Protein Science (1995), *4*:1291–1304. Cambridge University Press. Printed in the USA.

Modeling of the spatial structure of eukaryotic ornithine decarboxylases

NICK V. GRISHIN,[1] MARGARET A. PHILLIPS,[1] AND ELIZABETH J. GOLDSMITH[2]

[1] Department of Pharmacology, University of Texas Southwestern Medical Center, Dallas, Texas 75235
[2] Department of Biochemistry, University of Texas Southwestern Medical Center, Dallas, Texas 75235

(RECEIVED March 16, 1995; ACCEPTED April 21, 1995)

Abstract

We used sequence and structural comparisons to determine the fold for eukaryotic ornithine decarboxylase, which we found is related to alanine racemase. These enzymes have no detectable sequence identity with any protein of known structure, including three pyridoxal phosphate-utilizing enzymes. Our studies suggest that the N-terminal domain of ornithine decarboxylase folds into a β/α-barrel. Through the analysis of known barrel structures we developed a topographic model of the pyridoxal phosphate-binding domain of ornithine decarboxylase, which predicts that the Schiff base lysine and a conserved glycine-rich sequence both map to the C-termini of the β-strands. Other residues in this domain that are likely to have essential roles in catalysis, substrate, and cofactor binding were also identified, suggesting that this model will be a suitable guide to mutagenic analysis of the enzyme mechanism.

Keywords: alanine racemase; β/α-barrel; decarboxylase; folding; molecular evolution; ornithine; sequence analysis; structure prediction

Pyridoxal phosphate-dependent ornithine decarboxylase is an important target for therapeutic intervention in the treatment of cancer and parasitic infections (McCann & Pegg, 1992). The structure of the eukaryotic enzyme or of any of its homologues has not been determined.

Similarities in sequence, secondary structure, and hydrophobicity profiles have been used to classify proteins of unknown structure with proteins that have related spatial structures (Bowie et al., 1990, 1991; Burbaum et al., 1990; Jones et al., 1992; Luthy et al., 1992; Yue & Dill 1992; Bowie & Eisenberg, 1993; Bryant & Lawrence, 1993; Luthardt & Frommel, 1994), even when the identity levels appear random (5–10%). Recent progress has been made in using alignments of distantly related proteins to improve upon the accuracy of secondary structure predictions (Nishikawa & Ooi, 1986; Levin et al., 1993; Rost et al., 1993; Rost & Sander, 1993a, 1993b, 1994).

Using these methods, together with biochemical information, we have classified the sequenced PLP-utilizing enzymes into seven fold types. This analysis suggests that the eukaryotic ODCs and alanine racemases may belong to the same fold type and that they are not homologous to any of the three known structural classes of PLP-dependent enzymes. Furthermore, the N-terminal PLP-binding domain of ODC is likely to belong to the β/α-barrel fold. Given this sequence and structural alignment, striking similarities emerge between the ODCs and the FMN dehydrogenases, for which a structure is available. These similarities have allowed a model for the PLP-binding site to be constructed. Residues in the area of the active site, which might be important for the mechanism of action, were identified. We have carried out mutagenic experiments on *Trypanosoma brucei* ODC that demonstrate the importance of one of these residues, Glu 274, in the activity of this enzyme, providing support for the model (Osterman et al., 1995). Recent crystallographic analysis of ALR has confirmed that it is a β/α-barrel. (D. Ringe, pers. comm.).

Reprint requests to: Elizabeth J. Goldsmith, Department of Biochemistry, University of Texas Southwestern Medical Center, 5323 Harry Hines Boulevard, Dallas, Texas 75235; e-mail: betsy@howie.swmed.edu.

Abbreviations: ODC, eukaryotic ornithine decarboxylase; ALR, alanine racemase; AAT, aspartate aminotransferase; WSY, tryptophan synthase; GDC, glycine decarboxylase; SCS, selenocysteine synthase; ACCD, 1-aminocyclopropane-1-carboxylate deaminase; PDB, Protein Data Bank (Bernstein et al., 1977); PLP, pyridoxal phosphate.

Results and discussion

Sequence families

The PLP-dependent enzymes with reported sequences (312 in total) have for the most part previously been grouped into 16 Prosite (Bairoch, 1993) entries or protein families. We placed those not found in the Prosite entries into either one of the existing families, e.g., alanine aminotransferases, or into new families when no sequence similarity was apparent to any of the previously described families (Table 1). The sequences of all

Table 1. *PLP-dependent enzymes*[a]

	n	EC number	Prosite ID	Sequences
Fold type I (AAT)	198			
Aminotransferases class I	51			
Aspartate aminotransferase (AAT)	29	EC 2.6.1.1	PS00105	D00252, D14673, D25322, L09702, Z25466, X63428
Tyrosine aminotransferase	4	EC 2.6.1.5	PS00105	
Aromatic aminotransferase	1	EC 2.6.1.57	PS00105	
1-Aminocyclopropane-1-carboxylate synthase	13	EC 4.4.1.14	PS00105	U03294, X65982, L07882
Alanine aminotransferase	3	EC 2.6.1.2	NA	Alat_Human, Alat_Rat, Ala2_Panmi
MALY protein of *Escherichia coli*	1		NA	Maly_Ecoli
Aminotransferases class II	26			
Glycine acetyltransferase	1	EC 2.3.1.29	PS00599	
5-Aminolevulinic acid synthase (δ-ALA synthase)	11	EC 2.3.1.37	PS00599	
8-Amino-7-oxononanoate synthase	4	EC 2.3.1.47	PS00599	U00010
Histidinol-phosphate aminotransferase	8	EC 2.6.1.9	PS00599	
Serine palmitoyltransferase	1	EC 2.3.1.50	NA	Lcb1_Yeast
Pseudomonas denitrificans cobC protein	1		NA	CobC_Psede
Aminotransferases class III	30			
Acetylornithine aminotransferase	3	EC 2.6.1.11	PS00600	
Ornithine aminotransferase	6	EC 2.6.1.13	PS00600	L15426
ω-Amino acid-pyruvate aminotransferase	1	EC 2.6.1.18	PS00600	
4-Aminobutyrate aminotransferase (GABA transaminase)	6	EC 2.6.1.19	PS00600	U00011
DAPA aminotransferase	3	EC 2.6.1.62	PS00600	
2,2-Dialkylglycine decarboxylase	1	EC 4.1.1.64	PS00600	
Glutamate-1-semialdehyde aminotransferase (GSA)	8	EC 5.4.3.8	PS00600	
Lysine-ε aminotransferase	2	EC 2.6.1.36	NA	Lat_Nocla, Lat_Strcl
Aminotransferases class V	22			
Phosphoserine aminotransferase	5	EC 2.6.1.52	PS00595	
Serine-pyruvate aminotransferase	4	EC 2.6.1.51	PS00595	
Isopenicillin-*N*-epimerase (gene cefD)	2	EC 5.-.-.-	PS00595	
NifS—a protein of the nitrogen fixation operon of prokaryotes	9		PS00595	X68444, M98822
The small subunit of cyanobacterial soluble hydrogenase	2	EC 1.12.-.-	PS00595	
Others	69			
Serine hydroxymethyltransferase (SHMT)	15	EC 2.1.2.1	PS00096	L22529, L22528
Tryptophan indole-lyase (tryptophanase)	5	EC 4.1.99.1	PS00853	D14297
Tyrosine phenol-lyase (β-tyrosinase)	4	EC 4.1.99.2	PS00853	D13002
Prokaryotic ornithine decarboxylase (ODC)	2	EC 4.1.1.17	PS00703	
Prokaryotic lysine decarboxylase (LDC)	3	EC 4.1.1.18	PS00703	
Escherichia coli biodegradative arginine decarboxylase	1	EC 4.1.1.19	PS00703	
Glutamate decarboxylase (GAD)	8	EC 4.1.1.15	PS00392	M81883, Jh0423, M74826
Histidine decarboxylase (HDC)	6	EC 4.1.1.22	PS00392	X70644
Aromatic-L-amino-acid decarboxylase	11	EC 4.1.1.28	PS00392	M96070
Fruit fly α-methyldopa hypersensitive protein	1		NA	L2am_Drome
Cystathionine γ-lyase (γ-cystathionase)	3	EC 4.4.1.1	PS00868	
Cystathionine γ-synthase	1	EC 4.2.99.9	PS00868	
Cystathionine β-lyase (β-cystathionase)	2	EC 4.4.1.8	PS00868	
OAH/OAS sulfhydrylase	1	EC 4.2.99.-	PS00868	
Glycine decarboxylase	5	EC 1.4.4.2	NA	Gcsp_Chick, Gcsp_Human, Gcsp_Pea, Z25857, Gcsp_Ecoli
Selenocysteine synthase	1		NA	Sela_Ecoli
Fold type II (WSY)	61			
Tryptophan synthase β subunit (or domain)	24	EC 4.2.1.20	PS00168	M76685, L14596
Cysteine synthase (CSase)	8	EC 4.2.99.8	PS00901	D28777
Cystathionine β-synthase	3	EC 4.2.1.22	PS00901	
L-Serine dehydratase	3	EC 4.2.1.13	PS00165	
D-Serine dehydratase	2	EC 4.2.1.14	PS00165	
Threonine dehydratase	8	EC 4.2.1.16	PS00165	
Threonine synthase	8	EC 4.2.99.2	PS00165	D14071
1-Aminocyclopropane-1-carboxylate deaminase	2	EC 4.1.99.4	NA	1a1d_Psesp, A1ad_Pses0
Alliin lyase	3	EC 4.4.1.4	NA	Alln_Allas, Alln_Allce, Alln_Allsa

(*continued*)

Table 1. *Continued*

	n	EC number	Prosite ID	Sequences
Fold type III (ODC)	30			
Eukaryotic ornithine decarboxylase (ODC)	14	EC 4.1.1.17	PS00878	Drome (Rom & Kahana, 1993), U03059
Prokaryotic diaminopimelic acid decarboxylase (DAPDC)	6	EC 4.1.1.20	PS00878	L18879
Pseudomonas syringae pv. *tabaci* protein tabA	1		PS00878	
Bacterial and plant biosynthetic arginine decarboxylase	3	EC 4.1.1.19	PS00878	L16582
Alanine racemase (ALR)	6	EC 5.1.1.1	PS00395	
Fold type IV (ILV)	6			
Aminotransferases class IV	6			
Branched-chain amino-acid aminotransferase	2	EC 2.6.1.42	PS00770	
Protein ECA39	1		NA	Ec39_Mouse
D-Alanine aminotransferase	1	EC 2.6.1.21	PS00770	
4-Amino-4-deoxychorismate (ADC) lyase (gene pabC)	2	EC 4.-.-.-	PS00770	
Fold type V (PHS)	14			
Glycogen phosphorylase	14	EC 2.4.1.1	PS00102	
Fold ???	2			
Succinyldiaminopimelate aminotransferase	1	EC 2.6.1.17	NA	Dapd_Ecoli
Pseudomonas syringae hypothetical protein	1		NA	Yta3_Psesz
Fold ???	1			
Valine-pyruvate aminotransferase	1	EC 2.6.1.66	NA	Avta_Ecoli

[a] Enzyme families with a representative of known structure are shown in bold. Only the sequences (Swiss-Prot ID or GenBank accession number) absent in Prosite (release 12.1, 10/94; Bairoch, 1993) are shown. Others are listed in corresponding Prosite entries. *n*, Number of sequences considered in the analysis.

families were further classified into seven structural superfamilies as described below.

Structural superfamilies

Known three-dimensional structures of PLP-enzymes correspond to three different folds, originally described for: AAT, fold type I, WSY (PDB entry 1WSY), fold type II, and glycogen phosphorylase (PDB entry 1PYG), labeled here fold type V. Structures are available for nine PLP-dependent enzymes, seven of which belong to fold type I: PDB entries 1AAT, 7AAT, 1ASD, and ω-amino acid:pyruvate aminotransferase (Watanabe et al., 1989), tyrosine phenol-lyase (Antson et al., 1993), dialkylglycine decarboxylase (Toney et al., 1993, 1995), ornithine decarboxylase from *Lactobacillus*, solved recently (Momany et al., 1995). All of these enzymes are α/β proteins. Further, there are two regions in the PLP-binding sites that are present in all three different folds: a lysine residue, which forms a Schiff base with PLP, and a glycine-rich loop, which interacts with the phosphate of PLP. However, in fold types I and II, a distinct pattern of the placement of these two moieties in the sequence is observed (Fig. 1). In fold type I, the glycine-rich loop is found early in the sequence, whereas the Schiff base lysine is near the C-terminus. By contrast, in WSY the Schiff base lysine is near the N-terminus, whereas the phosphate-binding loop is closer to the C-terminus. The C-terminal domain of glycogen phosphorylase has a lactate dehydrogenase fold (Adams et al., 1970; Abad-Zapatero et al., 1987), and no other PLP-dependent proteins are similar.

Fold type I

Previous studies have classified several PLP-dependent enzymes as belonging to the same superfamily as AAT (Alexander et al., 1994; Sandmeier et al., 1994). This enzyme is an α/β protein, comprised of two domains. The first (PLP-binding) domain consists of a doubly wound planar β-sheet with unusual topology (McPhalen et al., 1992). We took as a basis for fold type I the α-family defined by Alexander et al. (1994). The α-family comprises aminotransferase-related enzymes and decarboxylase groups II and III defined by Sandmeier et al. (1994) (Table 1; fold type I enzymes from Prosite entry PS00868, GDCs, and SCS). We found that the γ-family (Prosite entry PS00868) defined by Alexander et al. (1994), the GDCs (group I, Sandmeier et al. 1994), and SCS also are likely to belong to fold type I. Profile analysis alone does not allow their unambiguous unification (Alexander et al., 1994; Sandmeier et al., 1994; highest *Z*-score is 4.1). However, two structural features have an identifiable signature within the sequences of fold type I enzymes. First, the Schiff base lysine (K258 in AAT) is closer to the C-terminus than the glycine-rich region and directly follows a hydrophobic β-strand (Fig. 1). Second, an invariant aspartic acid, which binds the pyridoxal ring nitrogen atom (D222 in AAT), precedes the Schiff base lysine by 20–50 amino acids.

Secondary structure prediction and profile analysis were used to make an alignment (see supplementary material in the Electronic Appendix) that includes the entries listed in Table 1. The γ-family and GDCs fit in the alignment as well as the other

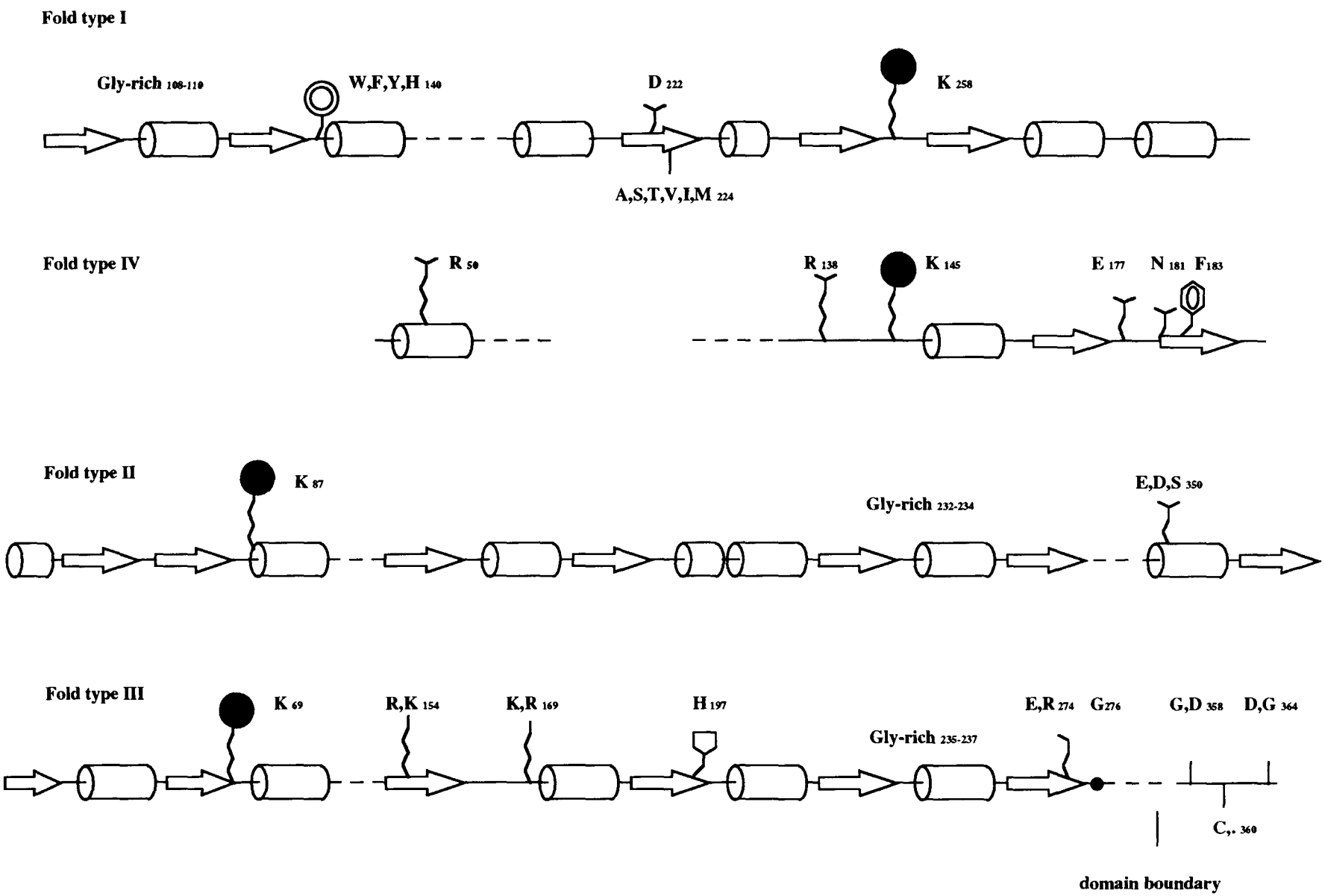

Fig. 1. Secondary structure elements and conserved residues in four of the PLP-binding fold types. Strands are shown as arrows, helices as cylinders, and loops as lines. Regions of sequence that are not shown are indicated by dashed lines. Conserved residue numbers correspond to those for AAT (Aatc_Chick) in fold type I, WSY (Trpb_Salty) for fold type II, ODC (Dcor_Mouse) for fold-type III, and D-alanine aminotransferase (Daaa_Ecoli) for fold type IV. Schiff base lysine is marked by a solid circle.

families. Because only a single sequence is available for SCS, its identification with this superfamily is less certain. From the alignment the phosphate-binding loop was identifiable for each family (the glycine-rich region, Fig. 1). Additionally, we find an aromatic residue (W140 of AAT), which is involved in binding the pyridoxal ring, is conserved throughout, whereas A224 of AAT, also involved in binding PLP, is constrained to be one of only six residues, A, S, T, V, I, or M.

In contrast to Mehta et al. (1993), we found that the aminotransferases class IV defined in Prosite (PS00770) did not belong to fold type I based on differences in the position of conserved residues and differences in predicted secondary structure. We have given these a separate label, fold type IV (Table 1); conserved residues and identifiable secondary structural elements are shown in Figure 1.

Fold type II

Fold type II is typified by WSY. The PLP-binding β subunit (or C-terminal domain) of this enzyme contains two domains comprised of mixed α/β structure with PLP bound between these two domains (Hyde et al., 1988). Several dehydratases and synthases had previously been aligned with WSY (β-family by Alexander et al., 1994). Our analysis extends this family to include ACCD and alliin lyase. The alignment of ACCD was straightforward on the basis of several boxes of conserved residues, as well as the placement of the Schiff base lysine and phosphate binding glycine-rich loop. The placement of alliin lyase is less certain, because the PLP-binding residue has not been identified and the sequence similarity with the other type II enzymes is low (see supplementary material in the Electronic Appendix). In this alignment, E350 of WSY, which is involved in binding PLP, is largely conserved as E, D, or S.

Fold type III: ODC and ALR

A large number of sequences of ODCs and ALRs have been reported (Table 2). Both of these families share some characteristics with the fold type II PLP enzymes. The Schiff base lysine precedes the glycine-rich loop, α-helices and β-strands are predicted to alternate throughout the sequence, and a large insertion can be found for some sequences near the C-terminus. However, the following unique characteristics led us to place the ODCs and ALRs into a probable separate fold, which we label fold type III (Table 1). The first notable difference is that in the ODCs and ALRs the Schiff-base lysine immediately follows a hydrophobic β-strand (see Fig. 1 and supplementary material in the Electronic Appendix) rather than a loop structure as in fold type II enzymes. Second, whereas fold type II has two β-strands in sequence flanked by α-helices near the N-terminus, secondary structure prediction and profile analysis suggest that the corresponding sequences of the ODCs and ALRs are formed into two β/α units.

Although the ODCs and ALRs have differences relative to fold II families, in each respect they appear similar to one another. For both families, the Schiff base lysine is located similarly in the overall sequence and relative to the predicted secondary structure, as noted above, directly following a hydrophobic β-strand. The pattern of predicted secondary structure elements is very similar as well as regular. Using the maximum-of-matching doublet strategy (see Materials and methods) and alignment of these described characteristics, we were able to construct a sequence alignment containing both enzyme families (see Fig. 2 and supplementary material in the Electronic Appendix). From this alignment, invariant histidine and glycine residues (His 197 and Gly 276 in mouse ODC [Dcor_Mouse]) were identified in addition to the conserved Schiff base lysine (Lys 69). Further, two conserved residues corresponding to Gly 358 and Asp 364 of mouse ODC are present (see Fig. 1 and supplementary alignment in the Electronic Appendix). Interestingly, these two residues can both be either glycine or aspartic acid, with a single G and D in each sequence. The importance of this fact is unknown.

Glycogen phosphorylase and other enzymes

The structure of glycogen phosphorylase is known (1PYG), and the PLP-binding C-terminal domain of the enzyme belongs to the lactate dehydrogenase fold superfamily, which we label fold V. None of the other PLP-utilizing enzymes analyzed in this study possess a lactate dehydrogenase fold.

Of the known PLP-binding proteins two enzyme classes, the succinyldiaminopimelate aminotransferase-related proteins and valine-pyruvate aminotransferase are without detectable sequence similarities to other proteins, and could not be fit into any of five described superfamilies (Table 1). These enzymes are likely to be the prototypes for yet two additional fold types of PLP binding enzymes.

Fold type III is a β/α-barrel

The proteins in the ODC/ALR superfamily appear to be organized into two domains. Two thirds of the sequence, or 300 amino acids belong to the N-terminal domain. The alignment of the ODC/ALR superfamily contains a region of low similarity with insertions of variable length, separating the N-terminal from the C-terminal part of the molecule. The insertion region is the site of proteolysis in ALR (Galakatos & Walsh, 1987) and thus is likely to represent a domain boundary.

The N-terminal domain has the following characteristics: (1) the secondary structure algorithms predict it is organized into at least seven β-strands alternating with seven α-helices; (2) the predicted β-strands are mainly hydrophobic, implying the strands form a parallel β-sheet; (3) several invariant residues are found near the C-terminus of the predicted β-strands, suggesting that the active site is at the C-terminal edge of the β-sheet. These points suggest that the N-terminal domain consists of a buried parallel β-sheet flanked by α-helices. Conserved residues are found throughout the sequence (Lys 69, Asp 134, Arg 154, His 197, Asp 233, Glu 274 of mouse ODC) at the C-terminus of β-strands or following loops (Fig. 2). Further, one such conserved residue is found in six of eight β/α units. This fact suggests that the C-termini of almost all the strands are close in space to form a compact active site.

By contrast, fold types I and II, which both contain planar β-sheets, have very different patterns of conservation (Fig. 1). In fold type I, conserved residues occur in two clusters, whereas in fold type II the conserved residues are distant in sequence from one another and do not map to the termini of the β-strands. Of the reported structures in the PDB database, only the β/α-barrel possesses all of the characteristics of secondary structure

Table 2. *Enzymes from ODC/ALR superfamily*

	Abbreviation	Organism	Comments
		Ornithine decarboxylases (ODCs) (EC 4.1.1.17, PS00878)	
1	Dcor_Mouse	*Mus musculus* (mouse)	
2	Dcor_Muspa	*Mus pahari* (shrew mouse)	
3	Dcor_Rat	*Rattus norvegicus* (rat)	
4	Dcor_Crigr	*Cricetulus griseus* (Chinese hamster)	
5	Dcor_Bovin	*Bos taurus* (bovine)	
6	Dcor_Human	*Homo sapiens* (human)	
7	Dcor_Chick	*Gallus gallus* (chicken)	
8	Dcor_Xenla	*Xenopus laevis* (African clawed frog)	
9	Dcor_Trybb	*Trypanosoma brucei*	
10	Drome	*Drosophila melanogaster* (fruit fly)	Absent in data banks
11	U03059	*Caenorhabditis elegans*	Absent in Swiss-Prot
12	Dcor_Leido	*Leishmania donovani*	
13	Dcor_Neucr	*Neurospora crassa*	
14	Dcor_Yeast	*Saccharomyces cerevisiae* (baker's yeast)	
		Diaminopimelate decarboxylases (DAPDCs) (EC 4.1.1.20, PS00878)	
15	Dcda_Myctu	*Mycobacterium tuberculosis*	
16	Dcda_Corgl	*Corynebacterium glutamicum*	
17	Dcda_Bacsu	*Bacillus subtilis*	
18	L18879	*Bacillus methanolicus*	Absent in Swiss-Prot
19	Dcda_Ecoli	*Escherichia coli*	
20	Dcda_Pseae	*Pseudomonas aeruginosa*	
		Taba protein (PS00878)	
21	Taba_Psesz	*Pseudomonas syringae* (pv. *tabaci*)	
		Arginine decarboxylases (ADCs) (EC 4.1.1.19, PS00878)	
22	Spea_Ecoli	*E. coli*	Biosynthetic ADC
23	Spe1_Avesa	*Avena sativa* (oat)	
24	L16582	*Lycopersicon esculentum* (tomato)	Absent in Swiss-Prot
		Alanine racemases (ALRs) (EC 5.1.1.1, PS00395)	
25	Alr_Bacst	*Bacillus stearothermophilus*	
26	Alr_Bacsu	*B. subtilis*	
27	Alr2_Ecoli	*E. coli*	Biosynthetic ALR
28	Alr2_Salty	*Salmonella typhimurium*	Biosynthetic ALR
29	Alr1_Ecoli	*E. coli*	Catabolic ALR
30	Alr1_Salty	*S. typhimurium*	Catabolic ALR

Fig. 2. (*See next three pages.*) Sequence alignment of ODC/ALR superfamily projected on the spatial structure alignment of β/α-barrel proteins. Data bank accession numbers (see Materials and methods) are used to identify the individual sequences, which are also listed in Tables 2 and 3. Long insertions are not shown and their lengths are indicated in numbers between tildes (~). The numbers of the first and the last residue in each sequence for each alignment block are shown before the block or after it. Total numbers of residues in the sequence are shown in parentheses at the end of each line. Mouse ODC (Dcor_Mouse) numbering is used to define residue position numbers for the family and is displayed in the ruler at the top of the figure. Secondary structure consensus for barrels of known structure is shown in the line SS structure: @, helical residues; =, sheet residues. Predictions of secondary structure elements are shown in line Dc PHD, if based on the ODC alignment and in line Alr PHD, if based on the ALR alignment: H, predicted helix; E, predicted strand; L, predicted loop. Secondary structure elements of the $(\beta/\alpha)_8$ motif are numbered above the alignment (strands: from 1S to 8S, helices: from 1H to 8H, inserted after 2S and 8S small helices are marked as 2H′ and 8H′). Doublets of amino acids that match between any of the proteins of unknown structure and any of the proteins with determined structure are underlined; asterisks mark the alignment positions with residues from at least one matching doublet. Side chains predicted to pack between a β-sheet and α-helices are marked in bold. Invariant residues in the ODC/ALR superfamily, Lys 69, which forms a Schiff base with PLP (▲), and His 197 (○) are marked. Suggested phosphate-binding site residues (●) (with invariant Gly 276 among them) and the negatively charged residues in decarboxylases, which the model places near the pyridine nitrogen of PLP (■), are also marked.

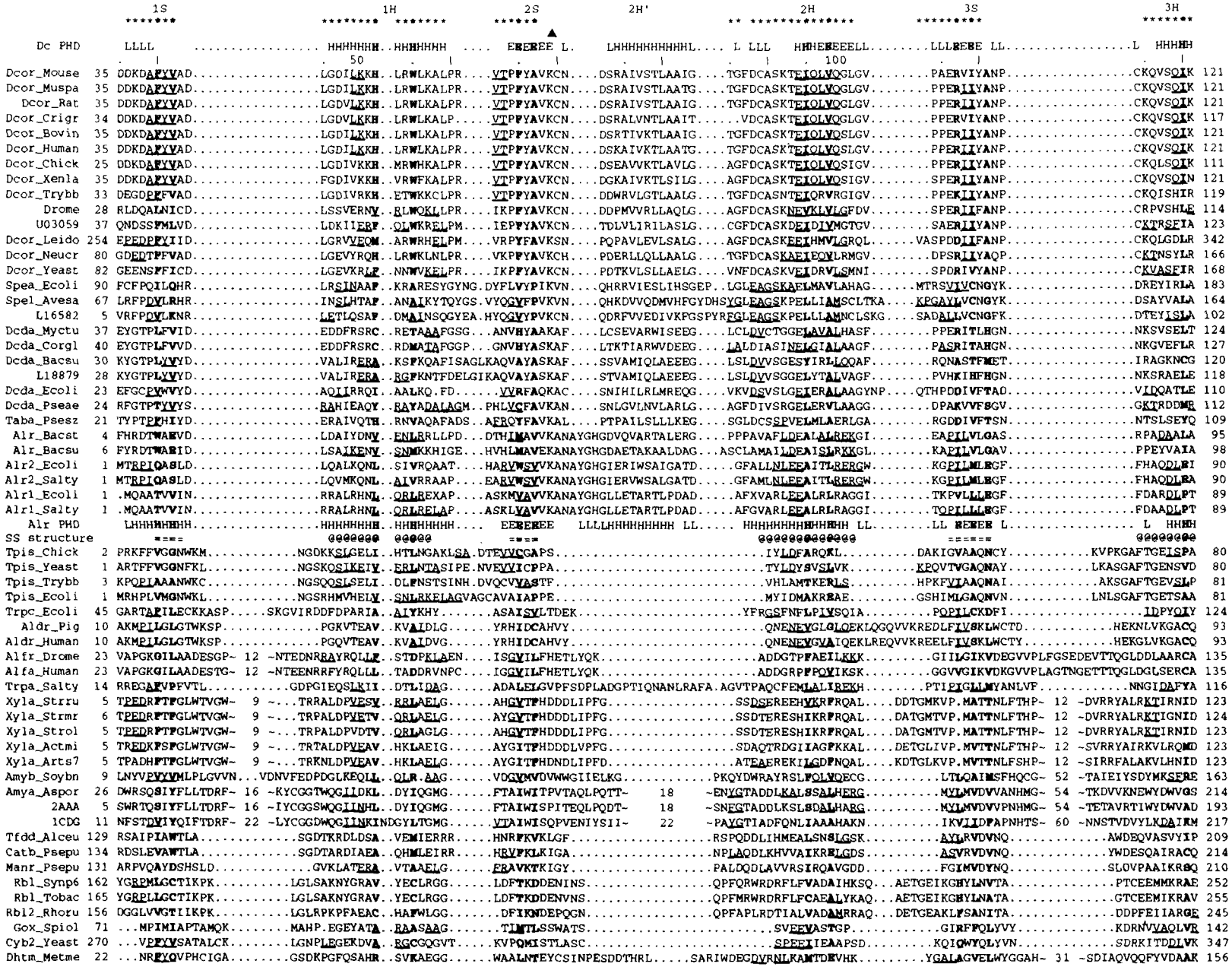

Fig. 2. *Caption appears on facing page, and figure continues on next two pages.*

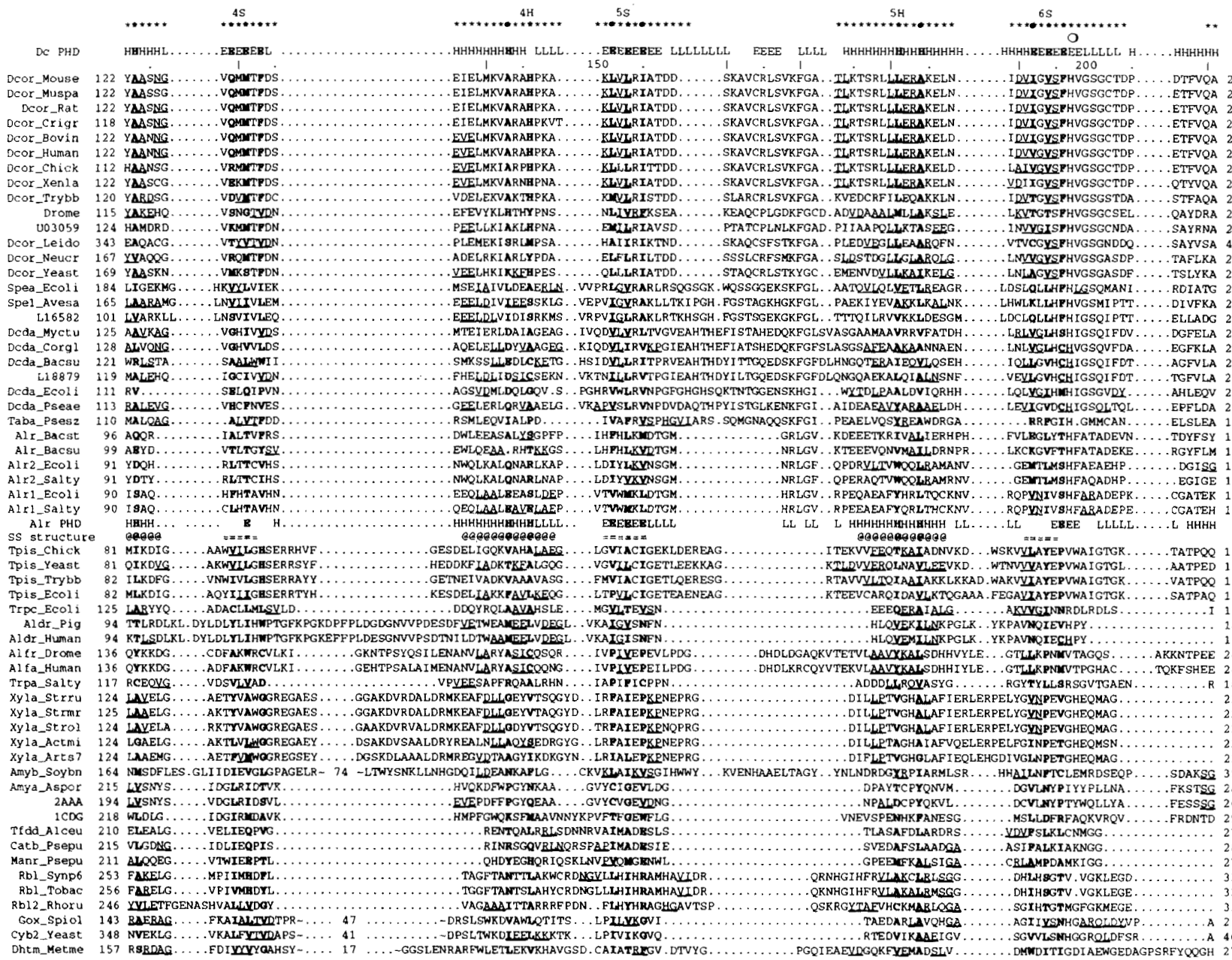

Fig. 2. *Continued.*

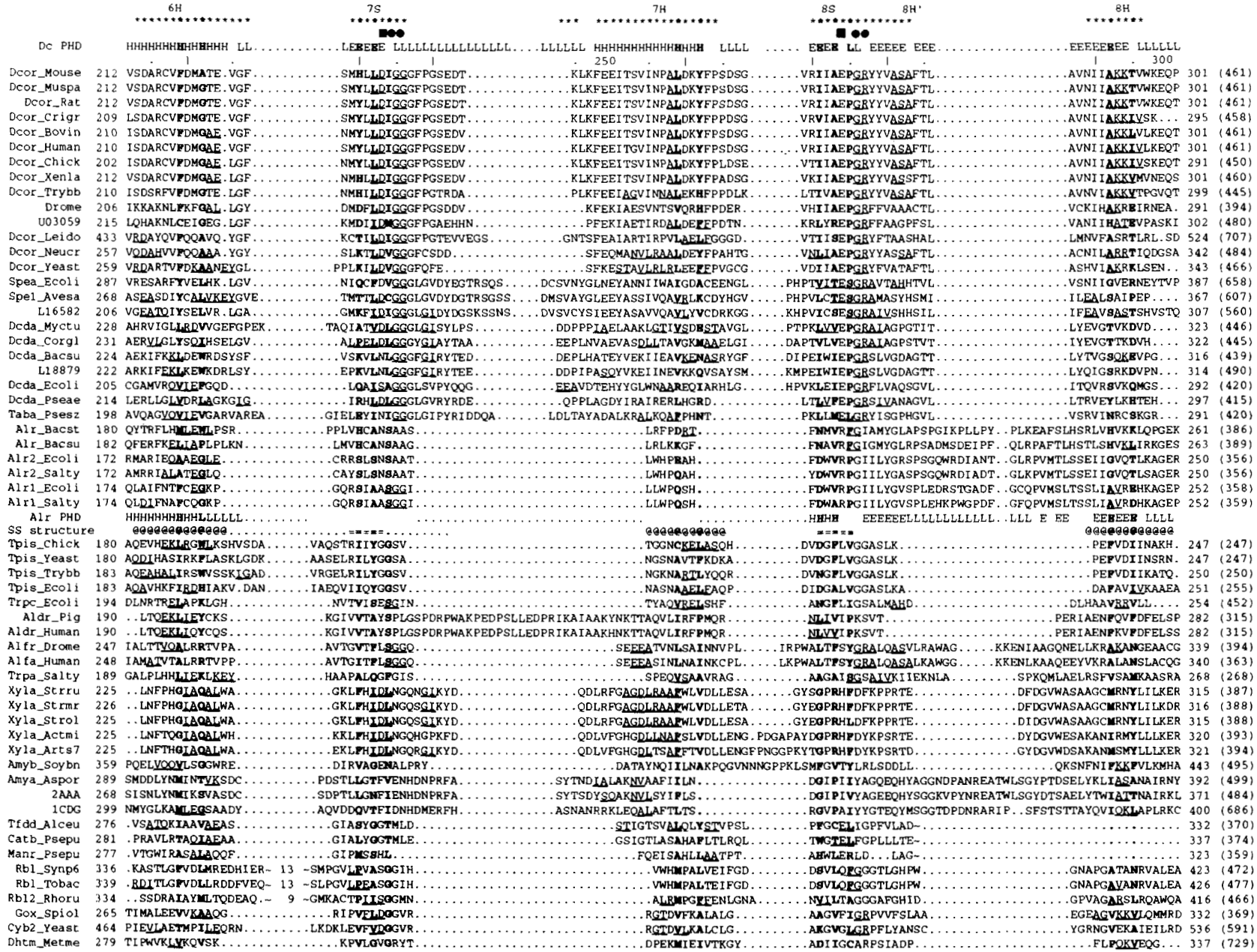

Fig. 2. *Continued.*

and sequence conservation listed above. The presence of these features in the ODC/ALR superfamily suggests that their N-terminal domain is a β/α-barrel.

The C-terminal domain is less similar among members of the ODC/ALR superfamily (see supplementary material in the Electronic Appendix), and secondary structure predictions of this part are less certain. Some similarities of this domain with cytochrome *b*5 can be detected by profile analysis (*Z*-score more than 5).

Structural alignment of β/α-barrels

About 20 families of enzymes share the β/α-barrel fold (Farber, 1993), which is comprised of an eight-stranded parallel circular β-sheet in which the first β-strand forms hydrogen bonds with the eighth β-strand. The β-sheet is surrounded by eight α-helices. β/α-Barrels bind a variety of substrates and catalyze a variety of reactions; many small deviations from the perfect eight β/α unit barrels have been catalogued (Farber & Petsko, 1990; Farber, 1993).

The structures of 28 β/α-barrel proteins (Table 3) were aligned based on the placement and orientation of residues within the β/α unit using local superpositions. Normally, 24 C_α atoms from each structure were compared, including C_α atoms from three β-strands and two α-helices. β-Strands were aligned to match side-chain orientation (toward the center of the barrel or toward the α-helices) and hydrogen bonding interactions. The α-helices were most variable, but we found that two hydrophobic residues, related by $i:i+3$, one in the middle of the helix and one near the C-terminus, point toward the β-sheet. The average RMS deviation over the 24 atoms compared was 1.6 Å. This procedure gave rise to the sequence alignment given in Figure 2. The alignment is more comprehensive than previous alignments (Pickett et al., 1992; Taylor et al., 1994), and is distinguished by being based on local rather than global superpositions.

Mapping of ODC/ALR onto structural alignment of β/α-barrels

The sequences for the ODC/ALR superfamily were mapped onto the structural alignment of the β/α-barrels, described above, maximizing the number of matching doublets (Fig. 2; see Materials and methods). If the hypothesis concerning the β/α-barrel structure of the ODC/ALR superfamily is correct, the alignment can be used to model elements of their secondary and spatial structure by extrapolation from known β/α-barrel proteins. Some contradictions between secondary structure predictions (Fig. 2, lines Dc PHD and Alr PHD) and consensus of secondary structure pattern for known barrel proteins (line SS structure) are present. Elements in ODC/ALRs, corresponding to the strand 1S and the helix 8H in the barrel proteins are predicted by the secondary structure prediction program PHD (see Materials and methods) to be a helix (1S) and a strand (8H). Taking into account that secondary structure predictions usually are no more than 75% correct (in this case 14 secondary structure elements out of 16: 88% were predicted correctly), those elements were modeled to match the β/α-barrel pattern instead.

The alignment predicts the orientation of residues within the β-sheet. In Figure 3, the sequence of mouse ODC is shown in the context of the structural alignment. The model predicts that all residues pointing toward helices are hydrophobic, whereas those facing into the center of the barrel can be hydrophobic or hydrophilic. Several residues that are predicted to pack between the β-sheet and helices are polar or charged for some other members of the ODC/ALR superfamily. However, this composition is consistent with what has been observed for barrel families of known structure. The hydrophilic residues Glu 274, Arg 154 and Asp 233, Thr 132, and Ser 195, as well as the Schiff base lysine and glycine-rich sequence (Gly 235–Gly 237) map toward the C-terminus of the β-sheet.

Due to diversity of helix packing in known β/α-barrels, detailed modeling of helix packing around the circular sheet is not possible. However, every helix (except 8H′, which is expected to be on the domain interface) has a definite amphipathic character, thus the β/α-barrel structure of ODC/ALRs obeys the general rules of protein structure with a hydrophobic interior and a mostly hydrophilic surface.

Loops and junctions between strands and helices in the ODC/ALR superfamily are mostly hydrophilic and their lengths are compatible with the loop lengths in other β/α-barrel proteins, although the connections between 1S and 1H, 3S and 3H, and between 3H and 4S appear to be unusually short. This may be to accommodate packing of another domain and/or subunit in the dimer. Short loops could cause certain deviations from the ideal barrel shape for the ODC/ALR superfamily. One can note that helix 7H is absent in ALRs according to our alignment. Its replacement by a loop could also perturb the barrel shape.

PLP-binding site of ODC

With the alignment of ODC/ALR to other barrel proteins, we found that similarities emerged between the ODC and FMN dehydrogenases (Fig. 2). Conserved residues in the ODC map to positions of conserved residues in the FMN family: these include a positively charged residue in position 154 (Dcor_Mouse numbers, Fig. 2), Thr 132, Asp 134, Ser 195, His 197, Asp 233, Gly 235, Arg 277, and Gly 276. These positions encompass a total of 9 out of 11 active site residues identified in glycolate oxidase (Lindqvist & Branden, 1989). FMN and PLP both contain phosphate and planar heterocycles, which may account for these surprising correspondences.

Based on these similarities with the FMN-binding barrel, it was straightforward to build a model for the PLP-binding site of ODC (mouse sequence, Fig. 4). Lys 69 that forms the Schiff base with PLP (Poulin et al., 1992) is positioned at the end of strand 2S. In ODC, the Gly 235–Gly 237 follows 7S. Interestingly, in 13 barrel enzymes that bind phosphate-containing substrates or cofactors, the phosphate group binds near the C-termini of strands 7S and 8S, interacting with a glycine-rich loop following 7S (Bernstein et al., 1977; Wilmanns et al., 1991). Further, in glycolate oxidase Arg 309 is involved in phosphate binding, in ODC Arg 277 is found in this position. These facts allow the approximate positioning of PLP in the active site and suggest that the function of the β/α-barrel domain is to bind PLP rather than substrate.

Given the structural alignment, some of the similarities between glycolate oxidase (Lindqvist & Branden, 1989) and ODC/ALR are interesting. First, His 254 in glycolate oxidase, which is thought to act as a catalytic base, is an invariant residue in the ODC/ALR superfamily (His 197 of Dcor_Mouse). Second, in glycolate oxidase a conserved Asp–Ser ion-dipole is in-

Table 3. *β/α-Barrel proteins with known spatial structure*

	Abbreviation	Protein	EC number	Organism	Prosite accession	Brookhaven ID	Resolution, Å
1	Tpis_Chick	Triosephosphate isomerase (TIM)	EC 5.3.1.1	*Gallus gallus* (chicken)	PS00171	1TPH	1.8
2	Tpis_Yeast	Triosephosphate isomerase (TIM)	EC 5.3.1.1	*Saccharomyces cerevisiae* (baker's yeast)	PS00171	1YPI	1.9
3	Tpis_Trybb	Triosephosphate isomerase (TIM)	EC 5.3.1.1	*Trypanosoma brucei*	PS00171	5TIM	1.83
4	Tpis_Ecoli	Triosephosphate isomerase (TIM)	EC 5.3.1.1	*Escherichia coli*	PS00171	1TRE	2.6
5	Trpc_Ecoli	Indole-3-glycerol phosphate synthase (IGPS)	EC 4.1.1.48	*Escherichia coli*	PS00614	1PII	2
6	Aldr_Pig	Aldose reductase (AR)	EC 1.1.1.21	*Sus scrofa* (pig)	PS00798	1DLA	3
7	Aldr_Human	Aldose reductase (AR)	EC 1.1.1.21	*Homo sapiens* (human)	PS00798	1ADS	1.6
8	Alfr_Drome	Fructose-bisphosphate aldolase	EC 4.1.2.13	*Drosophila melanogaster* (fruit fly)	PS00158	1FBA	1.9
9	Alfa_Human	Fructose-bisphosphate aldolase A	EC 4.1.2.13	*Homo sapiens* (human)	PS00158	2ALD	2
10	Trpa_Salty	Tryptophan synthase α chain	EC 4.2.1.20	*Salmonella typhimurium*	PS00167	1WSY	2.5
11	Xyla_Strru	Xylose isomerase	EC 5.3.1.5	*Streptomyces rubiginosus*	PS00172	1XIS	1.6
12	Xyla_Strmr	Xylose isomerase	EC 5.3.1.5	*Streptomyces murinus*	PS00172	1DXI	2.6
13	Xyla_Strol	Xylose isomerase	EC 5.3.1.5	*Streptomyces olivochromogenes*	PS00172	1XYB	1.96
14	Xyla_Actmi	Xylose isomerase	EC 5.3.1.5	*Actinoplanes missouriensis*	PS00172	1XIM	2.2
15	Xyla_Arts7	Xylose isomerase	EC 5.3.1.5	*Arthrobacter* sp. (strain NRRL B3728)	PS00172	4XIA	2.3
16	Amyb_Soybn	β-Amylase	EC 3.2.1.2	*Glycine max* (soybean)	PS00506	1BTC	2
17	Amya_Aspor	α-Amylase	EC 3.2.1.1	*Aspergillus oryzae*	NA	6TAA	2.1
18	2AAA	α-Amylase	EC 3.2.1.1	*Aspergillus niger*	NA	2AAA	2.1
19	1CDG	Cyclomaltodextrin glucanotransferase	EC 2.4.1.19	*Bacillus circulans* (strain 251)	NA	1CDG	2
20	Tfdd_Alceu	Chloromuconate cycloisomerase	EC 5.5.1.7	*Alcaligenes eutrophus*	PS00908	1CHR	3
21	Catb_Psepu	Muconate cycloisomerase I	EC 5.5.1.1	*Pseudomonas putida*	PS00908	1MLE	2.5
22	Manr_Psepu	Mandelate racemase (MR)	EC 5.1.2.2	*Pseudomonas putida*	PS00908	2MNR	1.9
23	Rbl_Synp6	Ribulose bisphosphate carboxylase (rubisco)	EC 4.1.1.39	*Synechococcus* sp. (strain PCC 6301)	PS00157	1RBL	2.2
24	Rbl_Tobac	Ribulose bisphosphate carboxylase (rubisco)	EC 4.1.1.39	*Nicotiana tabacum* (common tobacco)	PS00157	1RLD	2.5
25	Rbl2_Rhoru	Ribulose bisphosphate carboxylase (rubisco)	EC 4.1.1.39	*Rhodospirillum rubrum*	PS00157	5RUB	1.7
26	Gox_Spiol	Glycolate oxidase (GOX)	EC 1.1.3.15	*Spinacia oleracea* (spinach)	PS00557	1GOX	2
27	Cyb2_Yeast	Flavocytochrome *b2*	EC 1.1.2.3	*Saccharomyces cerevisiae* (baker's yeast)	PS00557	1FCB	2.4
28	Dhtm_Metme	Trimethylamine dehydrogenase (TMADH)	EC 1.5.99.7	Unclassified methylotrophic bacterium W3A1	NA	2TMD	2.4

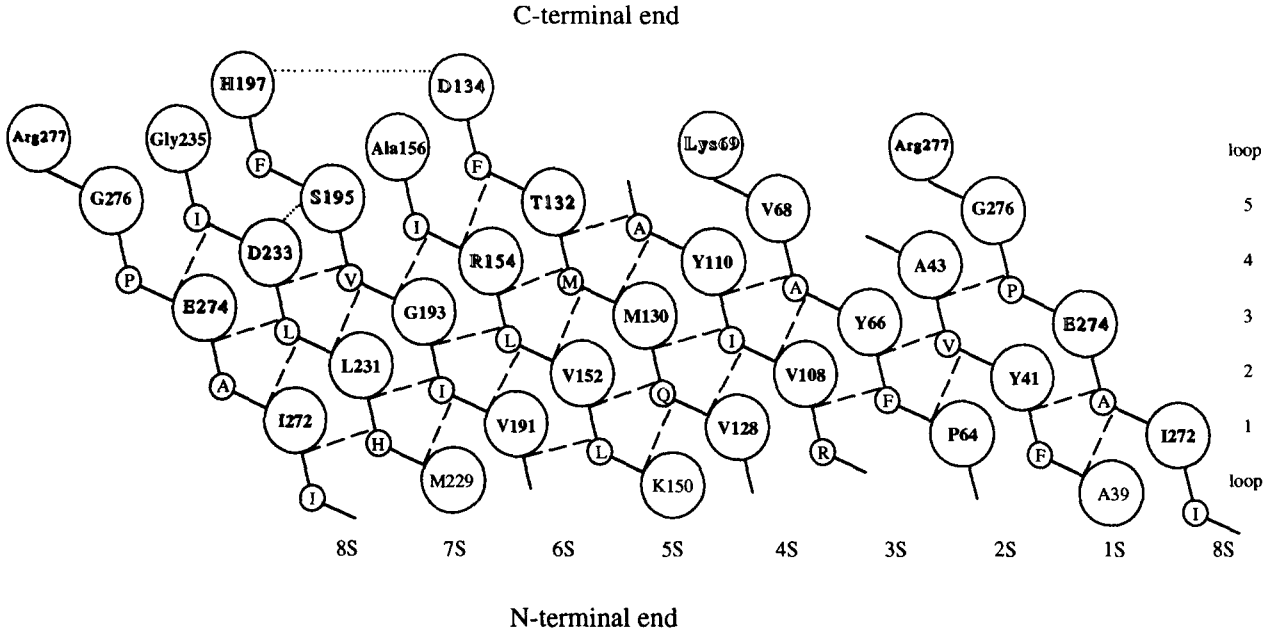

Fig. 3. A model of the putative hydrogen bonding pattern and side-chain orientations in the parallel β-sheet of mouse ODC (Dcor_Mouse). The eight-stranded sheet of the barrel is represented by a nine-stranded sheet in which the first and the last strands are identical. Strands are numbered according to their order in the sequence (1S–8S; see Fig. 2). Solid lines represent covalent bonds between neighboring residues, dashed lines represent main-chain hydrogen bonds, and dotted lines show putative side-chain hydrogen bonds. Amino acids with side chains pointed toward α-helices (between β-sheet and α-helices) are shown in small circles and amino acids with side chains pointed toward the interior of the barrel (toward the reader on the picture) are shown in large circles. Residues of the sheet, pointed toward the center, are arranged in five rows (row numbers at the right); the hydrophilic residues among them are shadowed. Residues in N-terminal loops (preceding the β-strands) are shown in plain text.

volved in cofactor binding. This Ser–Asp pair is conserved in the ODCs (Ser 195, Asp 233) (Figs. 2, 4). Third, position 154 of ODC aligns with Lys 230 in glycolate oxidase, which is involved in binding the FMN cofactor. This position is always occupied by arginine or lysine in the ODC/ALR superfamily, conserving the positive charge present in glycolate oxidase. Fourth, in ODC the model positions Glu 274 and Asp 233 near the pyridine nitrogen atom of PLP. Glu 274 has no negatively charged analog in glycolate oxidase.

Amino acid substrate-binding site

The model provides less information about the amino acid substrate-binding site. Interaction of the substrate carboxylate with the N-terminal domain is supported by analysis of the substrate-binding profiles of cross-species heterodimers formed between mouse and *T. brucei* ODC (Osterman et al., 1994). In glycolate oxidase and flavocytochrome *b*2, the carboxylate of substrate is bound to the β/α-barrel to Arg 257, in the loop following S6 (Mathews & Xia, 1987; Lindqvist & Branden, 1989). Thus, this region may be the carboxylate-binding site in the ODCs. However, the type III decarboxylases possess shared active sites with residues from both subunits of a dimer contributing to the active site (Tobias & Kahana, 1993). Thus, the substrate-binding site is likely to be at the subunit interface, with residues from both domains contributing to it. This is supported by the finding that mutation of Asp 361 to Ala increased K_m by 2,000-fold without affecting k_{cat}, demonstrating that Asp 361 is a substrate-binding determinant (Osterman et al., 1995).

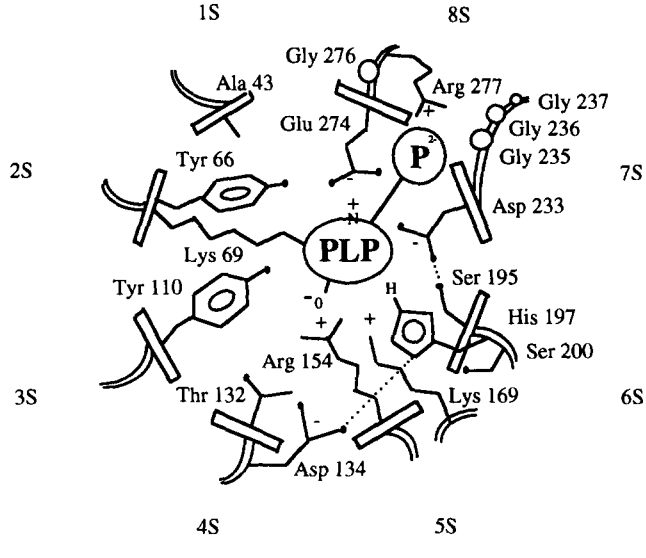

Fig. 4. A model of the PLP-binding site of mouse ODC (Dcor_Mouse). The view is parallel to the barrel central axis with the C-terminus of the β-sheet above. C-terminal ends of the β-strands (rectangles) are pointed toward the reader and numbered as in Figure 2 from 1S to 8S. N-termini of the loops connecting β-strands and α-helices are also displayed. PLP is shown as a Schiff base with the Lys 69. Phosphate of PLP is predicted to interact with the glycine-rich loop Gly 235–Gly 237 at the end of the seventh strand and with Gly 276, Arg 277, and the N-terminus of the inserted helix in the loop after the eighth strand. Tyr 66 and Tyr 110 are likely to tilt PLP toward strands 5–8. Functions of various charged residues surrounding PLP (Arg 154, Asp 233, Glu 274) are discussed in the text.

We are presently using the model for the β/α-barrel domain as a guide to site-specific mutagenesis to analyze the functional roles of residues in the active site (Osterman et al., 1995). Mutation of Glu 274 to Ala caused substantial reductions in the catalytic efficiency of the enzyme. However, analogously to the result obtained for Asp 222 to Ala mutant aspartate aminotransferase (Onuffer & Kirsch, 1994), activity could be restored to near wild-type levels for the Glu 274 to Ala mutant enzyme by replacing the PLP cofactor with *N*-methyl-pyridoxal phosphate. These results demonstrate that Glu 274 forms an interaction with the protonated pyridine nitrogen of PLP and support the proposed model for the structure of the N-terminal domain of ODC.

In conclusion, we have shown that extensive sequence information can in some cases be used to determine the fold of a protein, and where structural correspondences exist, catalytic and substrate-binding residues can be identified. We have predicted from this analysis that the N-terminal domain of ODC folds into a β/α-barrel, which contains the major PLP-binding determinants.

Materials and methods

Sequences were gathered from the databases Swiss-Prot (release 28, 09/94), GenBank (release 82, 05/94), Prosite (release 12.1, 10/94; Bairoch, 1993), and Brookhaven (PDB) (version 2.0, 06/94; Bernstein et al., 1977), and from the primary literature. The Brookhaven, Swiss-Prot, GenBank, and Prosite IDs or accession numbers are used as references for the sequences and structures; a complete bibliography can be found in these database entries.

Sequence alignments were carried out using the GCG (Genetics Computer Group, 1991) programs. Homologous sequences were found with StringSearch and Fasta, multiple alignments were constructed using Pileup, and profile analysis (Gribskov et al., 1987) was carried out using ProfileMake, ProfileSearch, and ProfileGap. Secondary structure predictions were calculated with the programs PHD (Rost et al., 1992; Rost & Sander, 1993a, 1994), Simulant (N.V. Grishin, unpubl.), and ALB (Ptitsyn & Finkelstein, 1983). The Simulant program uses empirically derived interaction energies considered to predict local conformation and is available by request. InsightII (Biosym) was used for spatial structure analysis.

Multiple alignments, produced by Fasta (default parameters) were adjusted by eye to maximize agreement of primary sequence, the predicted secondary structure, and hydrophobicity profiles. To produce alignment of two multiple alignments (i.e., the ODC and ALR families) with low identity between any pair of sequences, a program was written that searches for the maximum number of matching amino acid doublets as follows. The corresponding regions of secondary structure in two alignments were identified by hydrophobicity profile and secondary structure predictions. These regions, all short stretches of 10–15 amino acids, were aligned to ensure the maximal number of identities of amino acid doublets between the two alignments. Each doublet was counted only once.

Doublet-based strategies are widely used in fast homology search programs (e.g., Fasta in the GCG package). Doublets are much less likely to occur at random than single matches, and doublet matches may reflect features of the structure such as helix polarity and periodicity. We used a simple unitary matrix for doublets in our method: 1 for a perfect match and 0 otherwise.

Supplementary material in the Electronic Appendix

Subdirectory Grishin.SUP of the SUPLEMNT directory in the Electronic Appendix contains a file of alignments for enzymes of fold types I–IV.

Acknowledgments

We thank Dr. A.L. Osterman for helpful discussions. This work was supported by grants to M.A.P. from the Welch foundation (I-1257) and the National Institute of Health (R01 AI34432-01A2), and to E.J.G. by the Welch Foundation (I-1128).

References

Abad-Zapatero C, Griffith JP, Sussman JL, Rossmann MG. 1987. Refined crystal structure of dogfish apo-lactate dehydrogenase. *J Mol Biol 198*: 445–467.

Adams MJ, Ford GC, Koekoek R, Lentz PJ, McPherson A, Rossmann MG, Smiley IE, Schevitz RW, Wonacott AJ. 1970. Structure of lactate dehydrogenase at 2.8 Å resolution. *Nature 227*:1098–1103.

Alexander FW, Sandmeier E, Mehta PK, Christen P. 1994. Evolutionary relationships among pyridoxal-5′-phosphate-dependent enzymes. *Eur J Biochem 219*:953–960.

Antson AA, Demidkina TV, Gollnick P, Dauter Z, Von Tersch RL, Long J, Berezhnoy SN, Phillips RS, Harutyunyan EH, Wilson KS. 1993. Three-dimensional structure of tyrosine phenol-lyase. *Biochemistry 32*:4195–4206.

Bairoch A. 1993. The PROSITE dictionary of sites and patterns in proteins, its current status. *Nucleic Acids Res 21*:3097–3103.

Bernstein FC, Koetzle TF, Williams GJB, Meyer EF Jr, Brice MD, Rogers JR, Kennard O, Shimanouchi T, Tasumi M. 1977. The Protein Data Bank: A computer-based archival file for macromolecular structures. *J Mol Biol 112*:535–542.

Bowie J, Eisenberg D. 1993. Inverted protein structure prediction. *Curr Opin Struct Biol 3*:437–444.

Bowie JU, Clarke ND, Pabo CO, Sauer RT. 1990. Identification of protein folds: Matching hydrophobicity patterns of sequence sets with solvent accessibility patterns of known structures. *Proteins Struct Funct Genet 7*:257–264.

Bowie JU, Luthy R, Eisenberg D. 1991. A method to identify protein sequences that fold into a known three-dimensional structure. *Science 253*:164–170.

Bryant SH, Lawrence CE. 1993. An empirical energy function for threading protein sequence through the folding motif. *Proteins Struct Funct Genet 16*:92–112.

Burbaum JJ, Starzyk RM, Schimmel P. 1990. Understanding structural relationships in proteins of unsolved three-dimensional structure. *Proteins Struct Funct Genet 7*:99–111.

Farber GK. 1993. An α/β-barrel full of evolutionary trouble. *Curr Opin Struct Biol 3*:409–412.

Farber GK, Petsko GA. 1990. The evolution of α/β-barrel enzymes. *Trends Biochem Sci 15*:228–234.

Galakatos NG, Walsh CT. 1987. Specific proteolysis of native alanine racemase from *Salmonella typhimurium*: Identification of the cleavage site and characterization of the clipped two-domain proteins. *Biochemistry 26*:8475–8480.

Genetics Computer Group. 1991. *Program manual for the GCG package, version 7, April 1991.* Madison, Wisconsin: Genetics Computer Group.

Gribskov M, McLachlan AD, Eisenberg D. 1987. Profile analysis: Detection of distantly related proteins. *Proc Natl Acad Sci USA 84*:4355–4358.

Hyde CC, Ahmed SA, Padlan EA, Miles EW, Davies DR. 1988. Three-dimensional structure of the tryptophan synthase $\alpha_2\beta_2$ multienzyme complex from *Salmonella typhimurium*. *J Biol Chem 263*:17857–17871.

Jones DT, Taylor WR, Thornton JM. 1992. A new approach to protein fold recognition. *Nature 358*:86–89.

Levin JM, Pascarella S, Argos P, Garnier J. 1993. Quantification of secondary structure prediction improvement using multiple alignments. *Protein Eng 6*:849–854.

Lindqvist Y, Branden CI. 1989. The active site of spinach glycolate oxidase. *J Biol Chem 264*:3624–3628.

Luthardt G, Frommel C. 1994. Local polarity analysis: A sensitive method that discriminates between native proteins and incorrectly folded models. *Protein Eng 7*:627–631.

Luthy R, Bowie JU, Eisenberg D. 1992. Assessment of protein models with three-dimensional profiles. *Nature 356*:83–85.

Mathews FS, Xia ZX. 1987. *Flavins and flavoproteins*. Berlin: Walter de Gruyter. pp 124–132.

McCann PP, Pegg AE. 1992. Ornithine decarboxylase as an enzyme target for therapy. *Pharmacol & Ther 54*:195–215.

McPhalen CA, Vincent MG, Jansonius JN. 1992. X-ray structure refinement and comparison of three forms of aspartate aminotransferase. *J Mol Biol 225*:495–517.

Mehta PK, Hale TI, Christen P. 1993. Aminotransferases: Demonstration of homology and division into evolutionary subgroups. *Eur J Biochem 214*:549–561.

Momany C, Ghosh R, Hackert ML. 1995. Structural motifs for pyridoxal-5′-phosphate binding in decarboxylases: An analysis based on the crystal structure of the *Lactobacillus* 30a ornithine decarboxylase. *Protein Sci 4*:849–854.

Nishikawa K, Ooi T. 1986. Amino acid sequence homology applied to the prediction of protein secondary structures, and joint prediction with existing methods. *Biochim Biophys Acta 871*:45–54.

Onuffer JJ, Kirsch JF. 1994. Characterization of the apparent negative cooperativity induced in *Escherichia coli* aspartate aminotransferase by the replacement of Asp 222 with alanine. Evidence for an extremely slow conformational change. *Protein Eng 7*:413–424.

Osterman A, Grishin NV, Kinch LN, Phillips MA. 1994. Formation of functional cross-species heterodimers of ornithine decarboxylase. *Biochemistry 33*:13662–13667.

Osterman A, Kinch LN, Grishin NV, Phillips MA. 1995. Acidic residues important for substrate binding and cofactor reactivity in eukaryotic ornithine decarboxylase identified by alanine scanning mutagenesis. *J Biol Chem 270*:11797–11802.

Pickett SD, Saqi AS, Sternberg MJE. 1992. Evaluation of the sequence template method for protein structure prediction. *J Mol Biol 228*:170–187.

Poulin R, Lu LI, Ackermann B, Bey P, Pegg A. 1992. Mechanism of the irreversible inactivation of mouse ornithine decarboxylase by α-difluoromethylornithine. *J Biol Chem 267*:150–158.

Ptitsyn OB, Finkelstein AV. 1983. Theory of protein secondary structure and algorithm of its prediction. *Biopolymers 22*:15–25.

Rom E, Kahana C. 1993. Isolation and characterization of the *Drosophila* ornithine decarboxylase locus: Evidence for the presence of two transcribed ODC genes in the *Drosophila* genome. *DNA Cell Biol 12*:499–508.

Rost B, Sander C. 1993a. Prediction of protein secondary structure at better than 70% accuracy. *J Mol Biol 232*:584–599.

Rost B, Sander C. 1993b. Improved prediction of protein secondary structure by use of sequence profiles and neural networks. *Proc Natl Acad Sci USA 90*:7558–7562.

Rost B, Sander C. 1994. Combining evolutionary information and neural networks to predict protein secondary structure. *Proteins Struct Funct Genet 19*:55–72.

Rost B, Sander C, Schneider R. 1992. PHD—An automatic mail server for protein secondary structure prediction. *CABIOS 10*:53–60.

Rost B, Schneider R, Sander C. 1993. Progress in protein structure prediction? *Trends Biochem Sci 18*:120–123.

Sandmeier E, Hale TI, Christen P. 1994. Multiple evolutionary origin of pyridoxal-5′-dependent amino acid decarboxylases. *Eur J Biochem 221*: 997–1002.

Taylor WR, Flores TP, Orengo CA. 1994. Multiple protein structure alignment. *Protein Sci 3*:1858–1870.

Tobias KE, Kahana C. 1993. Intersubunit location of the active site of mammalian ornithine decarboxylase as determined by hybridization of site-directed mutants. *Biochemistry 32*:5842–5847.

Toney MD, Hohenester E, Cowan SW, Jansonius JN. 1993. Dialkylglycine decarboxylase structure: Bifunctional active site and alkali metal sites. *Science 261*:756–759.

Toney MD, Hohenester E, Keller JW, Jansonius JN. 1995. Structural and mechanistic analysis of two refined crystal structures of the pyridoxal phosphate-dependent enzyme dialkylglycine decarboxylase. *J Mol Biol 245*:151–179.

Watanabe N, Sakabe K, Sakabe N, Higashi T, Sasaki K, Aibara S, Morita Y, Yonaha K, Toyama S. 1989. Crystal structure analysis of ω-amino acid:pyruvate aminotransferase with a newly developed Weissenberg camera and an imaging plate using synchrotron radiation. *J Biochem 105*:1–3.

Wilmanns M, Hyde CC, Davies DR, Kirschner K, Jansonius JN. 1991. Structural conservation in parallel β/α-barrel enzymes that catalyze three sequential reactions in the pathway of tryptophan biosynthesis. *Biochemistry 30*:9161–9169.

Yue K, Dill K. 1992. Inverse protein folding problem: Designing polymer sequences. *Proc Natl Acad Sci USA 89*:4163–4167.