Atomic Structure of the Cubic Core of the Pyruvate Dehydrogenase Multienzyme Complex

ANDREA MATTEVI, GALYA OBMOLOVA, EGBERT SCHULZE, KOR H. KALK, ADRIE H. WESTPHAL, ARIE DE KOK, WIM G. J. HOL*

The highly symmetric pyruvate dehydrogenase multienzyme complexes have molecular masses ranging from 5 to 10 million daltons. They consist of numerous copies of three different enzymes: pyruvate dehydrogenase, dihydrolipoyl transacetylase, and lipoamide dehydrogenase. The three-dimensional crystal structure of the catalytic domain of *Azotobacter vinelandii* dihydrolipoyl transacetylase has been determined at 2.6 angstrom (Å) resolution. Eight trimers assemble as a hollow truncated cube with an edge of 125 Å, forming the core of the multienzyme complex. Coenzyme A must enter the 29 Å long active site channel from the inside of the cube, and lipoamide must enter from the outside. The trimer of the catalytic domain of dihydrolipoyl transacetylase has a topology identical to chloramphenicol acetyl transferase. The atomic structure of the 24-subunit cubic core provides a framework for understanding all pyruvate dehydrogenase and related multienzyme complexes.

THE PYRUVATE DEHYDROGENASE MULTIENZYME COMPLEX (PDC) catalyzes a key reaction in the aerobic energy-generating glucose metabolism because it occurs at the junction of the glycolysis and the citric acid cycle (*1*). The PDC is composed of numerous copies of at least three different enzymes, pyruvate decarboxylase (E1p; E.C. 1.2.4.1), dihydrolipoyl transacetylase (E2p; E.C. 2.3.1.12), and lipoamide dehydrogenase (E3; E.C. 1.8.1.4), yielding a total molecular mass of 5 to 10 million daltons. The complex catalyzes the decarboxylation of pyruvate and the acetylation of coenzyme A (CoA) by a multistep reaction involving not less than five different cofactors: thiamin pyrophosphate (TPP), dihydrolipoamide [Lip(SH)$_2$], nicotinamide adenine dinucleotide (NAD), flavin adenine dinucleotide (FAD), and CoA (Fig. 1). Deficiencies in the activity of this system lead to various pathological states, whose typical clinical manifestations are metabolic acidosis and neurological dysfunctions (*2, 3*).

In all organisms studied the structural core of the complex is made up by the transacetylase, although the architecture of the PDC differs among the various sources. In mammals and Gram-positive bacteria, 60 copies of E2p are associated to form the icosahedral core of the complex, whereas in Gram-negative bacteria the inner part of the system consists of 24 E2p molecules arranged with octahedral symmetry [for reviews, see (*2, 4–8*)].

The number of potential binding sites for E1p and E3 is related to the symmetry of the core. Although the compositions of isolated complexes vary (*9*), reconstitution experiments with the proteins from Gram-negative sources show that optimum catalytic activity is obtained with complexes that contain 12 dimers of E1p and 6 dimers of E3 per 24 molecules of E2p (*10, 11*).

The E2p component also plays a central role in the functioning of the complex. Each E2p subunit is composed of several domains (*8, 12, 13*): (i) one or more lipoyl domains of ~100 residues, each carrying a lipoamide molecule covalently attached to a lysine; (ii) an E1p-E3 binding domain of ~50 residues; and (iii) a catalytic COOH-terminal domain of 250 residues that contains the catalytic center and the intersubunit E2p binding site. The domains are linked to each other by segments rich in alanine and proline that have a high degree of conformational flexibility (*14*) and that have an essential role for PDC activity (*15*). They provide the lipoyl domains with the mobility required for interacting with the catalytic centers located on the three different enzymes in the complex.

The PDC appears to be the prototype for the entire class of 2-oxoacid dehydrogenase complexes (*1, 8*), which comprises the oxo-glutarate dehydrogenase (OGDH) and the branched-chain oxoacid dehydrogenase (BCODH). They all share a common organization, with the transferase enzyme (E2p in PDC, succinyl transferase or "E2o" in OGDH, and transacylase or "E2b" in BCODH) forming the core of these assemblies. Furthermore, E2p, E2o, and E2b show a considerable homology in amino acid sequence (*16*). In almost all organisms the same lipoamide dehydrogenase dimer occurs in all three multienzyme complexes (*8*), with the only known exception of *Pseudomonas putida* (*17*).

The PDC from Gram-negative *Azotobacter vinelandii* is well characterized (*18*). The isolated E2p core consists of 24 subunits,

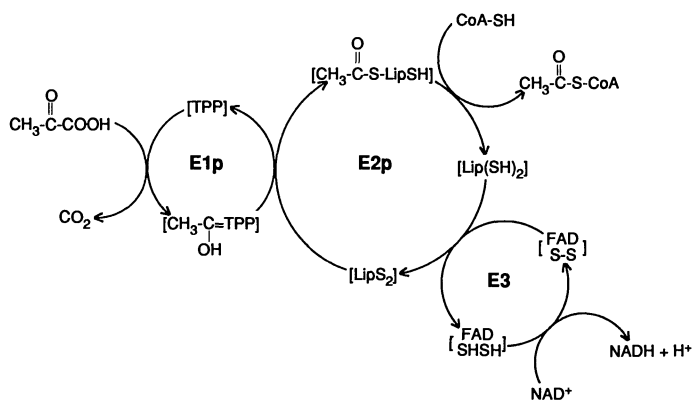

Fig. 1. Reaction sequence in the pyruvate dehydrogenase complex (*1*).

A. Mattevi, G. Obmolova, K. H. Kalk, and W. G. J. Hol are in the BIOSON Research Institute, Department of Chemistry, University of Groningen. Nijenborgh 4, 9747 AG Groningen, The Netherlands. E. Schulze, A. H. Westphal, and A. de Kok are in the Department of Biochemistry, Agricultural University of Wageningen, Dreijenlaan 3, 6703 HA Wageningen, The Netherlands.

*To whom correspondence should be addressed.

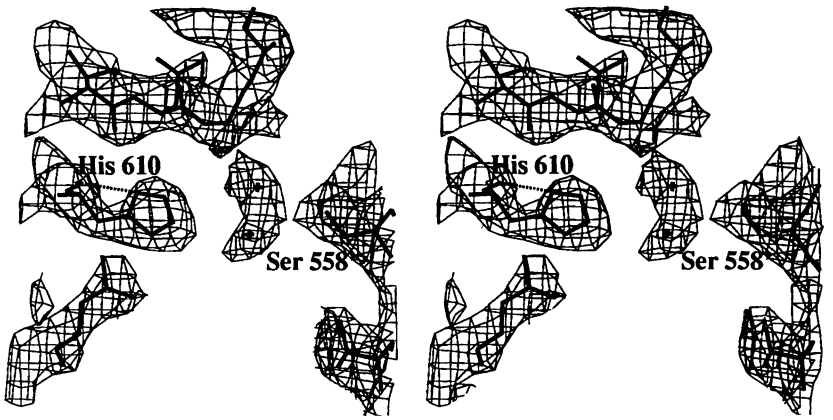

Fig. 2. Quality of the MIRAS electron density map calculated with coefficients $m|F_{obs}|\exp(i\alpha_{MIRAS})$, where m is the figure of merit. The region shown comprises the residues in the active center. The map is calculated with phases obtained from all three heavy-atom derivatives (Table 1) including anomalous data. The unusual conformation of His^{610}, which allows a hydrogen bond between Nδ1 and the carbonyl O of the same residue, is emphasized by the dashed line connecting the two atoms.

which, unlike complexes from other organisms, dissociates into smaller oligomers upon binding of E1p or E3 or both (*19*).

Attempts to crystallize the entire *A. vinelandii* E2p molecule failed to produce crystals suitable for high-resolution diffraction studies, probably because of the high flexibility of the lipoyl domains. However, the cloning and overexpression of the catalytic domain of *A. vinelandii* E2p allowed the growth of good-quality crystals (*20*) that enabled us to elucidate the atomic structure of the core of the complex, which has the rather uncommon octahedral 432 symmetry.

Structure determination. The E2p catalytic domain (E2pCD) was expressed as a fusion protein that contained: (i) the catalytic domain (residues 409 to 638 of *A. vinelandii* E2p); (ii) the Ala-Pro–rich sequence (residues 382 to 408), which connects this part of the structure to the E1p-E3 binding domain; and (iii) six amino acids from β-galactosidase at the NH_2-terminus (*20*). A multiple isomorphous replacement map including anomalous data (MIRAS) with an average figure of merit of 0.74 (Table 1) was calculated which allowed an unambiguous tracing of most of the polypeptide chain (Fig. 2). A model was built by using the fragment search option of the program O (*21*), and the initial structure was improved by crystallographic refinement with the energy minimization option of the program X-PLOR (*22*) and by manual model rebuilding. The current model contains highly restrained individual temperature factors and coordinates for 2182 protein atoms and 37 water molecules. The crystallographic R factor is 0.188 for 10,336 independent reflections between 10.0 Å and 2.6 Å resolution. The root-mean-square (rms) deviations from the ideal values for bonds and angles are 0.017 Å and 3.3°, respectively. All backbone dihedral angles (φ, ψ pairs) fall within the allowed regions, except for Ala^{156}, for which the φ angle is 10° outside an allowed region.

The first 19 residues are not visible in the electron density map probably because they are disordered. The first 6 amino acids belong to the fusion sequence, whereas the next 13 belong to the Ala-Pro–rich sequence at the NH_2-terminus of E2pCD, thereby confirming the high level of flexibility of this part of the molecule. Since the amino acids of β-galactosidase present at the NH_2-terminus of the fusion protein are not visible in the electron density map, we use the residue numbering of *A. vinelandii* E2p to describe the structure.

The tight association of the trimer. Each subunit of E2pCD consists of a mixed β sheet composed of three central parallel strands (strands E, G, and I) and three antiparallel ones (strands F, J, and B) at its borders. Moreover, a threefold-related subunit provides a seventh additional strand (strand H), that is adjacent and antiparallel to strand B. This large β sheet is flanked on one side by helix H1 and on the other by the remaining four α helices to generate a kind of open face sandwich structure (Fig. 3A). In total the individual subunit contains five α helices, one 3_{10}-helix (H6), and ten β strands, which comprise 64 percent of the total number of residues. A most remarkable feature is the NH_2-terminal "elbow" formed by residues 395 to 415, which has no nonbonded contacts with the single globular domain of an individual subunit.

Each monomer is tightly associated with two other subunits related by the threefold rotation axis (Fig. 3B). The intersubunit interactions are extensive as is indicated by the fact that 3900 $Å^2$, or 25 percent, of the monomer solvent accessible surface is buried upon trimerization, involving 12 intersubunit hydrogen bonds. Out of 243 residues, a total of 66 of each monomer has one or more atoms contributing to the intratrimer interface. In particular, several hydrophobic side chains (Ile^{395}, Trp^{429}, Leu^{430}, Phe^{438}, Thr^{566}, and Pro^{589}) are involved in the interactions along the threefold axis and are in van der Waals contact with their symmetry-related partners. The cylinder-shaped trimer has a

Table 1. Data collection, multiple isomorphous replacement, and refinement statistics. Crystals were obtained by vapor diffusion at pH 7.0 with 18 percent ammonium sulfate as a precipitant with the use of macroseeding technique (space group *F*432, cell axis 224.8 Å). There are four 24-subunit complexes in the unit cell, which correspond to one subunit per asymmetric unit and a solvent content of 73 percent (*45*). Diffraction data were collected on a FAST television area detector and evaluated by the software MADNES followed by profile fitting (*46*). One Hg and two Pt derivatives were used for phasing. The heavy-atom positions were located with the direct methods option of SHELXS-86 (*47*) and by a vector search program (W.G.J.H.). The mercury acetate (HgAc) derivative had one site, and the $PtCl_4$ and $Pt(NO_3)_3(NH_3)_2$ (PtNIt) derivative each had two sites. The heavy-atom parameters were refined including anomalous data by the program PHARE written by G. Bricogne. In particular, the Hg derivative was of excellent quality. The final model includes 2182 protein atoms and 37 solvent molecules, and has an R factor of 0.188 at 2.6 Å resolution and an average B factor of 16.0 $Å^2$. There are 73 main-chain hydrogen bonds per 100 residues (*48, 49*). The rms deviations from ideality are 0.017 Å for bond lengths and 3.3° for bond angles.

Parameter	Native	HgAc	$PtCl_4$	PtNIT
Measured reflections	68,417	52,164	48,771	47,572
Unique reflections	10,391	9,581	8,891	8,810
Completeness to 3.0 Å	0.779	0.931	0.881	0.870
Completeness to 2.6 Å	0.721			
R_{merge}*	0.025	0.031	0.035	0.037
R_{nat}†		0.210	0.197	0.099
$<F_H>/E$‡		3.1	1.5	1.7
R_{cullis}§		0.482	0.741	0.622

*The merging R factor is defined as $R^{merge} = \Sigma|I_i - <I>|/\Sigma<I>$. †$R_{nat}$ is given by $R_{nat} = \Sigma||F_{PH}| - |F_N||/\Sigma|F_N|$, where F_{PH} and F_N are the measured structure factors for the heavy-atom derivative and the native crystals, respectively. ‡$<F_H>/E$ is the "phasing power" calculated as (rms calculated F_H)/(rms lack of closure error), where F_H is the heavy-atom structure factor. §The Cullis R factor is defined as $R_{cullis} = \Sigma||F_{PH} \pm F_P| - |F_H||/\Sigma|F_{PH} - F_P|$, only for centric reflections.

height of ~46 Å along the threefold axis and a radius of 27 Å.

An important contribution to the intratrimer interactions comes from the NH_2-terminal "elbow," which extends over a neighboring monomer (Fig. 3A). The first part of this "elbow" comprises a pentapeptide (396 to 400) in an extended conformation, with four trans Pro residues (Pro^{396}, Pro^{397}, Pro^{399}, and Pro^{400}). After a sharp "elbow" bend of 120°, residues 409 to 412, form a short β strand (strand A) that, together with residues 486 to 498 (strands C and D; Fig. 3C) of a threefold-related subunit, generates a small β sheet located on the outside of the molecule.

Comparison with chloramphenicol acetyl transferase (CAT) and location of the catalytic center. The topology of E2pCD is virtually identical to that of CAT (*23*), confirming a prediction by Guest (*24*) based on a sequence alignment, which showed 19 percent amino acid identity mainly localized on the residues around the catalytic center. A structure-based sequence alignment (Fig. 3C) shows that, among the numerous insertions and deletions, a noticeable feature is the significantly shorter NH_2-terminal helix (residues 417 to 432) in CAT. Most remarkably, these helices occupy virtually the same position in the trimers of E2pCD

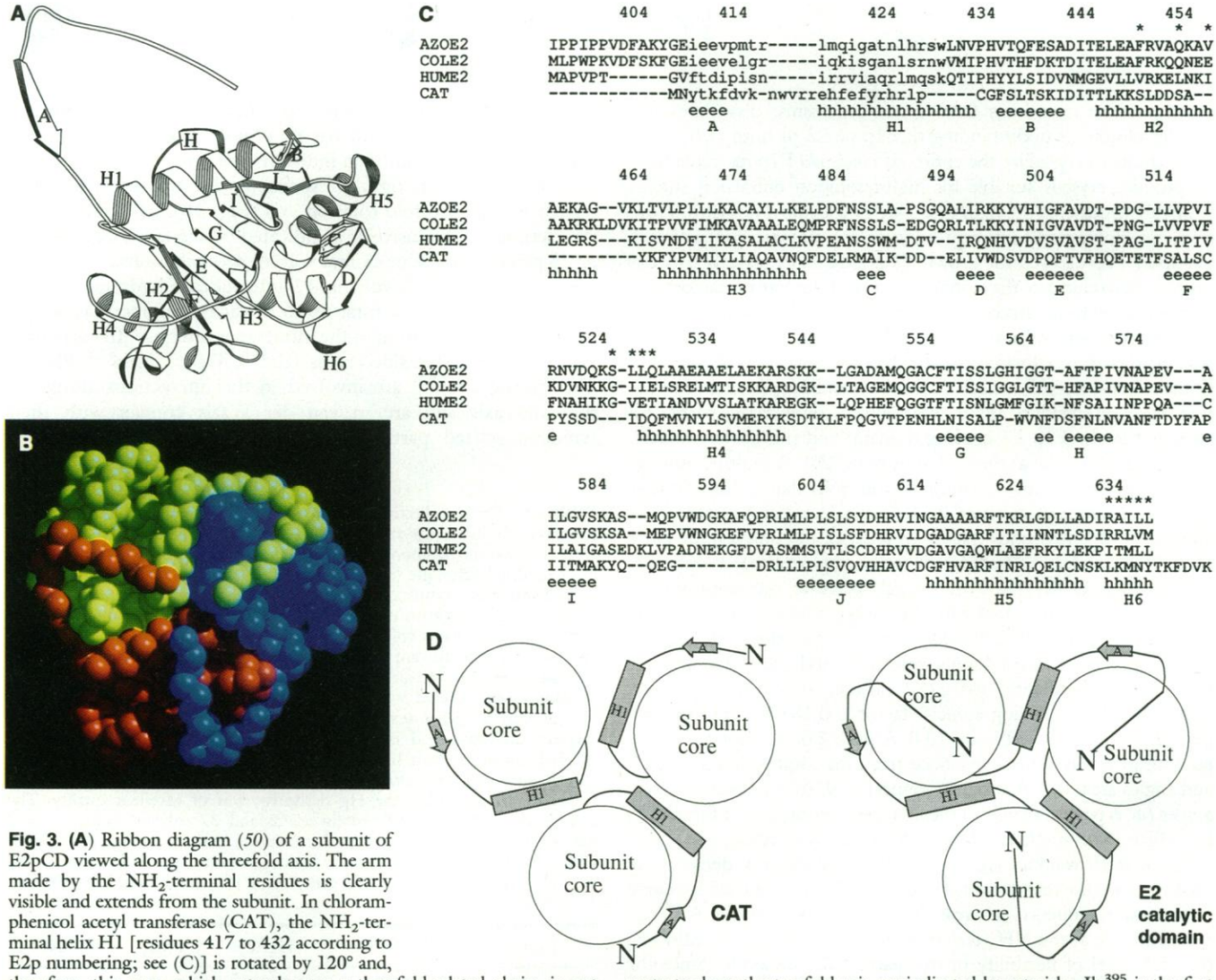

Fig. 3. (A) Ribbon diagram (*50*) of a subunit of E2pCD viewed along the threefold axis. The arm made by the NH_2-terminal residues is clearly visible and extends from the subunit. In chloramphenicol acetyl transferase (CAT), the NH_2-terminal helix H1 [residues 417 to 432 according to E2p numbering; see (C)] is rotated by 120° and, therefore, this arm, which extends over a threefold-related chain, is not present in CAT. The central β sheet of the E2pCD subunit (strands B, E, F, G, J, and I) is extended by strand H′, which is on a threefold-related monomer and runs antiparallel to B. In an analogous way, strand A forms a β sheet together with strands C″ and D″ from another threefold-related subunit. (**B**) Space-filling model (only Cα atoms) of a trimer of subunits viewed along the threefold axis with the red subunit in the same orientation as in (A). Note the close association of the subunits along the threefold axis. The direction of view is from the outside toward the center of the E2p 24-mer. (**C**) Structure-based alignment of the sequences of the *A. vinelandii* (residues 395 to 637), *E. coli* (*13*), and human (*51*) E2pCD and of CAT (*52*). For the structural comparisons, the coordinates of CAT deposited in the Brookhaven Protein Data Bank (3CLA) have been used. Residues belonging to α helices and β strands according to the program DSSP (*49*) are outlined by the letters h and e, respectively, and are named following the scheme used for CAT (*23*). The amino acids involved in the intertrimer contacts along the twofold axis are indicated by asterisks. Ile^{395} is the first amino acid visible in the electron density map. This alignment is strictly speaking only valid as of residue 434. Amino acids from E2pCD and CAT before this residue are not structurally equivalent when one single subunit of E2pCD is used for the superposition. However, when the NH_2-terminal arm (residues 409 to 427) of an adjacent subunit of E2pCD is superimposed onto CAT, then the rms deviation is 1.7 Å for these 19 positionally equivalent residues. These residues are indicated by lowercase in the alignment. Only one of these 21 residues, Leu^{425} in E2pCD, is identical in the two proteins. (**D**) Schematic drawing showing the totally different organization of the NH_2-terminal strand A (residues 409 to 412) and helix H1 (416 to 432) in CAT with respect to E2pCD. In E2pCD the NH_2-terminal residues form an arm connecting two threefold-related subunits, whereas in CAT they curl back and interact with amino acids of their own monomer. The trimers are viewed along the threefold axis from the outside toward the center of the cubic core in the same orientation as in (B).

and CAT, but belong to different subunits. In CAT the NH_2-terminal helix is rotated by 120° with respect to the corresponding helix of the E2pCD (Fig. 3D). Comparison of the helices with the use of the operations obtained by superposition of the globular domains of E2pCD and CAT yields an rms deviation of 1.7 Å for 19 residues, with only one identical side chain (Leu^{425}).

Overall, the rms displacement between 121 superimposable Cα atoms of E2pCD and CAT is 1.6 Å. The homology can be used to identify residues likely be involved in substrate binding and catalysis, which are located at the trimer interface and involve amino acids of two threefold-related subunits. Particularly, His^{195} and Ser^{148}, known to be essential for CAT activity (*25*), are conserved in E2pCD (Ser^{558} and His^{610}), as predicted by Guest *et al.* (*7, 24*). The distance between Oγ of Ser and Nε2 of His is 7.1 Å in CAT and 6.9 Å in E2pCD. These residues are in the middle of a 29 Å long channel, which provides the binding pockets for the two substrates (Fig. 4). The walls of this "catalytic" channel are defined by strands E, F, G, and H of one subunit and helix H5 of a threefold-related one. Mutagenesis experiments have established the essential role played in catalysis by Ser^{558} (homologous to 148 of CAT) in *Escherichia coli* E2p (*26*), where replacement of Ser by Ala resulted in a significant decrease in the enzyme activity.

The structural similarity between His^{610} of E2pCD and His^{195} of CAT is striking. In both enzymes (Fig. 2), this amino acid has an unusual conformation, with (ϕ, ψ) values of (−61°, −19°) and a (χ_1, χ_2) combination of (−155°, −7°) in E2pCD, whereas in CAT these dihedrals are (−67°, −21°) and (−148°, −32°), respectively. Such a conformation allows a hydrogen bond between the imidazole nitrogen Nδ1 and the carbonyl O of the same amino acid. In CAT this interaction has been related to the role played by this residue in catalysis (*23, 25*) and its conservation is suggestive of an analogous function also in E2pCD. This interpretation is supported by the mutagenesis studies on the *E. coli* enzyme (*27*) but not by those on the yeast molecule (*28*), where the replacement of this His by Ala does not affect the activity.

In spite of these similarities, the catalytic centers of E2pCD and CAT also show several significant differences. Among them, it is quite remarkable that the salt bridge Asp^{199}-Arg^{18}, known to be functionally and structurally essential in CAT (*29*), is absent in E2pCD because Asp^{199} of CAT is replaced by Asn^{614} and Arg^{18} of CAT does not have a counterpart in E2pCD because of a deletion between strand A and helix H1 (Fig. 3C). Furthermore, in E2pCD the residues flanking the catalytic His (Asp^{609} and Arg^{611}) form a salt bridge (Fig. 4), which might be of functional relevance, because the tripeptide Asp-His-Arg is conserved in all E2 sequences (*16*). This feature is not observed in CAT, where Asp^{609} and Arg^{611} of E2pCD are replaced by His and Ala, respectively (Fig. 3C).

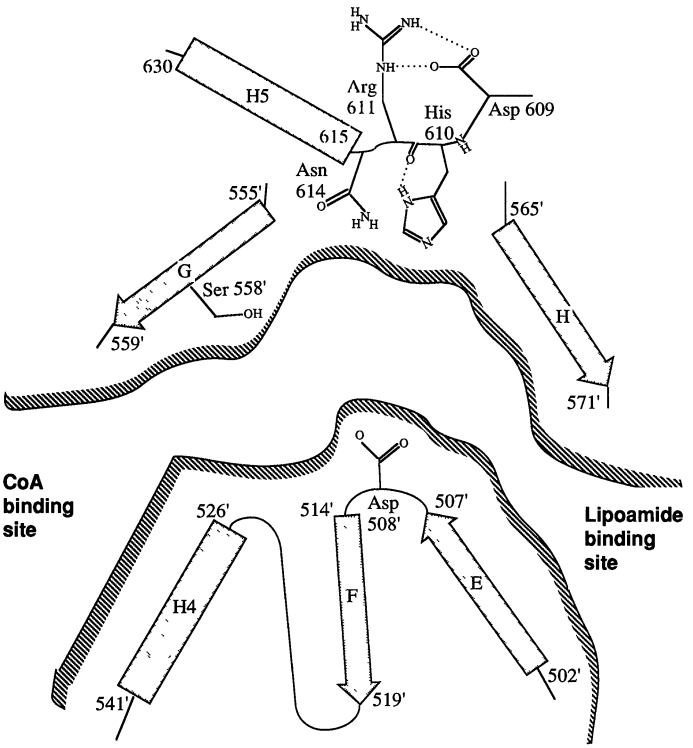

Fig. 4. Schematic representation of E2pCD catalytic center. The 29 Å long channel goes across the subunit trimer, with the CoA and lipoamide binding sites located at the two opposite entrances, as derived by analogy with CAT. The similarity between the two enzymes is particularly evident for the CoA binding region. The sequence identity between the residues forming the walls of this part of the channel is 27 percent. However, the different specificity for the second substrate involved in the transacetylase reaction, dihydrolipoamide in E2pCD and chloramphenicol in CAT, leads to a much lower degree of similarity for the other side of the channel, indicated by the lack of any sequence identity between the residues forming the lipoamide-chloramphenicol binding pocket.

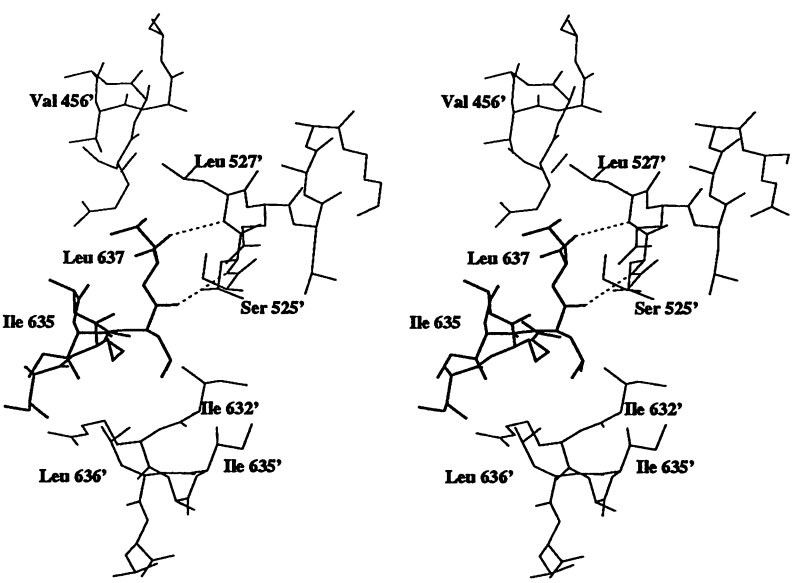

Fig. 5. Stereo drawing of the interactions made by the COOH-terminal residues with the amino acids of the twofold-related subunit, viewed approximately along the twofold axis. The hydrogen bonds are shown by the dashed lines. Most of the side chains involved in these types of contacts are hydrophobic. Note that the COOH-terminal helix 632 to 637 of one subunit (shown in fat lines) forms a "knob" fitting into a "hole" formed by the other twofold-related subunit whose residues numbers are indicated with primes.

The truncated cube. The E2pCD of *A. vinelandii* E2pCD can assemble into oligomers of 24 subunits arranged with 432 symmetry. In contrast to the extensive intratrimer contacts, the intertrimer interactions, which lead to the generation of the oligomer, are based on a quite limited number of contacts (Figs. 5 and 6B). No subunit interactions across the fourfold axis are present. The association of the trimers along the twofold axis buries only 8 percent (1150 $Å^2$) of the monomer-accessible surface. The residues of the two subunits involved in this type of intersubunit interactions are: (i) the COOH-terminal residues 632 to 637, which form a 3_{10}-like helix (H6); (ii) the NH_2-terminal part of the α helix H4; and (iii) the side chains of Phe^{449}, Gln^{453}, and Val^{456} along helix H2. Only two intertrimer hydrogen bonds are formed (see Fig. 5), which involve the carbonyl O of Leu^{636} and the negatively charged carboxylate O of the COOH-terminal residue Leu^{637}. These atoms interact with the main-chain nitrogens of Leu^{526} and Leu^{527} located at beginning of helix H4 of the opposite subunit. Because of the direction of the helix dipole (*30*) and the negative charge of the carboxylate, this interaction appears to be electrostatically favorable.

The side chains of the dimer interface are hydrophobic in character (Fig. 5). The COOH-terminal residues 632 to 637 form a sort of hydrophobic "knob" that is buried into a "hole" of the twofold-related subunit. The chemical nature of these side chains is conserved among all known E2p sequences whether they form octahedral or dodecahedral arrangements [Fig. 3C (*16*)]. Moreover, a similar arrangement is likely to be maintained in the entire class of the oxo-acid dehydrogenase complexes because the COOH-terminal residues are conserved in all of the transferase sequences (*16*).

The assembly of eight identical trimers leads to the generation of the 24-subunit complex, which corresponds to the inner core of the PDC. The shape of the multisubunit structure is that of a truncated cube whose corners are made by the trimers, which only partially fill the volume of the cube. They generate a hollow cage whose internal cavity is connected to the outside through large channels crossing the faces of the cube (Fig. 6A). Overall, less than 50 percent of its volume is occupied by protein atoms. The length of edge of the cube is 125 Å, but due to its truncation, the distance from corner (taking the Cα of Ile^{395}) to center is only 80 Å, instead of 108 Å expected for an "untruncated" cube. The radius of the channels going through the faces is ~30 Å. The largest sphere inside the cube that does not touch any atom has a radius of 23 Å.

Interestingly, whereas the entrance of the acetyllipoamide binding site points toward the outside of the oligomeric truncated cube structure, CoA has to approach its binding pocket from the internal cavity of the oligomer.

The tightly associated trimer appears to be the actual "building block" for the entire cubic core because the interactions at the trimer interface are by far the most extensive ones in the oligomer. The isolated 24-subunit complex of *A. vinelandii* E2p is known to dissociate into smaller suboligomers upon binding of the peripheral

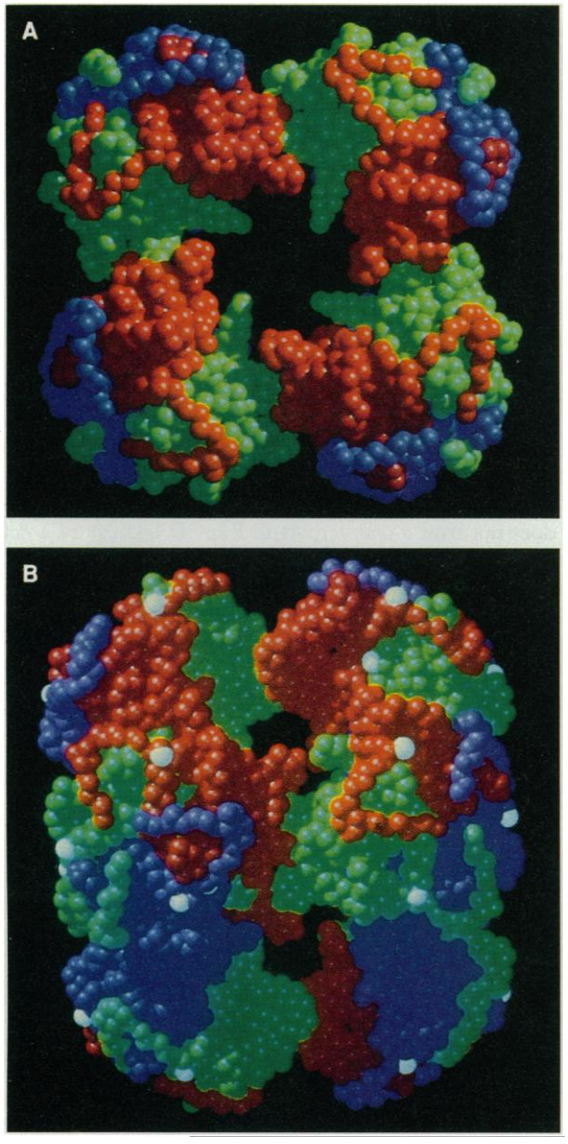

Fig. 6. The 24-subunit truncated cube forming the core of the *A. vinelandii* pyruvate dehydrogenase complex. (**A**) Space-filling representation (only Cα atoms) of the 24 E2pCD molecules (eight trimers), as viewed along the fourfold axis, obtained by applying the fourfold- and twofold-symmetry operations to the trimer shown in Fig. 3B. The distance left to right in this view is ~125 Å. (**B**) The same as in (A) but viewed along the twofold axis. The white spheres indicate the position of Arg^{416}, which according to mutagenesis studies (*32*) is involved in E1p binding.

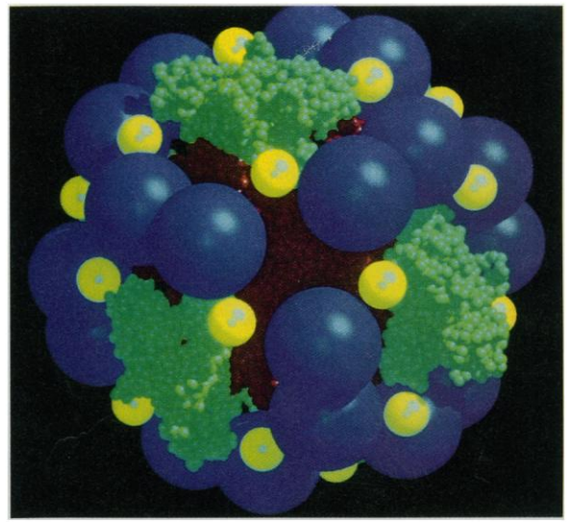

Fig. 7. Model of the cubic-octahedral PDC viewed along the threefold axis. In red and green, space-filling models (only Cα atoms) of *A. vinelandii* E2pCD and E3 are shown, respectively. The structure of E1p is approximated by a dimer of spheres (blue) with a radius of 29 Å, whereas each lipoyl domain is indicated by a sphere (yellow) with a radius of 17 Å. For clarity, the other two lipoyl domains of *A. vinelandii* E2p as well as the E2p-E3 binding domain have been omitted. Figures 3B, 6A, 6B, and 7 were generated by the program Raster3d (*53*).

components E1p or E3, or both (*19*). In view of the number and type of interactions at the trimer interface, the products of dissociation are probably trimers.

After ferritin (*31*), E2pCD provides the second multisubunit structure with 432 symmetry known at atomic level. The organization of the two quaternary structures appears, however, to be very different. In E2pCD the strongest intersubunit interactions are along the threefold axis with no contacts along the fourfold, but in ferritin the most extensive interactions are along the twofold, with weaker associations along the threefold and fourfold axes.

Toward an atomic model of a multienzyme complex. In *A. vinelandii* E2p, removal of the E1p-E3 binding domain, which is located at the NH_2-terminus of the catalytic domain and not present in our structure, prevents the binding of the peripheral components of the multienzyme complex. However, mutagenesis experiments on *A. vinelandii* E2p show a dramatic decrease in the affinity for E1p caused by the replacement of Arg^{416} at the NH_2-terminus of the catalytic domain by an Asp (*32*), suggesting a direct contact between this side chain and the E1p molecule. This result is in agreement with studies in other related organisms [for example, in *E. coli* PDC; see (*7*) and references therein] in which the E1p-E3 binding domain seems to be necessary only for E3 and not for E1p binding, indicating that the latter enzyme must interact also with residues of the catalytic domain. Arg^{416} is located (Fig. 6B) on the edge of the truncated cube, at a distance of 75 Å from the center of the oligomer. This distance agrees well with results obtained by mass-distribution analysis of the BCODH complex, which is closely related to the PDC (*33*), demonstrating that the inner radial limit for E1p mass is ~80 Å.

Models derived from "negatively stained" electron microscopy images show that the 12 copies of E1p present in the PDC bind along the edges of the cube-like core (*4*). E1p is a dimer of identical subunits with a molecular weight of 2 × 90,000. Its atomic structure is not known, but for the present purposes, by assuming a packing density of 0.85 dalton per cubic angstrom, the dimer can be approximated by two spheres with a radius of 29 Å each. In order to define their location with respect to the PDC core, we used the position of Arg^{416} as a point of contact and, as a second criterion, the condition that E1p should not block the catalytic channel of E2p. This approach leads to the model shown in Fig. 7, where the 12 pairs of spheres, approximating the E1p molecules, are situated at the edges of the cube, with their twofold axes coinciding with that of the 24-subunit complex, as derived by Reed and co-workers (*4*).

Electron microscopy analysis of PDC (*4, 34*) and the low-resolution studies of crystals of E2p/E3 subcomplexes (*35*) led to a model where the six E3 dimers sit on the centers of each face of the cubic core. Taking this as a starting point, a tentative model of the E2p/E3 association can be generated, because the atomic structure of the *A. vinelandii* E3 dimer (molecular mass 2 × 53,000 daltons) has been determined at 2.2 Å resolution (*36*) in our laboratories. First, the molecular twofold axis has been aligned with the fourfold axis of the PDC core. Next, the orientation and position that allow the minimum distance between the centers of gravity of E2p and E3, while keeping all interatomic contacts longer than 4.0 Å, have been searched (*37*). This analysis leads to the structure shown in Fig. 7. In this tentative model the highly mobile lipoyl domains, as spheres with radii of 17 Å, are also included to show that the space required for reaching and coupling the catalytic centers on the various enzymes is available.

In the model, the distance between a catalytic center of E2p and the nearest E3 active site is 45 Å, whereas the separation between E2p catalytic centers related by a threefold axis is significantly shorter (34 Å). This result confirms spectroscopic observations (*38*), which indicate a distance of at least 40 Å between the catalytic sites of different enzymes. Moreover, it suggests that a single lipoyl group covalently attached to a lipoyl domain is capable of coupling the active centers of all three subunits on the corner of the cube. This observation substantiates the model of a random-coupling mechanism (*39*), proposed on the bases of a wealth of biochemical and mutagenesis data [for reviews, see (*4, 5, 7, 8*)].

Images obtained by cryoelectron microscopy (*40*) seem to indicate that the peripheral components are well separated (40 to 50 Å) from the core of the complex. Although for E1p this result seems to be contradicted by the above-mentioned mutagenesis results, it might be indicative of a degree of flexibility in the mode of binding (*41*) of E3 to the complex.

A model for the dodecahedral inner core of human PDC. The core of the PDC of mammals, yeast, fungi, and the Gram-positive bacterium *Bacillus stearothermophilus* consists of 60 E2p subunits arranged in a dodecahedronlike structure with 532 symmetry (*4, 5*). The mammalian and yeast systems also possess additional components that are involved in the regulation (kinase and phosphorylase) and possibly in the assembly (the so-called protein X) of the complex (*42*). Sequence analysis reveals that, in spite of the different quaternary structure, the homology between the human (a prototype for the dodecahedral arrangement) and the *A. vinelandii* E2p is significant [28 percent sequence identity; Fig. 3C (*16*)]. Therefore, the conservation of the protein fold and a close similarity of the extensive molecular interactions along the threefold axis can be predicted for dodecameric PDC's. Furthermore, the conserved active-site residues (Fig. 3C) are located at the trimer interface, further substantiating the assumption that the trimer provides the "building block" not only for the 24-subunit complex but also for the 60-subunit complex.

As mentioned by Oliver and Reed (*4*), this similarity of interactions in multimers with different point-group symmetry is related to the concept of "quasi-equivalence," initially developed by Caspar and Klug (*43*). We propose that the modifications in the intertrimer interactions along the twofold axis required to construct a dodecahedral oligomer from a cubic one can be accommodated without major structural modifications of individual subunits. It is likely that the "knob" formed by the COOH-terminal helix fitting into the

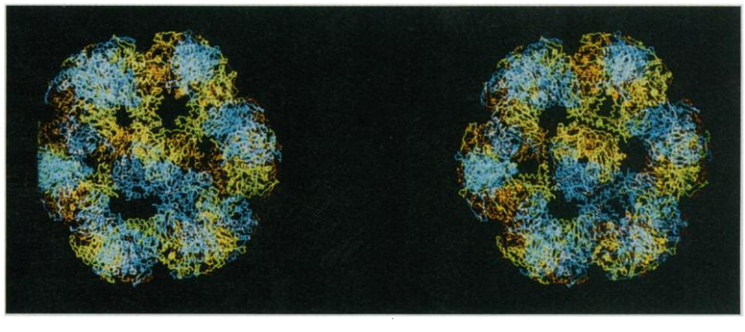

Fig. 8. Stereo picture of the dodecahedral complex (only Cα trace) viewed along a threefold axis obtained by modifying the mutual orientation and position of a pair of twofold-related trimers in order to make them compatible with the 532 symmetry, with the restrain of causing the least possible changes at the twofold interface. This approach led to a rotation of one trimer with respect to the other by 28°. Because of "knob-into-hole" architecture of the subunit association along the twofold axis (Fig. 5), this reorientation can be readily accommodated without the requirement of major conformational changes. The two trimers have been then expanded to the truncated dodecahedron by applying the 532 symmetry operators. The three chains forming each trimer are colored in blue, yellow, and green.

"hole" provided by a twofold-related subunit (Fig. 5) can accommodate the rotations required for building a truncated dodecahedron from a truncated cube. In addition, the "knob" itself may rotate slightly with respect to the main body of the subunit to which it is covalently connected.

A model for the dodecahedral 60-subunit oligomer is shown in Fig. 8. As for the cube-like structure, its shape is that of a hollow cage, whose volume is only partially filled by protein atoms. The inner radius of the dodecahedron is ~100 Å, whereas the "holes" in the pentameric faces have a radius of 35 Å. These dimensions are in good agreement with those observed by electron microscopy (*44*).

REFERENCES AND NOTES

1. L. J. Reed, *Acc. Chem. Res.* 7, 40 (1974).
2. S. L. Yeaman, *Trends. Biochem. Sci.* **11**, 293 (1986).
3. M. S. Patel and D. J. Carothers, *Endocrine and Biochemical Development of Fetus and Neonate* (Plenum, New York, 1990).
4. R. M. Oliver and L. J. Reed, in *Electron Microscopy of Proteins*, J. R. Harris, Ed. (Academic Press, New York, 1982), vol. 18, pp. 77–95.
5. L. J. Reed and M. L. Hackert, *J. Biol. Chem.* **265**, 8971 (1990).
6. M. S. Patel and T. E. Roche, *FASEB J.* **4**, 3225 (1990).
7. J. R. Guest, S. J. Angier, G. C. Russel, *Ann. N.Y. Acad. Sci.* **573**, 76 (1989).
8. R. N. Perham, *Biochemistry* **30**, 8501 (1991).
9. L. C. Packman, G. Hale, R. N. Perham, *EMBO J.* **3**, 1315 (1984).
10. A. de Kok and A. H. Westphal, *Eur. J. Biochem.* **152**, 35 (1985).
11. L. J. Reed *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **72**, 3068 (1975).
12. D. M. Bleile, P. Munk, R. M. Oliver, L. J. Reed, *ibid.* **76**, 4385 (1979).
13. P. E. Stephens, M. G. Darlison, H. M. Lewis, J. R. Guest, *Eur. J. Biochem.* **133**, 481 (1983).
14. R. N. Perham, H. W. Duckworth, G. C. K. Roberts, *Nature* **292**, 474 (1981).
15. J. S. Miles, J. R. Guest, S. E. Radford, R. N. Perham, *J. Mol. Biol.* **202**, 97 (1988).
16. G. C. Russell and J. R. Guest, *Biochim. Biophys. Acta* **1076**, 225 (1991).
17. G. Burns, P. J. Sykes, K. Hatter, J. R. Sokatch, *J. Bacteriol.* **171**, 665 (1989).
18. R. Hanemaaijer, A. H. Westphal, T. Van der Heiden, A. de Kok, C. Veeger, *Eur. J. Biochem.* **179**, 287 (1988).
19. H. J. Bosma, A. de Kok, A. H. Westphal, C. Veeger, *ibid.* **141**, 541 (1984).
20. E. Schulze *et al.*, *ibid.* **201**, 561 (1991).
21. T. A. Jones, J. Y. Zou, S. W. Cowan, M. Kjeldgaard, *Acta Crystallogr.* **A47**, 110 (1991).
22. A. T. Brünger, J. Kuriyan, M. Karplus, *Science* **235**, 458 (1987).
23. A. G. W. Leslie, *J. Mol. Biol.* **213**, 167 (1990).
24. J. R. Guest, *FEMS Microbiol. Lett.* **44**, 417 (1987).
25. A. Lewendon, I. A. Murray, W. Shaw, M. R. Gibbs, A. G. W. Leslie, *Biochemistry* **29**, 2075 (1990).
26. G. C. Russell and J. R. Guest, *Proc. R. Soc. London* **243**, 155 (1991).
27. ———, *Biochem. J.* **269**, 443 (1990).
28. X. D. Niu, J. K. Stoops, L. J. Reed, *Biochemistry* **29**, 8614 (1990).
29. M. R. Gibbs, P. C. E. Moody, A. G. W. Leslie, *ibid.*, p. 11261.
30. W. G. J. Hol, P. T. van Duijnen, H. J. C. Berendsen, *Nature* **273**, 443 (1978).
31. S. H. Banyard, D. K. Stammers, P. M. Harrison, *ibid.* **271**, 282 (1978); D. M. Lawson *et al.*, *ibid.* **349**, 541 (1991).
32. E. Schulze, A. H. Westphal, H. Boumans, A. de Kok, *Eur. J. Biochem.* **202**, 841 (1991).
33. M. L. Hackert *et al.*, *Biochemistry* **28**, 6816 (1989).
34. H. Yang, J. F. Hainfeld, J. S. Wall, P. A. Frey, *J. Biol. Chem.* **260**, 16049 (1985).
35. C. C. Fuller, L. J. Reed, R. M. Oliver, M. L. Hackert, *Biochem. Biophys. Res. Commun.* **90**, 431 (1979).
36. A. Mattevi, A. J. Schierbeek, W. G. J. Hol, *J. Mol. Biol.* **220**, 975 (1991).
37. The assumption that the E3 twofold axis coincides with the fourfold axis of the PDC inner core reduces the definition of E3 position and orientation to a three-parameter search: (i) the rotation (between 0° and 90°) about the fourfold axis of the core; (ii) the translation along the fourfold; and (iii) the rotation of the E3 dimer about an axis perpendicular to its twofold (which can differ by 180°). In order to analyze the latter, we refer to the "lower" and "upper" regions into which the dimer interface of lipoamide dehydrogenase can be divided [see figure 7 of (*36*)]. Orienting the "lower" E3 interface region toward the face of the truncated cube allows the shortest distance between the centers of gravity of E2p and E3, and on this basis it is preferred to the alternative with the "upper" interface toward the cube. However, the choice of E3 orientation has little effect on the separation between E3 and E2p catalytic centers, which in both cases is ~45 Å, taken as the distance between the reactive His^{450} of *A. vinelandii* E3 and His^{610} of E2pCD.
38. G. Shepherd and G. G. Hammes, *Biochemistry* **16**, 5234 (1977).
39. M. L. Hackert, R. M. Oliver, L. Reed, *Proc. Natl. Acad. Sci. U.S.A.* **80**, 2907 (1983).
40. T. Wagenknecht, R. Grassucci, D. Schaak, *J. Biol. Chem.* **36**, 22402 (1990).
41. H. J. Grande *et al.*, *Catalysis in Chemistry and Biochemistry. Theory and Experiment*, B. Pullman, Ed. (Reidel, Dordrecht, 1979), pp. 171–190.
42. J. A. Hodgson, O. de Marcucci, J. G. Linsay, *Eur. J. Biochem.* **158**, 595 (1986); J. M. Jilka, M. Rahmatullah, M. Kazemi, T. E. Roche, *J. Biol. Chem.* **263**, 1858 (1986).
43. D. L. Caspar and A. Klug, *Cold Spring Harbor Symp. Quant. Biol.* **27**, 1 (1962).
44. C. E. Henderson, R. N. Perham, J. T. Finch, *Cell* **17**, 85 (1979).
45. These crystals, although grown under different conditions, have virtually the same cell dimensions and space group as those of *Escherichia coli* dihydrolipoyl transsuccinylase obtained by De Rosier *et al.* [D. J. DeRosier, R. M. Oliver, L. J. Reed, *Proc. Natl. Acad. Sci. U.S.A.* **68**, 1135 (1971)], reflecting the same oligomeric arrangement of the PDC and the 2-oxoglutarate dehydrogenase multienzyme complexes.
46. W. Kabsch, *J. Appl. Crystallogr.* **21**, 916 (1988); A. Messershmidt and J. W. Pflugrath, *ibid.* **20**, 306 (1987).
47. G. M. Sheldrick, *Crystallographic Computing*, G. M. Sheldrick, C. Krüger, R. Goddard, Eds. (Oxford Univ. Press, Oxford, 1985), vol. 3, pp. 175–189.
48. The number of main-chain hydrogen bonds per 100 residues is calculated by the program DSSP (*49*). The value of 73 is similar to that observed in other protein molecules.
49. W. Kabsch and C. Sander, *Biopolymers* **22**, 2577 (1983).
50. J. P. Priestle, *J. Appl. Crystallogr.* **21**, 515 (1988).
51. T. J. Thekkumkara *et al.*, *FEBS Lett.* **240**, 45 (1988).
52. Abbreviations for the amino acid residues are: A, Ala; C, Cys; D, Asp; E, Glu; F, Phe; G, Gly; H, His; I, Ile; K, Lys; L, Leu; M, Met; N, Asn; P, Pro; Q, Gln; R, Arg; S, Ser; T, Thr; V, Val; W, Trp; and Y, Tyr.
53. The Preras3d/Raster3d package is written by D. Bacon, A. M. Berghuis, M. Israel, and M. E. P. Murphy.

54. We thank T. Smale and A. J. Schierbeek for their initial contributions to this project. The advice of T. Sixma and B. Strokopitov in the use of the heavy-atom refinement and direct methods programs is gratefully acknowledged. The continuous help and stimulating discussions with all of the members of the Groningen Protein Crystallography Group have been of invaluable importance for the determination and analysis of the structure. Supported by the Netherlands Foundation for Chemical Research (SON) with financial aid from the Netherlands Organization for Scientific Research (NWO).

31 October 1991; accepted 31 January 1992