

3D Structure and Significance of the GΦXXG Helix Packing Motif in Tetramers of the E1β Subunit of Pyruvate Dehydrogenase from the Archeon *Pyrobaculum aerophilum*^{†,‡}

Gary Kleiger, Jeanne Perry, and David Eisenberg*

Howard Hughes Medical Institute, UCLA-DOE Laboratory of Structural Biology and Molecular Medicine, Molecular Biology Institute, University of California, Los Angeles, Box 951570, Los Angeles, California 90095-1570

Received May 17, 2001; Revised Manuscript Received August 13, 2001

ABSTRACT: As part of a structural genomics project, we have determined the 2.0 Å structure of the E1β subunit of pyruvate dehydrogenase from *Pyrobaculum aerophilum* (PA), a thermophilic archaeon. The overall fold of E1β from PA is closely similar to the previously determined E1β structures from humans (HU) and *P. putida* (PP). However, unlike the HU and PP structures, the PA structure was determined in the absence of its partner subunit, E1α. Significant structural rearrangements occur in E1β when its E1α partner is absent, including rearrangement of several secondary structure elements such as helix C. Helix C is buried by E1α in the HU and PP structures, but makes crystal contacts in the PA structure that lead to an apparent β₄ tetramer. Static light scattering and sedimentation velocity data are consistent with the formation of PA E1β tetramers in solution. The interaction of helix C with its symmetry-related counterpart stabilizes the tetrameric interface, where two glycine residues on the same face of one helix create a packing surface for the other helix. This GΦXXG helix–helix interaction motif has previously been found in interacting transmembrane helices, and is found here at the E1α–E1β interface for both the HU and PP α₂β₂ tetramers. As a case study in structural genomics, this work illustrates that comparative analysis of protein structures can identify the structural significance of a sequence motif.

The α-keto acid dehydrogenase complexes belong to a ubiquitous family of diverse multienzyme systems that function at key points in carbohydrate metabolism (pyruvate dehydrogenase), in the citric acid cycle (α-ketoglutarate dehydrogenase), or in amino acid metabolism (branched-chain α-keto acid dehydrogenase) (1, 2). These complexes are among the largest and most complicated of known enzymes, with molecular masses of up to 10–20 million daltons. Both pyruvate dehydrogenase and branched-chain α-keto acid dehydrogenase are of clinical importance in humans (3). For example, mutations in the branched-chain α-keto acid dehydrogenase have been linked to maple syrup urine disease (4).

All α-keto acid dehydrogenase complexes are constructed from three protein components: E1, E2, and E3. The identities of E1, E2, and E3 depend on the particular class of enzyme complex, but usually the E1 component is constructed with both an α and a β subunit and functions as a substrate decarboxylase. The E2 component forms the core of the complex and acts as a dihydrolipoamide acetyltransferase. The E3 component is a dihydrolipoamide dehydrogenase.

Others have determined structures of the complete E1 α₂β₂-tetramer from both the bacterium *Pseudomonas putida* (PP)¹ and humans (5, 6). Both E1β structures are composed

of two domains. The N-terminal domain of E1β has a central six-stranded parallel β-sheet that is flanked by seven helices. The slightly smaller C-terminal domain has a five-stranded β-sheet surrounded by four helices. The two domains are connected by an extended loop. While these structures are derived from branched-chain α-keto acid dehydrogenases, they share significant sequence identity with pyruvate dehydrogenase E1β subunits and are expected to share similar structures.

Here we report the 2.0 Å structure of E1β from *P. aerophilum* in the absence of the E1α subunit. By comparing the three available E1β structures, we find both differences and similarities among the structures. A conserved packing motif was found in all three E1β structures that promotes a helix–helix interaction between E1β and E1α.

MATERIALS AND METHODS

Cloning the PA E1β Gene. Using the genomic sequence data for *P. aerophilum* (PA), we designed PCR primers for the sense (5'-GACGACGACAAGATGGTGGCTGGAGTG-GTTATGATG-3') and antisense (5'-GAGGAGAAGCCCG-GTTTACCTCATCACGTATTCTA TG-3') strands. With the primers and a PA genomic library plasmid containing the putative E1β ORF, the gene was amplified by the polymerase chain reaction (PCR). The PCR product was then incorporated into the pET30Ek/LIC expression vector (Novagen) by ligation-independent cloning.

[†] The work of G.K. was funded in part by USPHS Training Grant GM07185. This work was supported by the NIH and DOE.

[‡] PDB accession code 1IK6.

* Correspondence should be addressed to this author. Phone: (310) 825-3754. Fax: (310) 206-3914. Email: david@mbi.ucla.edu.

¹ Abbreviations: PA, *Pyrobaculum aerophilum*; PP, *Pseudomonas putida*; HU, human; BS, *Bacillus stearothermophilus*; Rms, root-mean-square; PDB, protein data bank.

Expression and Purification of PA E1β. A sequenced clone was transformed into *E. coli* BL21(DE3) cells by heat shock. A single colony was used to inoculate 10 mL of LB broth containing 50 μg/mL kanamycin. The cells were grown overnight at 37 °C with shaking. This culture was then used to inoculate 10 L of medium. The cells were grown until the A_{600} of the media measured 1.0, at which time expression was induced by adding IPTG to a final concentration of 1 mM. After 6 h of induction, the cells were harvested by centrifugation at 6000g for 15 min and stored at −20 °C.

Cell pellets were resuspended into Buffer A (20 mM Hepes, pH 8.0, 1 M NaCl, 20 mM imidazole, and 100 μg/mL PMSF) at room temperature. The addition of 1 M NaCl was based on empirical evidence that PA proteins tend to be stable only in high salt. To begin lysis, 5 mg of lysozyme was added to the cells. Lysis was further induced by sonification (5 cycles, 3 min per cycle) at a level of 70–80% of the maximum output. The cell debris was removed by centrifugation for 45 min at 33000g in a Sorvall SA-600 rotor. The supernatant was collected and filtered prior to chromatography.

To facilitate purification of E1β, a vector-encoded peptide was fused to the N-terminus. This peptide is 43 amino acids long and includes the S-tag and His-tag. PA E1β was purified on nickel-chelating resin suspended in a 5 mL Pharmacia Hi-trap column. Fractions were collected and assayed for E1β by SDS–PAGE. Several fractions containing E1β were at least 90% pure.

Fractions containing PA E1β were pooled and dialyzed into 20 mM Hepes buffer, pH 8.0, and 250 mM NaCl. This solution was heated to 65 °C for 15 min. Heat stability studies showed that PA E1β is stable at 65 °C, resulting in a significant purification step (data not shown). Denatured *E. coli* proteins were separated from PA E1β by centrifugation at 33000g for 30 min. The protein was concentrated to 27 mg/mL using a Centricon 10 centrifugal filter device (Amicon). E1β was shown to be greater than 95% pure by SDS–PAGE. In addition, MALDI-TOF (matrix-assisted laser desorption/ionization-time-of-flight) mass spectrometry revealed a single peak of approximately 39 900 Da; the expected molecular mass of PA E1β with the purification tag is 39 862 Da. This protein solution was filtered and immediately used for crystallization trials.

Crystallization and Data Collection. Screening for initial crystallization conditions was accomplished using crystal screen I by Hampton Research. Crystallization was attempted at 18 °C by hanging drop vapor diffusion against 48 conditions using equal volumes of protein and reservoir solutions. We found only one condition that produced crystals. The reservoir solution for this condition was 0.2 M CaCl₂, 0.1 M Na-Hepes, pH 7.5, and 28% PEG 400 (w/v). Several small crystals grew out of amorphous precipitate. The crystals measured only 50 μm × 50 μm × 25 μm.

Notice that the reservoir solution contains high concentrations of PEG 400, a common cryoprotectant. Crystals were mounted and frozen under liquid nitrogen. Initial native data were collected on a RAXIS IV detector using a Rigaku RU-200 rotating anode X-ray source, and a final native dataset was then collected at the Brookhaven National Laboratory on beamline X-8C using a MAR 300 imaging plate. The data diffracted to 2.0 Å, with an overall R_{merge} of 8.2%. The crystals belong to space group *I*222, with unit cell dimensions

Table 1: Statistics on Crystallographic Data and Refinement for the E1β Subunit of Pyruvate Dehydrogenase of PA

Data Collection	
resolution range high (Å)	2.0
resolution range low (Å)	50.0
data cutoff [$\sigma(F)$]	0.0
overall completeness (final res shell)	89.4 (96.0)
R_{merge} (%)	8.2 (25.7)
redundancy	5.6 (5.6)
no. of reflections (test set)	2235
no. of reflections (overall)	22189
Crystal	
space group	<i>I</i> 222
unit cell parameters	$a = 65.2 \text{ \AA}$, $b = 84.5 \text{ \AA}$, $c = 131.4 \text{ \AA}$
Refinement	
R -value (%)	21.2
free R -value (%)	25.5
mean B -value	
main-chain atoms (Å ²)	27
side-chain atoms (Å ²)	31
all atoms (Å ²)	29
rms deviations from ideal values	
bond lengths (Å)	0.013
bond angles (deg)	1.6

$a = 65.2 \text{ \AA}$, $b = 84.5 \text{ \AA}$, and $c = 131.4 \text{ \AA}$. The crystals contain one monomer in the asymmetric unit. Data were processed and scaled using the DENZO and SCALEPACK software package (7). Data collection statistics are summarized in Table 1.

Structure Determination of PA E1β by Molecular Replacement. The PA E1β structure was solved by molecular replacement using the RAXIS data and the program AmoRe (8). Using the crystal structure of the E1β subunit from *Pseudomonas putida* (PP) (6), a molecular replacement search model was generated. Using a sequence alignment between PP and PA E1β, all nonmatching residues were truncated to alanine. The highest peak in the rotation map had a correlation coefficient of 10.5, corresponding to a signal-to-noise ratio of 5.3 σ above the mean. The subsequent translation search unambiguously established the space group as *I*222, and the solution from this search had a correlation coefficient of 23.1 and a signal-to-noise ratio of 7.1 σ above the mean.

Refinement of the PA E1β Structure. Initial stages of the refinement were carried out using the RAXIS IV dataset and CNS (9). Shortly after the structure was determined, we measured an improved native dataset to 2.0 Å. For these datasets, reflections that were also present in the RAXIS IV dataset were assigned the same R_{free} flags. New reflections were assigned R_{free} flags randomly. Using these datasets, the electron density maps improved in quality, facilitating model building and completion of the structure. The final R -factor and R_{free} were 21.2 and 25.5%, respectively. Residues corresponding to the purification tag were disordered and not included in the structure. Therefore, the residues are numbered from the N-terminus of the E1β chain, discounting the 43-residue N-terminal tag. Refinement statistics are summarized in Table 1.

The following residues had weak or no density and therefore were not included in the refinement: 34–42, 89–92, 108–113, 125–134, and 299–310. The following residues had adequate density for modeling of backbone atoms; however, the side-chain densities were disordered. These residues were modeled as alanine residues: Glu 26, Met 76, Leu 93,

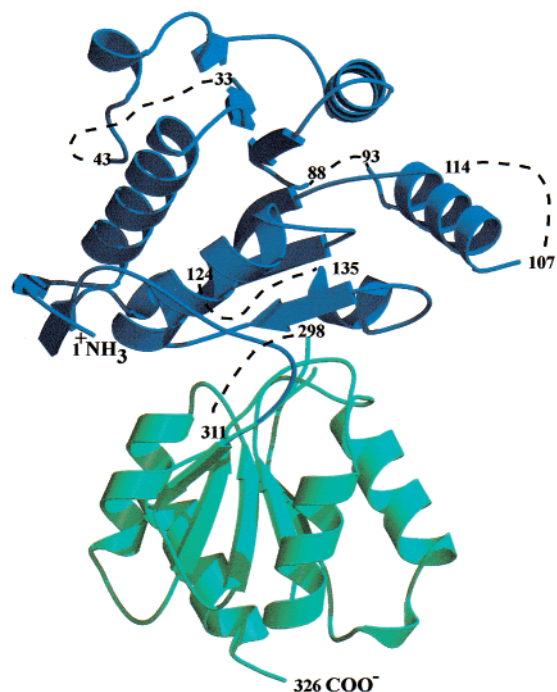


FIGURE 1: Ribbon representation of the 2.0 Å resolution crystal structure of the E1 β subunit of pyruvate dehydrogenase from *Pyrobaculum aerophilum* (PA). The N-terminal domain is dark blue. The C-terminal domain is cyan. Five loops in the structure have missing or weak electron density and are not included in the molecular coordinates. Loop 1 is located between Glu 33 and Leu 43; loop 2 is located between Val 88 and Leu 93; loop 3 is located between Tyr 107 and Lys 114; loop 4 is located between Gly 124 and Ser 135; loop 5 is located between Val 298 and Pro 311. The location of each residue adjacent to the missing loops is shown.

His 101, Ile 102, Lys 104, Tyr 107, Lys 114, Lys 178, Glu 205, Lys 229, Arg 275, Glu 314, and Arg 325.

Validation of Model Quality. Several structure validation methods were used to assess the quality of the PA E1 β model. No significant errors were found in the model using the programs VERIFY 3D, ERRAT, PROCHECK, and WHAT IF (10–12). For example, ERRAT calculated that 98.5% of the residues for the E1 β monomer are within the 95% confidence limit, indicating that the number and type of nonbonded interactions in the E1 β structure are consistent with other well-refined, high-resolution crystal structures. Approximately 94% of the residues are in the most favorable position of the Ramachandran plot, with only one residue in the disallowed region. Composite, simulated-annealing omit maps were calculated for the entire molecule. Density from the omit maps agreed very well with the atomic positions of nearly all residues, including the one residue in the disallowed region of the Ramachandran plot (Gln 238).

Velocity Sedimentation. Sedimentation velocity runs (13) were performed at 56 000 rpm and 20 °C in a Beckman Optima XL-A analytical ultracentrifuge using absorption optics at 228 nm and double sector cells. The samples were in 25 mM sodium phosphate, pH 7.0, and 125 mM Na₂SO₄. The sedimentation coefficients were determined from the slopes of $\ln(r)$ versus t plots where r was determined as the 50% position of the sedimenting boundaries (only scans where the two boundaries were clearly resolved were used). A partial specific volume of 0.742, calculated from the amino acid composition and corrected for temperature, was used.

Multiple Sequence Alignment. The amino acid sequences of 18 E1 β protein subunits have been published (SWISS-PROT IDs: odbb_bovin, odbb_human, odbb_rat, odpb_mycge, odpb_mycpn, odpb_achla, odpb_bacst, odpb_bacsu, odbb_bacsu, odbb_psepu, odpb_human, odpb_rat, odpb_ascsu, odpb_schpo, odpb_yeast, odpb_pea, odpb_porpu, acob_alceu). A multiple sequence alignment of these 18 sequences with PA E1 β was performed using the program PILEUP. These sequences represent E1 β subunits from all classes of α -keto acid dehydrogenases.

The amino acid sequences of 30 E1 α subunits have been published (SWISS-PROT IDs: odpt_mouse, odpt_rat, odpa_mouse, odpa_rat, odpa_human, odpa_pig, odpa_smima, odpt_human, odpa_ascsu, odpt_ascsu, odpa_cael, odpa_pea, odpa_soltu, odpa_arath, odpa_klula, odpa_yeast, odpa_schpo, acoa_alceu, odpa_porpu, odpa_mycge, odpa_mycpn, odpa_bacst, odpa_bacsu, odpa_achla, odpa_mouse, odpa_rat, odpa_human, odpa_bovin, odpa_bacsu, odpa_psepu). A multiple sequence alignment of these 30 sequences with PA E1 α was performed using PILEUP.

Calculation of Interaxial Helical Angles. The interaxial angle at which a given pair of helices cross each other was calculated using the following algorithm. A normalized eigenvector was calculated using all CA atoms in one helix. The eigenvector corresponds to the line of best fit along the helical axis. This procedure is then repeated for the other helix. The dot product of these vectors gives $\cos \theta$, where θ is the angle between the two helices.

RESULTS

(A) Structural Comparison of *P. aerophilum* E1 β with *P. putida* and Human E1 β

Similarities in Structure. Using a homology model derived from the *P. putida* (PP) E1 β structure (6) as a probe for molecular replacement, we solved the structure for *P. aerophilum* (PA) E1 β to 2.0 Å resolution (Figure 1).

As expected from the 48% sequence identity between the two protein sequences, the structures of PA and PP E1 β are highly similar. The structure of a third E1 β subunit from humans (HU) (5) was recently determined, and it shares both significant sequence (44%) and structural similarity with PA E1 β . Using the program ALIGN_v2 (14), the structure of PA E1 β was superimposed onto the structure of PP E1 β with a resulting rms distance between 1045 main-chain atom pairs of 1.11 Å. Using an identical procedure, we aligned PA E1 β to HU E1 β . The rms distance between 1046 aligned main-chain atom pairs is 1.07 Å. Finally, PP E1 β was structurally aligned to HU E1 β . The rms distance between 1213 main-chain atom pairs is 0.71 Å. As noted earlier, the E1 β fold is highly conserved.

The most striking similarity shared by all three E1 β subunits is a conserved tertiary interaction that connects the N-terminal domain of E1 β to the C-terminal domain. For PA E1 β , Gln 238 of the C-terminal domain forms a 3.0 Å hydrogen bond with Asn 157 of the N-terminal domain (Figure 3a). For HU E1 β , Arg 255 of HU E1 β forms a 2.9 Å hydrogen bond with Gln 170 (Figure 3b). For PP E1 β , Arg 255 forms a salt bridge to Asp 155 (Figure 3c). The residues forming these interactions occupy equivalent positions in these structures (Figure 2) (15). Therefore, the

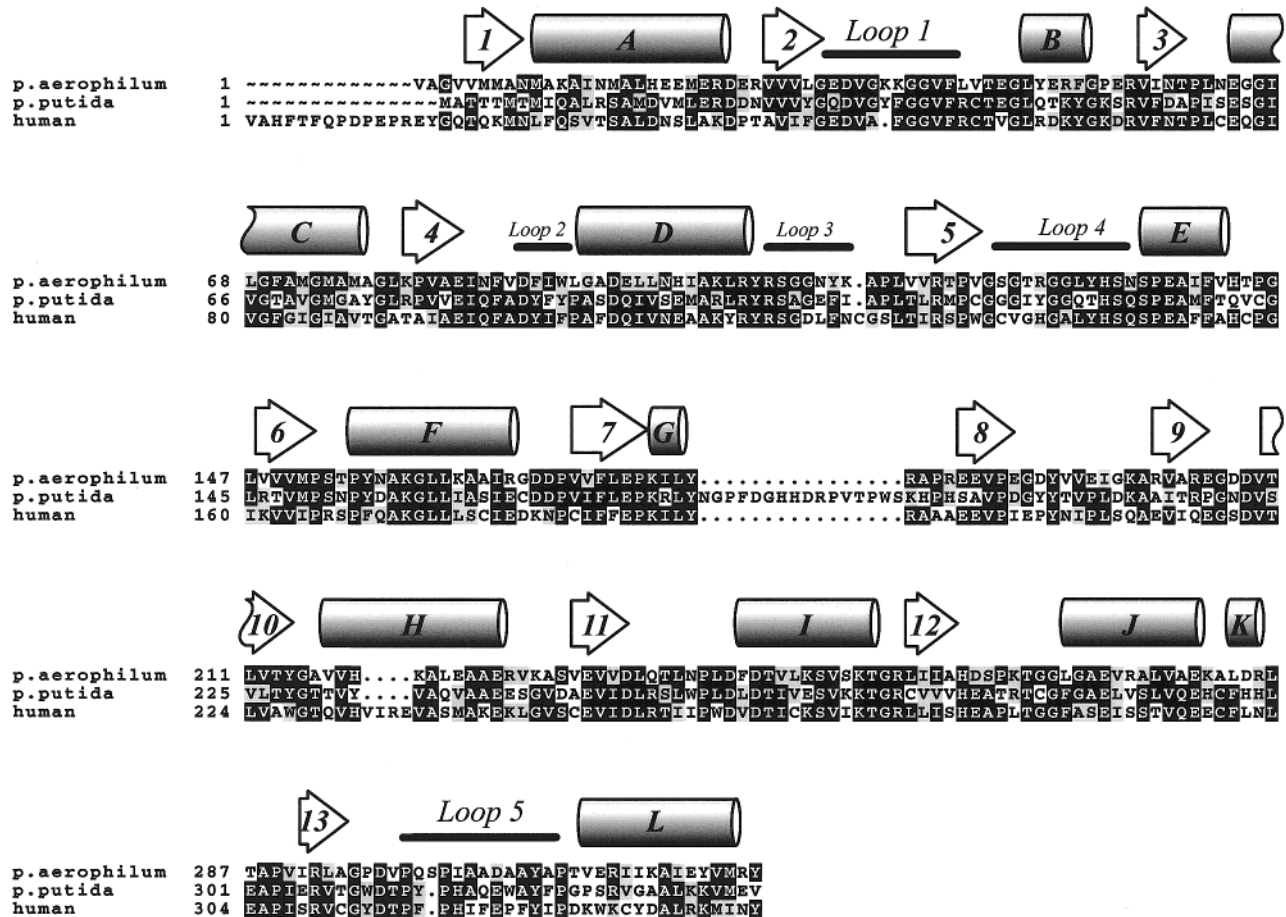


FIGURE 2: Sequence alignment of three E1 β subunits of known structure. The sequence of the E1 β subunit from PA is on top of the alignment, and the *Pseudomonas putida* (PP) and human (HU) sequences are in the middle and the bottom, respectively. The PA E1 β subunit is part of the multisubunit enzyme pyruvate dehydrogenase, whereas the PP and HU subunits are parts of branched-chain α -keto acid dehydrogenases. Aligned residues with similar physical-chemical characteristics are shaded gray, whereas identical residues are shaded black. Secondary structure assignment according to the program DSSP (30) for PA E1 β is shown above the alignment. Helices are shown as shaded cylinders labeled with alphabetical letters which indicate the N- to C-terminal order of the helices in the structure; β -strands are indicated as numbered arrows. Five black bars indicate where the missing loops in the PA E1 β structure are located in the alignment. Part of this figure was generated using the program BOXSHADE.

interaction between these residues is conserved, connecting equivalent positions on each of the E1 β structures. It is noteworthy that the backbone phi and psi angles for Gln 238 (PA, 68°, -61°), Arg 252 (PP, 70°, -42°), and Arg 255 (HU, 65°, -70°) are located in the disallowed region of the Ramachandran plot. Because the unfavorable energy of this backbone conformation is tolerated, it suggests that this tertiary interaction is important for stabilizing the relative orientation of the two domains of E1 β . A multiple sequence alignment of 19 available E1 β protein sequences shows that these residues tend to be conserved or conservatively substituted. Therefore, this interaction is expected to be conserved among all E1 β subunits.

Why is this tertiary interaction important? Both domains are involved in intermolecular interactions; the N-terminal domain interacts with the E1 α subunit, and the C-terminal domain interacts with the E2 subunit (6). The E1, E2, and E3 components associate to form a large, highly symmetric complex. Whereas a flexible linker between the N- and C-terminal domains would allow each domain to rotate independently, the tertiary interaction locks the N- and C-terminal domains into specific orientations. Precise orientation of each subunit and its associated symmetry elements is necessary to obtain an ordered complex (16).

Therefore, orientation of the N- and C-terminal domains in E1 β is likely to be important for proper dehydrogenase assembly and function.

Differences in Structure. While the overall E1 β fold is conserved, closer inspection of the three available structures suggested that they diverge from each other locally. Five loops are disordered in the PA E1 β structure, whereas these loops are all ordered in both the PP and HU E1 β structures. Electron density in both $2F_o - F_c$ and $F_o - F_c$ maps is either very weak or absent for these residues in the PA structure. Analysis of both the PP and HU $\alpha_2\beta_2$ -tetramer structures showed that all five loops contribute to intermolecular contacts at the E1 α -E1 β interface. The differences between the PA E1 β structure and the PP and HU E1 β structures are likely related to the fact that PA E1 β was isolated and crystallized in the absence of its interacting partner, E1 α . It is therefore likely that the loops will order upon binding to PA E1 α .

(B) Identification of a PA E1 β Tetramer and Its Relationship to the Oligomeric Interface of the E1 $\alpha_2\beta_2$ Tetramer

Potential Oligomeric Interfaces for PA E1 β . The E1 β subunit forms extensive interfaces with the α , α' , and the β'

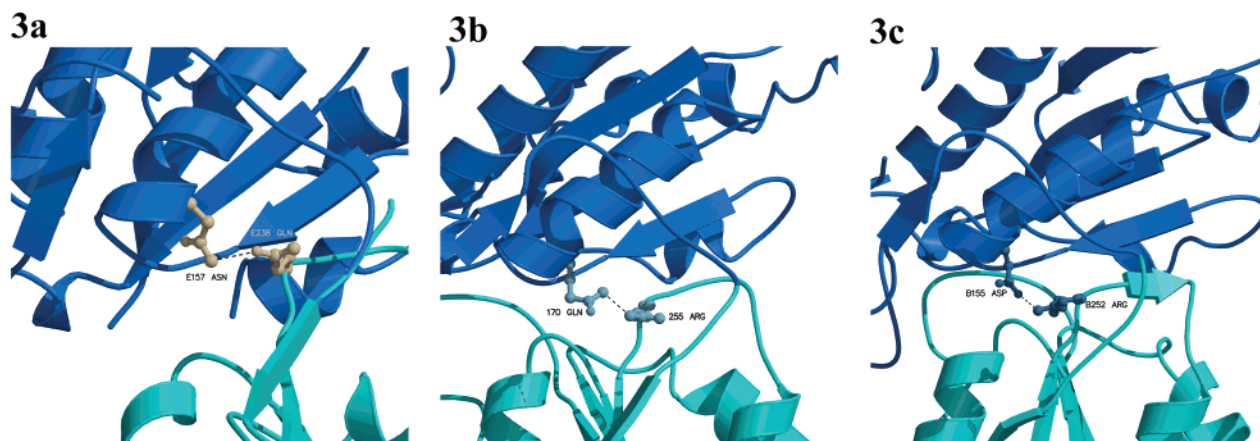


FIGURE 3: Conserved tertiary interaction that links the N- and C-terminal domains of the E1 β subunits. Gln 238 of PA E1 β , Arg 255 of HU E1 β , and Arg 252 of PP E1 β are structurally equivalent. They each form a conserved tertiary interaction to structurally equivalent residues in the N-terminal domain. (a) Gln 238 of the C-terminal of PA E1 β forms a hydrogen bond to Asn 157 of the N-terminal domain. The N-terminal domain is dark blue; the C-terminal domain is cyan. Both residues are tan. (b) Arg 255 of the C-terminal domain of HU E1 β forms a hydrogen bond to Gln 170 of the N-terminal domain. Each domain is colored as in (a); both residues are light blue. (c) Arg 252 of the C-terminal domain of PP E1 β forms a salt bridge to Asp 155 of the N-terminal domain. Each domain is colored as in (a); both residues are dark blue. The significance of this interaction is that it stabilizes the orientation of the N-terminal domain of E1 β relative to the C-terminal one and is likely to be important for dehydrogenase assembly.

subunits in both the PP and HU E1 $\alpha_2\beta_2$ -tetramer structures. While the asymmetric unit of the PA E1 β crystal contains only one subunit, two potential oligomeric interfaces were found at crystal contacts. The largest of these interfaces, referred to as interface I, preserves the $\beta\beta'$ interface observed in the PP and HU E1 $\alpha_2\beta_2$ tetramers. Approximately 1376 Å² of accessible surface area is buried for each PA E1 β subunit when they are brought together at interface I.

Another potential interface for PA E1 β , called interface II, was also identified at a crystal contact. Approximately 1322 Å² of surface area is buried from solvent at interface II. Here, two of the dimeric E1 β subunits are related by a crystallographic 2-fold to create a β_4 tetramer. The 1322 Å² of buried surface area is well within the range of values observed for interactions of known biological significance (17). The structure of the PA β_4 tetramer is shown in Figure 4a.

Consistent with our observations of crystal packing, PA E1 β also forms tetramers in solution. Sedimentation velocity experiments confirmed that PA E1 β forms tetramers in solution. Two peaks can be seen in the scan (Figure 5); the sedimentation coefficient of the slower peak corresponds to monomer. If the shape does not change drastically upon oligomerization, the ratio of the sedimentation coefficients (6.75 ± 0.02 for the fast peak, 2.75 ± 0.03 for the slow peak) can be used to calculate the ratio of the molecular masses. Here, the ratio of molecular masses for the two peaks is 3.84, confirming that PA E1 β forms tetramers in solution. Oligomerization is independent of the purification tag, which is disordered in the crystal structure and is distant from both intermolecular interfaces.

Analysis of the Atomic Interactions That Stabilize the PA β_4 Tetramer. The stability of the PA E1 β tetramer is most likely due to the close packing of helix C of the N-terminal domain with its symmetry mate helix C'' in the N-terminal domain of a 2-fold-related molecule. The interaction of helix C with helix C'' is stabilized by van der Waals interactions, where Phe 70 from one helix packs against the adjacent helix (Figures 4c and 6a). A flat surface for packing is created by

Gly 69'' and Gly 73''. Notice that the four-residue separation of these glycine residues aligns them on the same side of the helix. Phe 70 forms extensive van der Waals interactions primarily with Gly 73'', although favorable contacts are also observed with Gly 69''. Because helices C and C'' are related by a crystallographic 2-fold axis, Phe 70'' interacts in an identical manner with Gly 69 and Gly 73 of helix C. Also of importance are the interactions of Met 74 with helix C'' and Met 74'' with helix C.

Glycophorin A and the G Φ XXG Motif. Other examples of protein–protein interactions mediated by helix–helix association and glycine residues can be found in the PDB. Dr. William Russ drew our attention to glycophorin A, which has a single transmembrane helix which self-associates to form a symmetric, right-handed homodimer (18). Structural analysis of the helix–helix interaction by NMR showed that two glycines, Gly 79 and 83, form a flat surface onto which Val 80' from the symmetry-related helix packs against (Figures 4f and 6b). By symmetry, Val 80 forms an identical interaction with Gly 79' and Gly 83'.

The helix–helix packings in the PA E1 β tetramer and glycophorin A are nearly identical. The two structures superimpose with an rms distance of 1.03 Å for 126 main-chain atoms (residues 63–78, corresponding to helix C, and residues 63''–78'', corresponding to helix C''), were superimposed onto residues 73–88 and 73'–88' of the transmembrane helices of glycophorin A). Gly 69 and 73 of PA E1 β align with Gly 79 and 83 of glycophorin A. More importantly, Phe 70 of PA E1 β aligns with Val 80 of glycophorin A, demonstrating that these two proteins share highly similar motifs for stabilizing the helix–helix interactions. Therefore, we may infer a packing motif, G Φ XXG, that is in common between the PA β_4 and the glycophorin A helical interactions, where Φ represents the residue that packs on the glycine face.

Analysis of a Conserved Helix–Helix Interaction at the $\alpha\beta$ Interface for both the HU and PP $\alpha_2\beta_2$ E1 Tetramers. In the HU $\alpha_2\beta_2$ structure, helix C of E1 β interacts with helix 6 from the E1 α subunit (Figure 4b,d) in a manner reminiscent

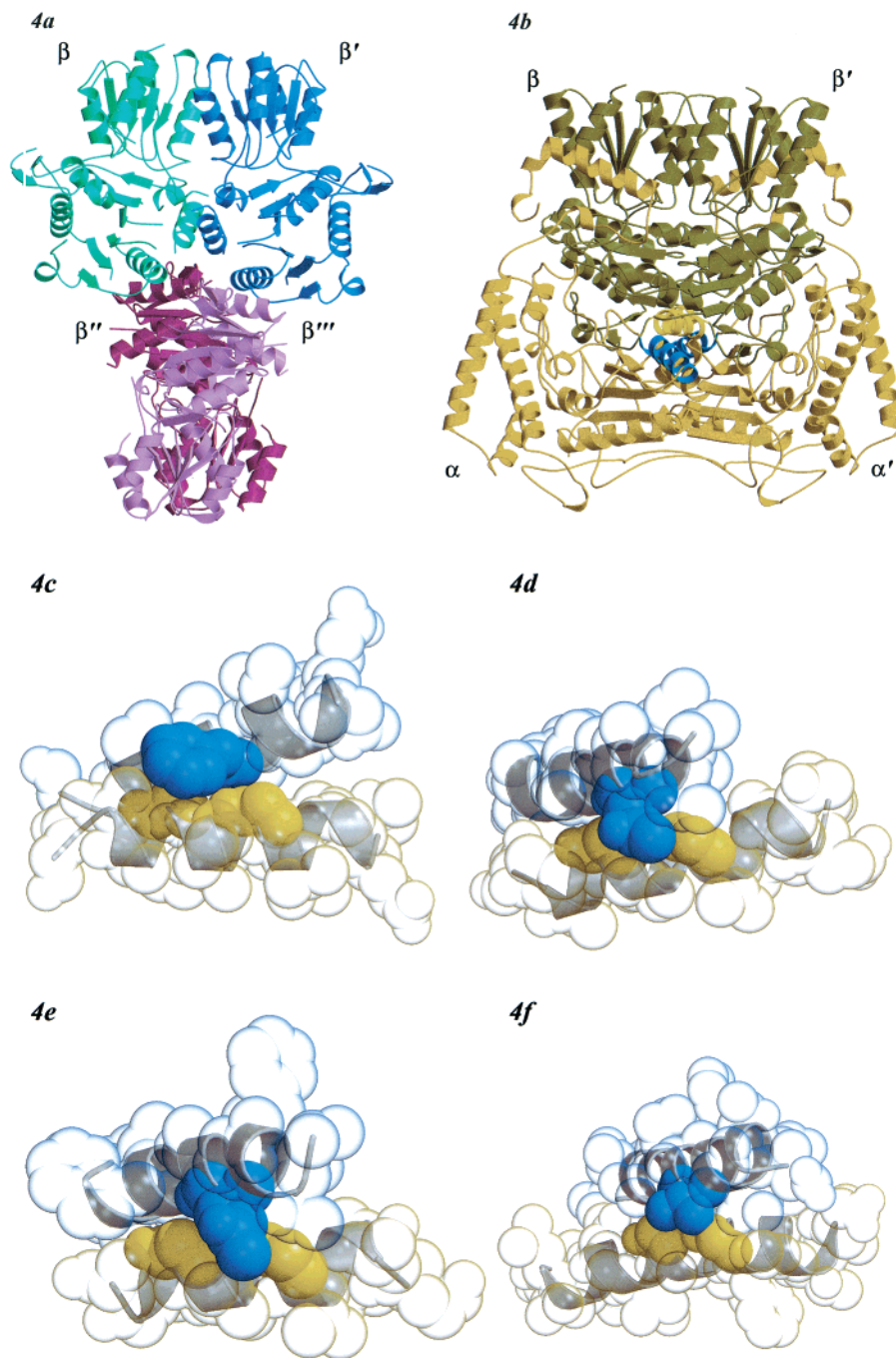


FIGURE 4: (a) Crystal structure of the PA E1 β β_4 tetramer. Each monomer is given a different color for clarity. The blue and cyan monomers form a dimer with an oligomeric interface referred to as interface I. The dark red and plum monomers also form a dimer with an identical interface. An additional interface occurs where the two dimers meet to form the tetramer. This interface is referred to as interface II. The four subunits in the tetramer are related by 222 symmetry. (b) Ribbon representation of the crystal structure of the HU $\alpha_2\beta_2$ -tetramer. Both of the β subunits are olive; the α subunits are light brown. Helix C of HU E1 β is gold; helix 6 of HU E1 α is blue. (c) Closer examination of the interaction between helices C and C' for the PA β_4 tetramer at interface II. Atoms from helix C' are shown as transparent blue spheres; atoms from helix C are transparent gold spheres. The backbone of each helix is shown as a ribbon. Phe 70' (dark blue spheres) of helix C' interacts with Gly 69 and Gly 73 (gold spheres) from helix C. The symmetry-related interaction of Phe 70 with Gly 69' and Gly 73' is not shown for clarity. These interactions bury the majority of protein surface area from solvent, suggesting that Phe 70 is the most important residue for stabilizing interface II. (d) Closer examination of the interaction between helix C of HU E1 β with helix 6 of HU E1 α . Atoms from helix C are transparent gold spheres; atoms from helix C are transparent blue spheres. Gln A170 (dark blue spheres) of HU E1 α packs against the flat surface provided by Gly B81 and B85 (dark gold spheres) of HU E1 β . (e) Same as 4d, except using the PP $\alpha_2\beta_2$ -tetramer. Here, Gln A190 (dark blue spheres) of PP E1 α packs against helix C and Gly B67 and B71 (dark gold spheres) of PP E1 β . (f) Helix-helix interaction observed in the transmembrane protein glycophorin A. Val 80 from one helix packs against Gly 79' and 83' of the symmetry-related one. Atoms from Val 80 are represented as dark blue spheres; atoms from Gly 79' and 83' are represented as dark gold spheres. The symmetry-related interaction is not shown for clarity. Notice the close similarity of this homodimeric interaction to the heterodimeric interaction of 4d and 4e.

of the helix C-helix C' interaction observed in the PA β_4 tetramer. We superimposed coordinates for the PA E1 β helix

C-helix C' interaction onto those for the helix-helix interaction found at the HU E1 α -E1 β interface (residues

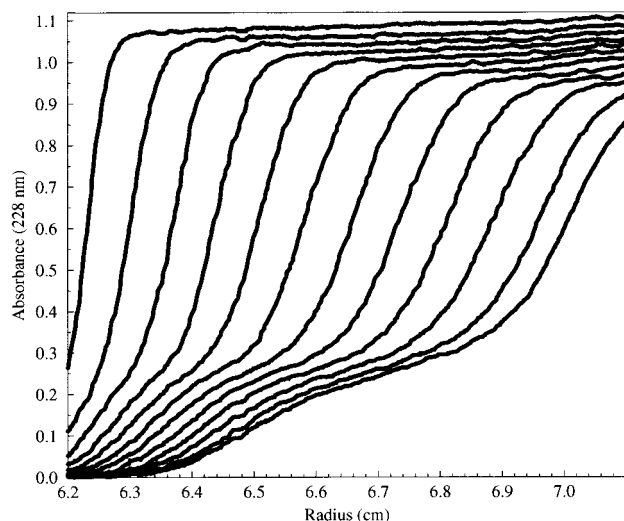


FIGURE 5: Samples of 0.2 mg/mL purified PA E1- β were examined in 12 mm double sector cells at 56 000 rpm and 20 °C in a Beckman Optima XL-A analytical ultracentrifuge operating at 228 nm. Plots of absorbance versus radial position in the cell were acquired every 3 min. Here consecutive scans at 9 min intervals are shown. Two components are present.

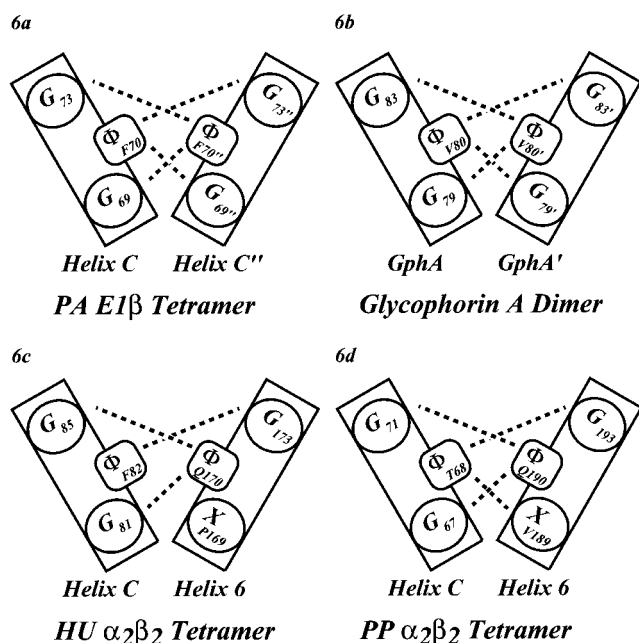


FIGURE 6: Schematic of the four helix-helix interactions which utilize the G Φ XXG motif. Residue pairs which form van der Waals interactions are connected with dotted lines. (a) Interaction of helix C and helix C'' in the PA E1 β tetramer. (b) Glycophorin A dimer. (c) Interaction of helix C and helix 6 at the HU E1 α -E1 β interface. (d) Interaction of helix C and helix 6 at the PP E1 α -E1 β interface.

A165–A180, corresponding to helix 6 from HU E1 α , and residues B75–B89, corresponding to helix C from HU E1 β). Helix C of HU E1 β packs with helix 6 of HU E1 α with an interaxial angle of -57° . Helix C and helix C'' of the PA β_4 tetramer pack with an interaxial angle of -46° . Despite the 11° difference in interaxial packing angles, the two backbones superimpose with only a 0.63 Å rms deviation for 107 aligned backbone atoms.

The atomic interactions which stabilize the HU E1 α -E1 β helix-helix interaction are similar to the interactions which stabilize the helix C-helix C'' interface in the PA β_4

tetramer, although helix 6 of HU E1 α does not conserve the first glycine of the G Φ XXG motif. Helix C of HU E1 β has two glycine residues, Gly B81 and B85, both on the same side of the helix, forming a complementary surface. These glycines structurally superimpose with the glycines of the G Φ XXG motif of PA E1 β helix C. Both glycines are conserved at equivalent positions in all 19 available E1 β sequences. In addition, Gln 170 of helix 6 (from HU E1 α) packs against Gly B81 and B85 of helix C. Gln 170 is located at the Φ position of helix 6. Likewise, Phe 82 (located at the Φ position of helix C) forms van der Waals interactions with Gly 173 of helix 6 (Figure 6c). Gly 173 corresponds to the second glycine of the G Φ XXG motif and is conserved at structurally equivalent positions for all 31 E1 α sequences. However, the first glycine of the motif is not conserved in E1 α subunits. For example, the residue at this position in HU E1 α is Pro 169 and does not interact with Phe 82 (Figure 6c). This demonstrates that, at least for the oligomerization of HU E1 β with E1 α , strict conservation of the G Φ XXG motif is not necessary to promote helix-helix association.

We also found that helix 6 of PP E1 α interacts with helix C of PP E1 β in an analogous manner to the HU helix-helix interaction. We superimposed coordinates for the PA E1 β helix C-helix C'' interaction onto those for the helix-helix interaction found at the PP E1 α -E1 β interface (residues A186–A200, corresponding to helix 6 from PP E1 α , and residues B61–B75, corresponding to helix C from PP E1 β). The two structures aligned with an rms deviation of 1.01 Å over 100 backbone atoms. In a highly analogous manner as the HU helical interaction, Gln 190 of PP E1 α is located at the Φ position on helix 6. Gln 190 interacts with helix C of PP E1 β , where Gly 67 and Gly 71 form the conserved packing surface (Figures 4e and 6d). Likewise, Thr 68 (located at the Φ position on helix C) forms van der Waals interactions with Gly 193 and Val 189 on helix 6 (Figure 6d). Therefore, the helix-helix interaction observed at the E1 α -E1 β interface for both the HU and PP $\alpha_2\beta_2$ -tetramers is structurally conserved.

DISCUSSION

Helix-Helix Interactions and the G Φ XXG Motif. A helix packing motif of G Φ XXG was identified in the soluble E1 β_4 -tetramer, essentially identical to that found previously in the glycophorin A dimer (18). This motif promotes homo-oligomerization by both permitting two helices to pack together and stabilizing the interface between them. Analysis of both the HU and PP E1 $\alpha_2\beta_2$ -tetramers showed that a similar mechanism is used to promote a helix-helix interaction and the oligomerization of the E1 α and E1 β subunits. This is the first example of two proteins in a nonmembrane environment that use the G Φ XXG motif and demonstrates that this motif may be used to drive both homo- and hetero-oligomerization of proteins.

It is noteworthy that both E1 and glycophorin A use similar packing motifs for association. In the former case, two soluble protein subunits, E1 α and E1 β , interact to form the $\alpha_2\beta_2$ -tetramer in the aqueous environment of the mitochondrial matrix (human) or cytosol (bacterial). In the latter case, a membrane protein self-associates to form a dimer in the apolar environment of the membrane lipid bilayer. It must be noted that the interaction of helix C and helix 6 is only

part of a very large interface between E1 α and E1 β . Many other secondary structure elements come together at the interface creating a hydrophobic environment. In this sense, also E1 assembly and glycophorin A oligomerization are similar: the environment of helix C and helix 6 is somewhat apolar and perhaps resembles the hydrophobicity of the membrane.

What makes the GΦXXG motif well adapted for helix–helix interactions? First, it provides a complementary surface for packing. Second, it permits the two helices to come into close contact with each other, facilitating contact between the other interfacial residues. Third, there is no loss of side-chain entropy for glycine residues upon dimerization, whereas residues with many conformations would lose entropy upon association and destabilize the interaction (19).

A recent study of sequences predicted to form helices in membrane proteins showed that the most statistically significant motif is GXXXG, suggesting that numerous membrane proteins utilize this motif for helix–helix association (20). Also, an *in vivo* selection system designed to search for probable transmembrane oligomerization motifs almost invariably came up with GΦXXG, where the residue located at the Φ position was usually a β-branched amino acid (21, 22). Apparently, this motif is common for helix–helix interactions in membrane proteins. Our results extend the role of the GΦXXG motif to helix–helix interactions between protein subunits located in an aqueous environment. Also consistent with the findings of Russ and Engelman (22), our results also demonstrate that strict conservation of the GΦXXG motif is not necessary to promote oligomerization. For example, position 189 of PP E1 α is the first position of the GΦXXG motif, yet the residue occupying this position is valine. Nevertheless, Val 189 forms van der Waals interactions with Thr 68 (located at the Φ position on helix C of PP E1 β), demonstrating how variation on the motif may be tolerated. In addition, the interaxial helical angle for the four helix–helix interactions differs by as much as 15° (data not shown), further showing the flexibility of the GΦXXG packing motif.

Possible Biological Significance of PA β_4 Tetramers. PA E1 β forms β_4 tetramers, both in the crystal and in solution. It was unexpected that PA E1 β is capable of forming highly ordered and stable oligomers. For example, attempts to express HU E1 β from the branched-chain α -keto acid dehydrogenase complex yielded insoluble aggregates (23). Studies have indicated that proper assembly of the $\alpha_2\beta_2$ -tetramer in humans is dependent on chaperone proteins (24, 25). Expression of bovine E1 β did yield soluble protein; however, high molecular weight aggregates were observed that did not behave as a single molecular species (26). In contrast, expression of pyruvate dehydrogenase E1 β from *Bacillus stearothermophilus* (BS) was successful (27). BS E1 β also forms tetramers in solution, providing a second example of an E1 β subunit that forms β_4 tetramers in the absence of E1 α (27). In addition, BS E1 α was also expressed independently. When mixed, the BS E1 α and E1 β subunits formed functional $\alpha_2\beta_2$ -tetramers *in vitro*, demonstrating that, at least for BS, reconstitution was possible (27).

Why do both PA and BS E1 β form tetramers? Our hypothesis, based on sequence alignment, is that all E1 β subunits contain the GΦXXG motif, which drives the association of protein helices. Placed on the exterior of a

protein subunit, such as E1 β , a helix with this sequence may wrongfully associate with other proteins or with itself. There appear to be at least two options to prevent wrongful aggregation. First, chaperones (such as the mitochondrial chaperone hsp60/hsp10) can be used to sequester and aid in the assembly of eukaryotic E1 α and E1 β after import into the mitochondria (28). For prokaryotes, chaperones such as GroEL/GroES can aid in the assembly of E1 α and E1 β in the bacterial cytosol. Second, helix C of E1 β can be covered so that it is no longer exposed to solvent. This is accomplished in the PA β_4 tetramer. Formation of a β_4 tetramer is not expected to affect normal E1 function *in vivo*. The interface between E1 α and E1 β is clearly more stable, burying nearly 3 times as much surface area as the β_4 tetramer. Once the E1 α subunit is present, the β_4 tetramer could dissociate into monomers that would then associate with E1 α to form an $\alpha\beta$ dimer. These dimers can then oligomerize to form functional $\alpha_2\beta_2$ -tetramers, similar to the assembly pathway of the E1 component of the human branched-chain α -keto dehydrogenase (25, 29).

Notice that the two examples of E1 β subunits that form β_4 tetramers in the absence of E1 α come from thermophilic organisms. Perhaps, at higher temperatures, it is even more imperative that helix C be covered to prevent inappropriate oligomers. As E1 β subunits from a more diverse group of organisms are expressed, additional examples of β_4 tetramers may be observed, revealing the degree of generality of this pattern of assembly.

ACKNOWLEDGMENT

We thank Michael Sawaya, Martin Phillips, and Mari Gingery for helpful discussion. We also thank Martin Phillips and Duilio Cascio for data collection.

REFERENCES

- de Kok, A., Hengeveld, A. F., Martin, A., and Westphal, A. H. (1998) *Biochim. Biophys. Acta* 1385, 353–366.
- Perham, R. N. (2000) *Annu. Rev. Biochem.* 69, 961–1004.
- Patel, M. S., Roche, T. E., and Harris, R. A. (1996) *Alpha-keto acid dehydrogenase complexes*, Birkhäuser Verlag, Boston.
- Chuang, D. T. (1998) *J. Pediatr.* 132, S17–23.
- Aevarsson, A., Chuang, J. L., Wynn, R. M., Turley, S., Chuang, D. T., and Hol, W. G. (2000) *Struct. Folding Des.* 8, 277–291.
- Aevarsson, A., Seger, K., Turley, S., Sokatch, J. R., and Hol, W. G. (1999) *Nat. Struct. Biol.* 6, 785–792.
- Otwinowski, Z. (1993) (Sawyer, L., Isaacs, N., and Bailey, S., Eds.) pp 56–62, SERC Daresbury Laboratory, Warrington, U.K.
- Navaza, J. (1994) *Acta Crystallogr., Sect. A* 50, 157–163.
- Brunger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J. S., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T., and Warren, G. L. (1998) *Acta Crystallogr., Sect. D: Biol. Crystallogr.* 54, 905–921.
- Colovos, C., and Yeates, T. O. (1993) *Protein Sci.* 2, 1511–1519.
- Laskowski, R. A., McArthur, M. W., Moss, D. S., and Thornton, J. M. (1993) *J. Appl. Crystallogr.* 26, 283–291.
- Luthy, R., Bowie, J. U., and Eisenberg, D. (1992) *Nature* 356, 83–85.
- Schuster, T. M., and Laue, T. M. (1994) *Modern analytical ultracentrifugation: acquisition and interpretation of data for biological and synthetic polymer systems*, Birkhäuser, Boston.

14. Cohen, G. H. (1986) *J. Mol. Biol.* 190, 593–604.
15. Kleiger, G., Beamer, L. J., Grothe, R., Mallick, P., and Eisenberg, D. (2000) *J. Mol. Biol.* 299, 1019–1034.
16. Padilla, J. E., Colovos, C., and Yeates, T. O. (2001) *Proc. Natl. Acad. Sci. U.S.A.* 98, 2217–2221.
17. Jones, S., and Thornton, J. M. (1996) *Proc. Natl. Acad. Sci. U.S.A.* 93, 13–20.
18. MacKenzie, K. R., Prestegard, J. H., and Engelman, D. M. (1997) *Science* 276, 131–133.
19. MacKenzie, K. R., and Engelman, D. M. (1998) *Proc. Natl. Acad. Sci. U.S.A.* 95, 3583–3590.
20. Senes, A., Gerstein, M., and Engelman, D. M. (2000) *J. Mol. Biol.* 296, 921–936.
21. Russ, W. P., and Engelman, D. M. (1999) *Proc. Natl. Acad. Sci. U.S.A.* 96, 863–868.
22. Russ, W. P., and Engelman, D. M. (2000) *J. Mol. Biol.* 296, 911–919.
23. Wynn, R. M., Davie, J. R., Cox, R. P., and Chuang, D. T. (1992) *J. Biol. Chem.* 267, 12400–12403.
24. Song, J. L., Wynn, R. M., and Chuang, D. T. (2000) *J. Biol. Chem.* 275, 22305–22312.
25. Wynn, R. M., Song, J. L., and Chuang, D. T. (2000) *J. Biol. Chem.* 275, 2786–2794.
26. Wynn, R. M., Chuang, J. L., Davie, J. R., Fisher, C. W., Hale, M. A., Cox, R. P., and Chuang, D. T. (1992) *J. Biol. Chem.* 267, 1881–1887.
27. Lessard, I. A., and Perham, R. N. (1994) *J. Biol. Chem.* 269, 10378–10383.
28. Martin, J. (1997) *J. Bioenerg. Biomembr.* 29, 35–43.
29. Chuang, J. L., Wynn, R. M., Song, J. L., and Chuang, D. T. (1999) *J. Biol. Chem.* 274, 10395–10404.
30. Kabsch, W., and Sander, C. (1983) *Biopolymers* 22, 2577–2637.

BI011016K