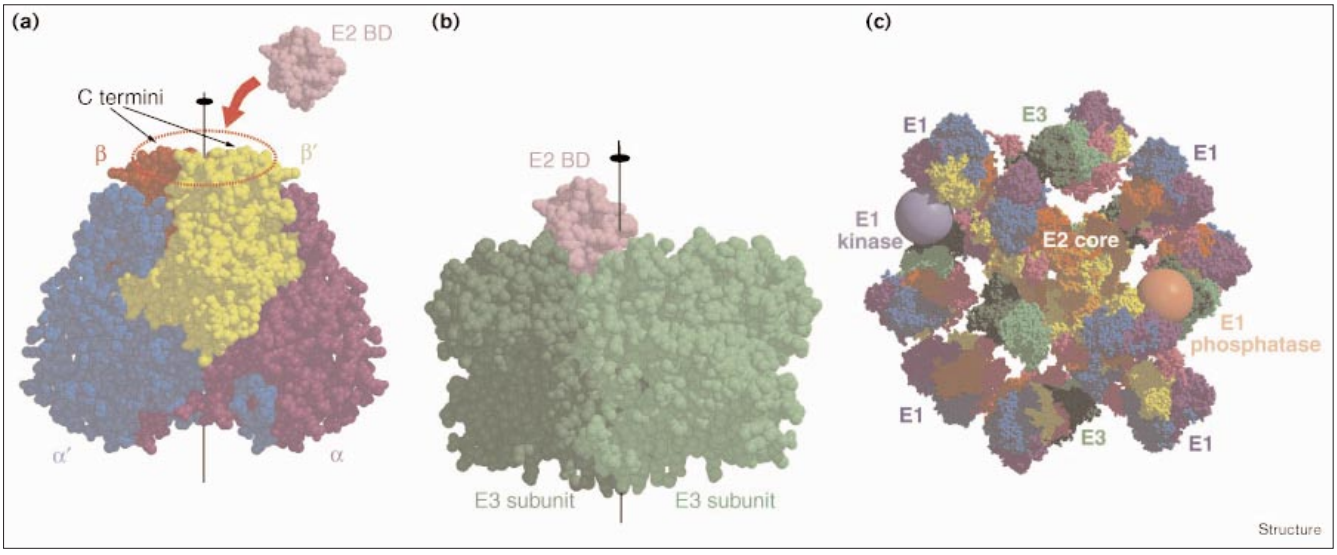


The proposed assembly pathway of the human E1b tetramer and the possible effects of mutations F364C-α, Y368C-α and Y393N-α on the assembly. The tetramer probably assembles through the formation of intermediate heterodimers (see text). Assembly of mutant proteins carrying the Y393N-α mutation results exclusively in the formation of heterodimers that can be predicted to correspond to αβ and α′β′ dimers because the α–β′ and α′–β interaction is impaired by the Y393N-α mutation. The same is probably the case for the F364C-α mutation. The Y368C-α mutation would have distinctly different effects on tetramer assembly because it occurs at different interfaces (α–β and α′–β′) compared with the other two mutations. The figure was created using MOLSCRIPT [48].

Interaction with other components in the multienzyme complex

The human E1b component interacts with a number of other components within the multienzyme complex: the binding domain and lipoyl domain of E2b as well as the specific protein kinase and phosphatase that control the activity of E1 through (de)phosphorylation. It has been shown previously that an E1 tetramer binds only one copy of the E2 binding domain with the binding site located on the β subunits [34,35]. This suggests that the binding site is located at or close to the twofold axis on the β subunits. Our reconstitution experiments employing different constructs of human E1b now provide more direct evidence supporting binding at this location of the tetramer. The results indicate that, under the same conditions, an E1b tetramer with an N-terminal His_6-tag on the α subunit readily forms a complex with the E2b 24-mer, whereas an E1b heterotetramer that carries a C-terminal His_6-tag on

Figure 8



Assembly of the BCKD complex. **(a)** Interaction between E1 and the binding domain of E2. Human E1b is shown on the left with the proposed binding site of the E2 binding domain at the top of the tetramer close to the C termini of the β subunits. Recombinant protein carrying a His-tag at the C termini of the β subunits is unable to bind the E2 binding domain. **(b)** The proposed position of the binding domain at the twofold axis of the E1 tetramer is analogous to the binding of the E2 binding domain at the twofold axis of the E3 dimer, as seen in the structure of *Bacillus stearothermophilus* E3 in complex with the E2 binding domain [36]. **(c)** Hypothetical composite model of a branched-chain dehydrogenase multienzyme complex (described in the text).

the β subunit cannot bind to E2b (data not shown). The latter construct was used to express protein for the structure determination and although the His_6 tags are invisible in our electron-density maps there is visible density up to the last residue preceding the His_6 tag at the C termini of the β subunits. It is interesting that the proposed binding of the E2 binding domain to E1 (Figure 8a) is analogous to the binding of a single copy of this domain to the E3 dimer as seen previously in the structure of *Bacillus stearothermophilus* E3 in complex with the E2 binding domain (Figure 8b) [36].

Given our ability to model approximate interactions of different multienzyme components, we can now assemble the building blocks together in an overall picture of the human E1b multienzyme complex analogs because examples of all components have been structurally characterized [4,16,37,38]. Combining this with a simple model of the mutual arrangement of components in multienzyme complexes that emerged from early electron microscopy work [39], we can arrive at the hypothetical model of a BCKD complex shown in Figure 8c. The model includes an E2 core with 24 subunits arranged as a cube with 432 (octahedral) symmetry as seen for the first time for the core of *Azotobacter vinelandii* pyruvate dehydrogenase [4]. An E1 tetramer is placed above each of the 12 edges of the cubic core with the twofold axis of the E1 molecule coinciding with the twofold axes of the core and the β subunits with the associated E2 binding domains oriented towards the core. In an analogous way an E3 dimer in complex with an E2 binding domain is placed at the fourfold axis above each of the six faces of the cubic core. The lipoyl domains are placed in the space between the E1 and E3 molecules allowing free access to the respective active sites. The resulting complex has a molecular mass close to four million Da and an overall diameter of approximately 400 Å.

As intact α-ketoacid dehydrogenase multienzyme complexes resist crystallization, our attempts to obtain a high resolution overall picture of these macromolecular assemblies focused on structural studies of the individual components providing pieces to a composite picture as the one outlined above. Further advances using electron microscopy, nuclear magnetic resonance (NMR), crystallography and biochemical techniques will continue to provide more refined models. These complexes are an elegant combination of a highly symmetric core surrounded by flexible 'arms' to which dozens of large additional enzymes are attached with variations in stoichiometry. One might particularly appreciate the substrate CoAs, which have to navigate around the swirling peripheral proteins to enter the hollow core of the complex [40] in order to retrieve a precious acyl group, and then to make the journey out of the complex to assure the proper delivery of this group to other critical, yet simpler, components in the mitochondrial matrix.

Biological implications

Mutations in components of the α-ketoacid dehydrogenase multienzyme complex can lead to potentially severe disorders in humans, including maple syrup urine disease (MSUD). We determined the crystal structure of a component (the E1b component) of the branched-chain α-ketoacid dehydrogenase multienzyme complex from humans.

The structure of human 2-oxoisovalerate dehydrogenase has enabled us to provide a detailed explanation for all the 19 point mutations of α-ketoacid dehydrogenase (E1b) that are known to be responsible for MSUD in human patients. In essentially all instances the severity of the disease corresponds well with the predicted effect of the mutation on the structure, stability, and cofactor or K^+ binding by the enzyme. In the severe cases cofactor binding is essentially prohibited, or proper assembly made impossible. In these cases our structural studies provide profound insight, but the structure does not easily lend itself as a platform for the development of small therapeutic agents that assist in treatment for the disease.

For some of the milder mutations, however, the human E1b structure presented here might lend itself perhaps to the discovery of compounds that could be therapeutically useful for several of these less severe forms of MSUD, even if these are caused by entirely different mutations. The principle of the proposed approach is that mutations such as M64T-α, I281T-α and R252H-α probably lead to a significant but not dramatic decrease in the stability of E1b. E1b heterotetramers with these mild MSUD mutations might be recovered by interactions with small molecules. Because tightly binding small molecules are able to increase the stability of proteins significantly during or after folding, the discovery of low molecular weight compounds with high affinity for human E1b that do not adversely affect the catalytic machinery of the enzyme, could be a starting point for the development of useful therapeutic agents. In particular, patients with the intermittent forms of the disease, who, therefore, would not require a life long use of therapeutic compounds, might benefit from such an approach. This approach might be generally applicable to other genetic diseases in which loss of protein stability causes milder forms of the disease (for instance, several forms of β-galactosialidosis are caused by relatively mild mutations in the human protective protein [41]).

Materials and methods

Construction of a vector for the expression of human E1 with a His_6 tag at the C terminus of the β subunit

To facilitate vector construction, an *Nco*I site was engineered into the N terminus of the mature human α sequence using polymerase chain reaction (PCR) amplification. This resulted in the incorporation of an extra glycine immediately upstream of the N-terminal serine residue. An *Nco*I/*Eco*RI fragment encoding the N-terminal sequence of the α subunit was restricted from the PCR product. An *Eco*RI/*Bam*HI fragment containing the downstream sequence of the subunit was excised from the previously described EBO-bhE1a vector [42]. To generate the insert for the mature β subunit, a sequence coding for the His_6 tag was incorporated immediately 3′ to the C-terminal tyrosine codon. The β-His_6 sequence was inserted into the *Nco*I/*Eco*RI-digested pKK233-2 vector from Amersham Pharmacia (Piscataway, NJ). The *Nco*I/*Eco*RI and the *Eco*RI/*Bam*HI fragments from the a subunit along with the *Bam*HI/*Sca*I gene cartridge containing the pTac promoter and the β-His_6 sequence were ligated into a *Nco*I/*Sca*I-digested pTrcHisB vector from Invitrogen (Carlsbad, CA). The resultant vector p-hE1-His_6 encodes an untagged α sequence and the C-terminal His_6-tagged β sequence of human E1b.

Expression and purification of recombinant human E1b protein

The p-hE1-His_6 plasmid and pGroESL plasmid overproducing chaperonins GroEL/ES were cotransformed into $GroES^{ts}$ CG-712 *E. coli*, and cells were induced with IPTG for hE1b expression as described previously [30,34]. The overexpression of chaperonins GroEL/ES in the host cells was essential for efficient expression of hE1b. The recombinant human E1b-His_6 protein was extracted from cell lysates with Ni-NTA resin (Qiagen, Valencia, CA), and eluted with a 15–250 mM imidazole gradient in a buffer containing 50 mM KPi, pH 7.5, 500 mM NaCl, 10 mM β-mercaptoethanol, 1 mM benzamidine, 1 mM PMSF and 5% glycerol. The Ni-NTA extracted protein was applied to a Resource Q FPLC column and eluted with a 50–300 mM KCl gradient in the above-mentioned buffer minus NaCl, and with 2 mM DTT replacing β-mercaptoethanol. The eluted protein was further purified on a Superdex 200 FPLC gel-filtration column equilibrated with 50 mM Hepes, pH 7.5, 250 mM KCl, 2 mM DTT and 5% glycerol. The purified human E1b-His_6 protein was concentrated on Ultra free-15 concentrator (Millipore, Bedford, MA) to a concentration of 20 mg/ml. The protein was homogenous as judged using SDS-PAGE and Coomassie blue staining. The final specific activity of the purified E1b-His_6 protein was 79 mU/mg as assayed radiochemically with 2,6-dichlorophenol indophenol as an electron acceptor [43].

Crystallization

After purification, the protein was kept in 0.1 M MES pH 7.5, 0.25 M KCl, 10% glycerol, 1 mM ThDP, 10 mM $MgCl_2$ and stored at −80°. The protein was crystallized using sitting drops with 35–40% ethanol, 0–2% PEG8000, 0.1 M MES pH 7.5 as a reservoir solution. Upon setting up crystallization drops, the protein solution, containing about 18 mg/ml protein, 0.1 M MES pH 7.5, 0.25 M KCl, 10% glycerol, 1 mM ThDP, 10 mM $MgCl_2$ and 1 mM inhibitor (either 2-chloroisocaproate or clofibrate), was usually mixed with the reservoir solution in a ratio between 1:0.1 and 1:1 protein solution:reservoir solution. In some cases the protein solution was not mixed with reservoir solution and crystallization proceeded through free diffusion of ethanol from the reservoir to the sitting drop. In either case, crystallization was rapid and the crystals reached their final size in one day. The crystals grew as long thin rods up to 20 mm in thickness. Attempts to cryoprotect crystals using artificial cryo solutions were often not very successful and severely reduced quality of diffraction. Eventually, the crystals could be frozen satisfactorily using cryo solution made by mixing reservoir solution and 100% glycerol in a ratio of 4 to 1. Because of continuous evaporation of ethanol, care was taken to rapidly remove crystals from the crystallization drop to the cryo solution and then freeze them either in liquid nitrogen for storage or directly in a nitrogen stream during crystal mounting on a detector. Crystals left in crystallization drops beyond a few weeks would deteriorate ('rubberize') and show no longer a diffraction pattern when exposed to the X-ray beam.

Data collection and structure determination

Data used for structure solution were collected using synchrotron sources at ESRF beamline ID02 (data set 1, co-crystallization with α-chloroisocaproate) and APS beamline 19-ID (data 2, co-crystallization

Table 2

Data collection and refinement statistics.

Data statistics	
Space group	$P3_121$
Cell dimensions (Å)	a = b = 143.8, c = 69.2
Resolution range (Å)	50.0–2.7
Number of observations	306,699
Unique reflections	22,910
Completeness (%)	99.9
R_{sum} (%)	14
Refinement statistics	
Resolution (Å)	50.0–2.7
Number of atoms	5671
Number of reflections	22,858
R/R_{free} (%)	22.4/27.9
Rmsd	
bond length (Å)	0.008
bond angles (°)	1.4
dihedrals (°)	22.5
improper (°)	0.92
B average (Å^2)	36.1
Ramachandran plot	
favored (%)	83.3
allowed (%)	15.7
generously allowed (%)	0.7
disallowed (%)	0.3

with clofibrate). Data collected from selenomethionine crystals suffered from crystal decay, despite cryo cooling, and did not allow successful phasing using MAD methods (unpublished observations). The structure could eventually be determined using molecular replacement methods (MR) using the previously determined structure of *P. putida* E1b [16] as a search model. An initial MR [44] solution was found using data set 1 but improved maps were obtained using data set 2 at 2.7 Å resolution. Iterative model building using the program O [45] and refinement, using X-PLOR [46] and at later stages using CNS with simulated annealing torsion angle refinement using maximum-likelihood target function [47], resulted in a model of residues 7–301 and 314–400 (out of 400) in the α subunit and residues 17–342 (out of 342 + His-tag) in the β subunit (Table 2). Residues missing in the current model were disordered and poorly or not visible in density maps. The asymmetric unit of the final model contains one α subunit, one β subunit, 48 water molecules, one ThDP cofactor molecule, one Mg^+ and two K^+.

Accession numbers

The coordinates have been deposited in the Protein Data Bank with accession code 1DTW.

Acknowledgements

We are thankful to other members of the Biomolecular Structure Center, University of Washington, for their support. We thank Francis Athappilly for preparing Figure 4a. A postdoctoral grant to AÆ from the Swedish Foundation for International Cooperation in Research and Higher Education (STINT) is gratefully acknowledged. WGJH acknowledges a major equipment grant form the Murdock Charitable Trust to the Biomolecular Structure Center. This study was also supported by grant DK-26758 from the NIH, grant I-1286 from the Welch Foundation, and grant 95G-074R from the American Heart Association, Texas Affiliate. Use of the Advanced Photon Source was supported by the U.S. Department of Energy, Basic Energy Sciences, Office of Energy Research, under Contract No. W-31-109-ENG-38. Use of the Argonne National Laboratory Structural Biology Center beamlines at the Advanced Photon Source, was supported by the U. S. W-31-109-ENG-38. We are grateful for the assistance of Stephan L Ginell, Randy Alkire and Frank Rotella in collecting the data at the SBC beamline. We also would like to thank Julien Lescar and Bjarne Rasmussen for their help at beamline ID02B at the European Synchrotron Radiation Facility.

References

1. Reed, L.J., Damuni, Z. & Merryfield, M.L. (1985). Regulation of mammalian pyruvate and branched-chain α-keto acid dehydrogenase complexes by phosphorylation–dephosphorylation. *Curr. Topics Cell Regul.* **27**, 41-49.
2. Reed, L.J. (1974). Multienzyme complexes. *Acc. Chem. Res.* **7**, 40-46.
3. Perham, R.N. (1991). Domains, motifs, and linkers in 2-oxo acid dehydrogenase multienzyme complexes: a paradigm in the design of a multifunctional protein. *Biochemistry* **30**, 8501-8512.
4. Mattevi, A., *et al.,* & Hol, W.G.J. (1992). Atomic structure of the cubic core of the pyruvate dehydrogenase multi-enzyme complex. *Science* **255**, 1544-1550.
5. Izard, T., *et al.,* & Hol, W.G.J. (1999). Principles of quasi-equivalence and Euclidean geometry govern the assembly of cubic and dodecahedral cores of pyruvate dehydrogenase complexes. *Proc. Natl Acad. Sci. USA* **96**, 1240-1245.
6. Chuang, D.T. & Shih, V.E. (1995). Disorders of branched-chain amino and α-ketoacid metabolism. In *The Metabolic and Molecular Basis of Inherited Disease* 7th edition. (Scriver, C.R., Beaudet, A.L., Sly, W.S. & Valle, D., eds), pp. 1239-1277, McGraw-Hill, New York.
7. Wynn, R.M., Davie, J.R., Meng, M. & Chuang, D.T. (1996). Structure, function and assembly of mammalian branched-chain α-ketoacid dehydrogenase complex. In *Alpha-Keto Acid Dehydrogenase Complexes* .(Patel, M.S., Roche, T.E. & Harris, R.A. eds), pp. 101-118, Birkhäuser Verlag, Basel.
8. Shimomura, Y., Nanaumi, N., Suzuki, M., Popov, K.M. & Harris, R.A. (1990). Purification and partial characterization of branched-chain α-ketoacid dehydrogenase kinase from rat liver and rat heart. *Arch. Biochem. Biophys.* **2**, 293-299.
9. Lee, H.Y., Hall, T.B., Kee, S.M., Tung, H.Y.L. & Reed, L.J. (1991). Purification and properties of branched-chain α-keto acid dehydrogenase kinase from bovine kidney. *Biofactors* **2**, 109-112.
10. Damuni, Z., Merryfield, M.L., Humphreys, J.S. & Reed, L.J. (1984). Purification and properties of branched-chain α-ketoacid dehydrogenase phosphatase from bovine kidney. *Proc. Natl Acad. Sci. USA* **81**, 4335-4338.
11. Wynn, R.M., Davie, J.R., Cox, R.P. & Chuang, D.T. (1992). Chaperonins GroEL and GroES promote assembly of heterotetramers ($\alpha_2\beta_2$) of mammalian mitochondrial branched-chain α-keto acid decarboxylase in *E. coli. J. Biol. Chem.* **267**, 12400-12403.
12. Chuang, J.L., Wynn, R.M., Song, J.L. & Chuang, D.T. (1999). GroEL/GroES-dependent reconsititution of $\alpha_2\beta_2$, tetramers of human mitochondrial branched chain α-ketoacid decarboxylase – obligatory interaction of the chaperonins with an alpha beta dimeric intermediate. *J. Biol. Chem.* **274**, 10395-10404.
13. Dancis, J., Hutzler, J. & Levits, M. (1960). Metabolism of the white blood cells in maple-syrup-urine disease. *Biochim. Biophys. Acta* **43**, 343-345.
14. Menkes, J.H., Hurst, P.L. & Craig, J.M. (1954). A new syndrome: Progressive familial infantile cerebral dysfunction associated with an unusual urinary substance. *Pediatrics* **14**, 462-466.
15. Taylor, J., Robinson, B.H. & Serwood, W.G. (1978). A defect in branched-chain amino acid metabolism in a patient with congenital lactic acidosis due to dihydrolipoyl dehydrogenase deficiency. *Pediatr. Res.* **12**, 60-62.
16. Ævarsson, A., Seger, R., Turley, S., Sokatch, J.R. & Hol, W.G.J. (1999). Crystal structure of 2-oxoisovalerate dehydrogenase and the architecture of 2-oxo acid dehydrogenase multienzyme complexes. *Nat. Struct. Biol.* **6**, 785-792.
17. Nikkola, M., Lindqvist, Y. & Schneider, G. (1994). Refined structure of transketolase from *Saccharomyces cerevisiae* at 2.0 Å resolution. *J. Mol. Biol.* **238**, 387-404.
18. Sundstrom, M., Lindqvist, Y., Schneider, G., Hellman, U. & Ronne, H. (1993). Yeast *TKL1* gene encodes a transketolase that is required for efficient glycolysis and biosynthesis of aromatic amino acids. *J. Biol. Chem.* **268**, 24346-24352.
19. Marshall, L. & DiGeorge, A. (1981). Maple syrup urine disease in the old order Mennonites. *Am. J. Hum. Genet.* **33**, 139A.
20. Hasson, M.S., *et al.,* & Ringe, D. (1998). The crystal structure of benzoylformate decarboxylase at 1.6 Å resolution: diversity of catalytic residues in thiamin diphosphate-dependent enzymes. *Biochemistry* **37**, 9918-9930.
21. Muller, Y.A., Lindqvist, Y., Furey, W., Schulz, G.E., Jordan, F. &

Schneider, G. (1993). A thiamin diphosphate binding fold revealed by comparison of the crystal structures of transketolase, pyruvate oxidase and pyruvate decarboxylase. *Structure* **1**, 95-103.
22. Chabrière, E., Charon, M.-H., Volbeda, A., Pieulle, L., Hatchikian, E.C. & Fontecilla-Camps, J.-C. (1999). Crystal structures of the key anaerobic enzyme pyruvate: ferredoxin oxidoreductase, free and in complex with pyruvate. *Nat. Struct. Biol.* **6**, 182-190.
23. Kern, D., *et al.*, & Hübner, G. (1997). How thiamine diphosphate is activated in enzymes. *Science* **275**, 67-70.
24. Schellenberger, A. (1998). Sixty years of thiamin diphosphate biochemistry. *Biochim. Biophys. Acta* **1385**, 177-186.
25. Schneider, G. & Lindqvist, Y. (1998). Crystallography and mutagenesis of transketolase: mechanistic implications for enzymatic thiamin catalysis. *Biochim. Biophys. Acta* **1385**, 387-398.
26. Zhao, Y., *et al.*, & Harris, R.A. (1994). Site-directed mutagenesis of phosphorylation sites of the branched chain α-ketoacid dehydrogenase complex. *J. Biol. Chem.* **269**, 18583-18587.
27. Hawes, J.W., Schnepf, R.J., Jenkins, A.E., Shimomura, Y., Popov, K.M. & Harris, R.A. (1995). Roles of amino acid residues surrounding phosphorylation site 1 of branched-chain α-ketoacid dehydrogenase (BCKDH) in catalysis and phosphorylation site recognition by BCKDH kinase. *J. Biol. Chem.* **270**, 31071-31076.
28. Shimomura, Y., Kuntz, M.J., Suzuki, M., Ozawa, T. & Harris, R.A. (1988). Monovalent cations and inorganic phosphate alter branched-chain α-ketoacid dehydrogenase-kinase activity and inhibitor sensitivity. *Arch. Biochem. Biophys.* **266**, 210-218.
29. Hawkins, C.F., Borges, A. & Perham, R.N. (1989). A common structural motif in thiamine pyrophosphate-binding enzymes. *FEBS Lett.* **255**, 77-82.
30. Wynn, R.M., Daive, J.R., Chuang, J.L., Cote, C.D. & Chuang, D.T. (1998). Impaired assembly of E1 decarboxylase of the branched chain α-ketoacid dehydrogenase complex in type IA maple syrup urine diesase. *J. Biol. Chem.* **273**, 13110-13118.
31. Matsuda, I., *et al.,* & Harada, H. (1990). A T-to-A substitution in the E1α subunit gene of the branched-chain α-ketoacid dehydrogenase complex in two cell lines derived from Mennonite maple syrup urine disease patients. *Biochem. Biophys. Res. Commun.* **172**, 451-464.
32. Fisher, C.R., Chuang, J.L., Cox, R.P., Fisher, C.W., Star, R.A. & Chuang, D.T. (1991). Maple syrup urine disease in Mennonites. Evidence that the Y393N mutations in E1α impedes assembly of the E1 components of branched-chain α-keto acid dehydrogrenase complex. *J. Clin. Invest.* **88**, 1034-1037.
33. Chuang, D.T. (1998). Maple syrup urine disease: it has come a long way. *J. Pediatr.* **132**, S17-S23.
34. Wynn, R.M., *et al.*, & Chuang, D.T. (1992). Cloning and expression in *Escherichia coli* of mature E1β subunit of bovine mitochondrial branced-chain α-ketoacid dehydrogenase complex. Mapping of the E1β-binding region on E2. *J. Biol. Chem.* 267, 1881-1887.
35. Lessard, I.A.D., Fuller, C. & Perham, R.N. (1996). Competitive interaction of component enzymes with the peripheral subunit-binding domain of the pyruvate dehydrogenase multienzyme complex of *Bacillus stearothermophilus*: kinetic analysis using surface plasmon resonance detection. *Biochemistry* **35**, 16863-16870.
36. Mande, S.S., Sarfaty, S., Allen, M.D., Perham, R.N. & Hol, W.G.J. (1996). Protein–protein interactions in the pyruvate dehydrogenase multi-enzyme complex: dihydrolipoamide dehydrogenase complexed with the binding domain of the dihydrolipoamide acetyltransferase. *Structure* **4**, 277-286.
37. Mattevi, A., Schierbeek, A.J., Oblomova-Teplyakova, G. & Hol, W.G.J. (1990). The three-dimensional crystal structure of lipoamide dehydrogenase from *Azotobacter vinelandii* at 2.2 Å resolution. In *Flavoenzymes and Flavoproteins* (Curti, A., Romchi, B. & Zanetti, C., eds), pp. 549-556. W. de Gryter & Co., Berlin.
38. Perham, R.N. (1996). Interaction of protein domains in the assembly and mechanism of 2-oxo acid dehydrogenase multienzyme complexes. In *Alpha-Keto Acid Dehydrogenase Complexes.* (Patel, M.S., Roche, T.E. & Harris, R.A. eds), pp. 1-15. Birkhäuser Verlag, Basel.
39. Oliver, R.M. & Reed, L.J. (1982). Multienzyme complexes. In *Electron Microscopy of Proteins.* (Harris, J.R., ed.), Vol. 2, pp. 1-48. Academic Press, London, UK.
40. Mattevi, A., Obmolova, G., Kalk, K.H., Teplyakov, A. & Hol, W.G.J. (1993). Crystallographic analysis of substrate binding and catalysis in dihydrolipoyl transacetylase (E2p). *Biochemistry* **32**, 3887-3901.
41. Rudenko, G., Bonten, E., Hol, W.G.J. & d'Azzo, A. (1998). The atomic model of the human protective protein/cathepsin A suggests a structural basis for galactosialidosis. *Proc. Natl Acad. Sci.* **95**, 621-625.
42. Fisher, C.R., Fisher, C.W., Chuang, D.T. & Cox, R.P. (1991). Occurrence of a Tyr393 Asn (Y393N) mutation in the E1a gene of the branched-chain α-ketoacid dehydrogenase complex in maple syrup urine disease patients from a Mennonite population. *Am. J. Hum. Genet.* **49**, 429-434.
43. Chuang, D.T. & Cox, R.P. (1988). Enzyme assays with mutant cell lines of maple syrup urine disease. *Methods Enzymol.* **166**, 135-146.
44. Navaza, J. (1994). AMoRe: an automated package for molecular replacement. *Acta Crystallogr. A* **50**, 157-163.
45. Jones, T.A., Zou, J.Y., Cowan, S.W. & Kjeldgaard, M. (1991). Improved methods for binding protein models in electron density maps and the location of errors in these models. *Acta Crystallogr. A* **47**, 110-119.
46. Brünger, A.T., Krukowski, A. & Erickson, J.W. (1990). Slow-cooling protocols for crystallographic refinement by simulated annealing. *Acta Crystallogr. A* **46**, 585-593.
47. Brünger, A.T., *et al.,* & Warren, G.L. (1998). Crystallography and NMR system: a new software suite for macromolecular structure determination. *Acta Crystallogr. D* **54**, 905-921.
48. Kraulis, P.J. (1991). MOLSCRIPT: a program to produce both detailed and schematic plots of protein structures. *J. Appl. Crystallogr.* **24**, 946-950.
49. Merritt, E.A. & Murphy, M.E.P. (1994). Raster3D version 2.0. A program for photorealistic molecular graphics. *Acta Crystallogr. D* **50**, 869-873.
50. Wallace, A.C., Laskowski, R.A. & Thornton, J.M. (1995). LIGPLOT: a program to generate schematic diagrams of protein–ligand interactions. *Protein Eng.* **8**, 127-134.
51. Nicholls, A., Shar, K.A. & Honig, B. (1991). Protein folding and association: insights from the interfacial and thermodynamic properties of hydrocarbons. *Proteins Struct. Funct. Gen.* **11**, 282-296.
52. Nobukuni, Y., *et al.*, & Matsuda, I. (1993). Heterogeneity of mutations in maple syrup urine disease (MSUD): screening and identification of affected E1α and E1β subunits of the branched-chain α-keto-acid dehydrogenase multienzyme complex. *Biochim. Biophys. Acta.* **1225**, 64-70.
53. Chuang, J.L., Davie, J.R., Chinsky, J.M., Wynn, R.M., Cox, R.P. & Chuang, D.T. (1995). Molecular and biochemical basis of intermediate maple syrup urine disease: occurrence of homozygous G245R and F364C mutations at the E1α locus of Hispanic-Mexican patients. *J. Clin. Invest.* **95**, 954-963.
54. McConnell, B.B., Burkholder, B. & Danner, D.J. (1997). Two new mutations in the human E1b subunit of branched chain α-ketoacid dehydrogenase associated with maple syrup urine disease. *Biochim. Biophys. Acta.* **1361**, 263-271.
55. Chuang, J.L., Fisher, C.R., Cox, R.P. & Chuang, D.T. (1994). Molecular basis of maple syrup urine disease: novel mutations at the E1α locus that impairs E1 ($\alpha_2\beta_2$) assembly or decrease steady-state E1α mRNA levels of branched-chain α-keto acid dehydrogenase complex. *Am. J. Hum. Genet.* **55**, 297-304.

Because *Structure with Folding & Design* operates a 'Continuous Publication System' for Research Papers, this paper has been published on the internet before being printed (accessed from http://biomednet.com/cbiology/str). For further information, see the explanation on the contents page.