The Crystal Structure of Benzoylformate Decarboxylase at 1.6 Å Resolution: Diversity of Catalytic Residues in Thiamin Diphosphate-Dependent Enzymes[†,‡]

Miriam S. Hasson,*[,§,||] Angelika Muscate,[⊥,#] Michael J. McLeish,[⊥,▽] Lena S. Polovnikova,[§] John A. Gerlt,[○] George L. Kenyon,[⊥] Gregory A. Petsko,[||] and Dagmar Ringe[||]

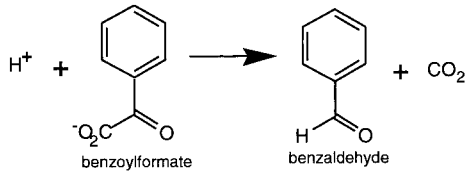
Department of Biological Sciences, Purdue University, West Lafayette, Indiana 47907-1392, Departments of Biochemistry and Chemistry, and the Rosenstiel Basic Medical Sciences Research Center, Brandeis University, Waltham, Massachusetts 02254-9110, Department of Pharmaceutical Chemistry, University of California, San Francisco, California 94143-0446, and Department of Biochemistry, University of Illinois, Urbana, Illinois 61801

Received December 11, 1997; Revised Manuscript Received May 7, 1998

ABSTRACT: The crystal structure of the thiamin diphosphate (ThDP)-dependent enzyme benzoylformate decarboxylase (BFD), the third enzyme in the mandelate pathway of *Pseudomonas putida*, has been solved by multiple isomorphous replacement at 1.6 Å resolution and refined to an *R*-factor of 15.0% (free $R = 18.6\%$). The structure of BFD has been compared to that of other ThDP-dependent enzymes, including pyruvate decarboxylase. The overall architecture of BFD resembles that of the other family members, and cofactor- and metal-binding residues are well conserved. Surprisingly, there is no conservation of active-site residues not directly bound to the cofactor. The position of functional groups in the active site may be conserved, however. Three classes of metal ions have been identified in the BFD crystal structure: Ca^{2+} bound to the cofactor in each subunit, Mg^{2+} on a 2-fold axis of the tetramer, and Ca^{2+} at a crystal contact. The structure includes a non-proline cis-peptide bond and an unusually long and regular polyproline type II helix that mediates the main contact between tetramers in the crystal. The high-quality electron-density map allowed the correction of errors totaling more than 10% of the amino acid sequence, which had been predicted from the reported sequence of the *mdlC* gene. Analysis of the BFD structure suggests that requirements for activation of the cofactor, the nature of the reaction intermediates, and architectural considerations relating to the protein fold have been dominant forces in the evolution of ThDP-dependent enzymes.

Benzoylformate decarboxylase (BFD;[1] EC 4.1.1.7) from *Pseudomonas putida* (ATCC 12633) is a thiamin-diphosphate (ThDP)-dependent enzyme that catalyzes the nonoxidative conversion of benzoylformate to benzaldehyde and carbon dioxide (Scheme 1). The enzyme is a component of the mandelate pathway, which allows some pseudomonads to utilize (*R*)-mandelic acid as a sole carbon source by converting it to benzoic acid, which is then metabolized by the

Scheme 1



β-ketoadipate pathway and the citric acid cycle. The mandelate pathway was the first pathway whose genes were recognized as being coordinately regulated (*1*). Its component enzymes also evidently physically associate in vivo to form a multienzyme complex (*2*). The mandelate pathway has been studied as a paradigm for the study of enzymic evolution (*3*). It is present in only a few organisms, suggesting that the components of the pathway have only recently evolved to assume their present roles. The comparison of a "new" enzyme, one that has recently diverged from another with a different catalytic ability, to "older" enzymes that catalyze similar reactions may reveal how proteins with related sequences evolve drastically different catalytic activity and substrate specificity (*4*).

BFD is homologous to three ThDP-dependent enzymes (*5, 6*) that are involved in central metabolic pathways and whose crystal structures have recently been solved. The structure of yeast pyruvate decarboxylase (PDC) was determined to 2.3 Å resolution (*7, 8*), that of pyruvate oxidase was determined to 2.1 Å resolution (*9, 10*), and the structure

[†] This research was supported by Grants GM-40570 (to J.A.G., G.L.K., and G.A.P.) and GM26788 (to G.A.P. and D.R.) from the National Institutes of Health, Fellowship DRG-1194 of the Cancer Research Fund of the Damon Runyon-Walter Winchell Foundation (to M.S.H.), and in part by a grant from the Lucille P. Markey Charitable Trust.

* To whom correspondence should be addressed (at Purdue Univ.).

[‡] Atomic coordinates and structure factors have been deposited in the Protein Data Bank with the deposition code 1BFD and are scheduled for immediate release.

[§] Purdue University.

[||] Brandeis University.

[⊥] University of California.

[#] Present address: Bioanalytical Research, Novartis Pharma AG, 4002 Basel, Switzerland.

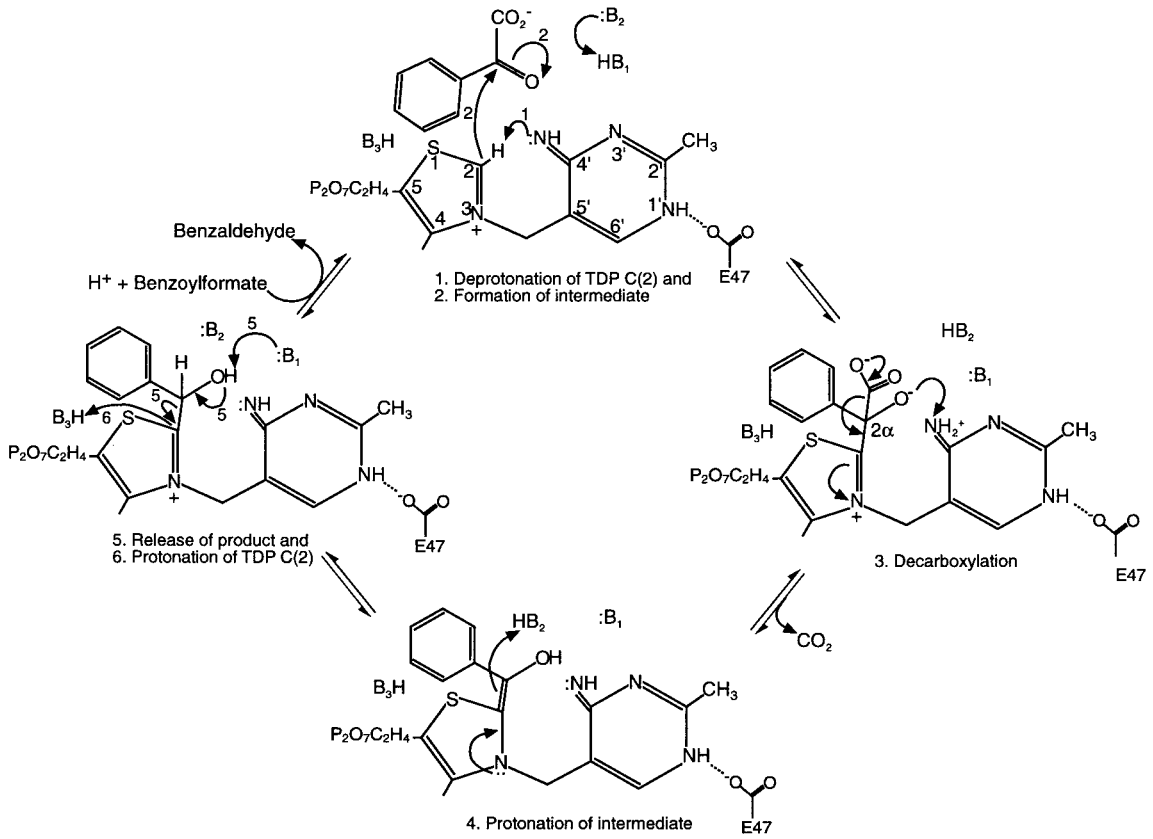
[▽] Present address: Victorian College of Pharmacy, Monash University, Parkville 3052, Australia.

[○] University of Illinois.

[1] Abbreviations: APMA, aminophenylmercuric acetate; BFD, benzoylformate decarboxylase; HEPES, *N*-(2-hydroxyethyl)piperazine-*N*′-(2-ethanesulfonic acid); MPD, 2-methyl-2,4-pentanediol; PDC, pyruvate decarboxylase; PEG 400, poly(ethylene glycol), average molecular mass 400 kDa; ppII, polyproline type II; rms, root-mean-square; ThDP, thiamin diphosphate; Tris, tris(hydroxymethyl)aminomethane.

S0006-2960(97)03047-X CCC: $15.00
Published on Web 06/23/1998

Scheme 2



of transketolase was determined to 2.0 Å resolution (*11*, *12*). The three enzymes share a common fold (*13*). Each enzyme is composed of three structurally similar domains, with each domain consisting of a β-sheet surrounded by α-helices. The active site is at the interface of two of the domains. The third domain binds an activator in PDC (*8*, *14*), and a second cofactor, FAD, in pyruvate oxidase (*9*). Although the three domains are in a different order in the primary sequence of transketolase, the chemical and structural details of cofactor binding are very similar in the three enzymes. In all three cases, ThDP is forced into a strained (V) conformation and is chemically activated by adjacent functional groups on the enzyme.

Catalysis by ThDP has been studied in great detail both on and off enzymes (for recent reviews, see refs *15*−*18*). Studies of benzoylformate decarboxylase itself suggest the involvement of at least two enzymic functional groups, one protonated and one unprotonated, in the course of the reaction (*19*). The accepted mechanism of catalysis by BFD is presented in Scheme 2. Abstraction of a proton from C2 of ThDP, which initiates the enzyme-catalyzed reaction, is accomplished by a base contributed by the ThDP itself, the 4′-amino group (*20*−*23*). The resulting ylide attacks the carbonyl carbon of benzoylformate to form the first tetrahedral complex. Decarboxylation results in a carbanion, stabilized by the formation of an enamine resonance form, which is protonated by an enzymic base. Dissociation of the second tetrahedral complex gives ThDP and benzaldehyde. The enzymic groups required by this mechanism have not yet been identified in BFD, nor have the analogous groups been unequivocally identified in other ThDP-dependent enzymes, although recently, progress has been made through analysis of site-directed mutants (*24*−*33*).

This study describes the crystal structure of BFD determined at 1.6 Å resolution. The physical description of the constellation of active-site residues of BFD is the first step toward the elucidation of their specific roles in ThDP-dependent catalysis. The structure of BFD is compared here to that of three other enzymes that utilize ThDP: pyruvate oxidase, PDC, and transketolase. Significant differences in the active site, even between the two decarboxylases, have implications for enzyme evolution. While the mode of cofactor binding is quite similar, there are, perhaps surprisingly, no other conserved residues in the active site. Yet, although catalytic residues are not conserved in the primary sequence of the enzymes, there remains a positional conservation of catalytic atoms in the active site relative to the cofactor and the protein scaffold.

EXPERIMENTAL PROCEDURES

Crystallization. BFD was purified and crystallized by vapor diffusion as described previously (*34*), except that 0.5% 2-methyl-2,4-pentanediol (MPD) was added to the well solution in all cases. Briefly, crystals were grown at room temperature by hanging-drop vapor diffusion against a well solution of 22% (v/v) poly(ethylene glycol) with an average molecular mass of 400 kDa (PEG 400), 0.15 M $CaCl_2$, 0.5% (v/v) MPD, and 0.1 M TrisCl (pH 8.5). Drops contained equal volumes (2 μL) of well solution and purified benzoylformate decarboxylase [10 mg/mL in 0.1 mM $MgCl_2$, 0.2 mM ThDP, and 25 mM NaHEPES (pH 7.0)]. Crystals are of space group *I*222, with unit cell lengths $a = 82.3$ Å, $b = 96.6$ Å, $c = 138.3$ Å.

Data Collection and Processing. Crystals were kept at a temperature of 4 °C during X-ray experiments with a stream

Table 1: Data Collection Statistics

crystal	native	HGI	PTCL6	PTCL8	APMA
cell dimension *a* (Å)	82.3	82.4	82.1	82.2	82.3
cell dimension *b* (Å)	96.6	96.4	96.9	97.0	97.0
cell dimension *c* (Å)	138.3	138.5	138.1	138.2	138.4
high-resolution limit (Å)	1.6	1.6	1.8	1.8	1.8
number of reflections $I/\sigma(I) > 1$					
total	217 078	143 607	93 158	114 799	183 555
unique	59 179	54 102	38 927	44 864	48 032
completeness (%)	81	74	77	87	93
completeness to 2.5 Å (%)	98	95	87	93	98
R_{merge} (%)[a]	5.1	5.1	5.4	4.9	8.0
average I/σ to 2.5 Å	17.1	13.8	8.6	15.2	9.7
R_{iso} to 2.5 Å (%)[b]		8.6	9.1	8.7	6.7
number of sites		3	3	5	3
F_H/E[c](6.2−4.8 Å)		2.3	1.3	1.6	1.9
F_H/E[c](2.8−2.5 Å)		1.1	0.74	0.68	0.76

[a] $R_{merge} = \sum\sum_i|I_i - \langle I\rangle|/\sum\langle I\rangle$, where I_i is the scale factor-corrected intensity for a reflection and $\langle I\rangle$ is the mean intensity for that reflection (calculated to high-resolution limit). [b] Mean fractional isomorphous difference, $R_{iso} = \sum(|F_{PH}| - |F_P|)/\sum(|F_P|)$, where F_{PH} and F_P are the derivative and native structure factors, respectively. [c] $F_H/E = [\sum F_H^2/\sum(F_{PHobs} - F_{PHcalc})^2]^{1/2}$.

Table 2: Heavy-Atom Positions

derivative[a]	x[b]	y	z	B[c] (Å^2)	occupancy[d]	ligand
HGI	0.023	0.246	0.068	15.3	0.054	C300
	0.132	0.369	0.157	13.3	0.028	C398
	0.233	0.092	0.136	32	0.027	C248
APMA	0.234	0.095	0.137	24.5	0.047	C248
	0.023	0.25	0.07	23.6	0.007	C300
	0.257	0.095	0.209	64.3	0.006	*e*
PTCL6	0.359	0.733	0.161	31.4	0.034	M365
	0.397	0.72	0.169	29	0.034	M365
	0.195	0.369	0.245	9.29	0.038	M453
PTCL8	0.369	0.275	0.997	63.8	0.046	H172
	0.406	0.649	0.053	122.2	0.027	N-terminus
	0.359	0.733	0.162	24	0.024	M365
	0.398	0.72	0.169	35.9	0.046	M365
	0.195	0.368	0.245	7.4	0.033	M453

[a] Compounds are described in the Experimental Procedures. [b] Fractional atomic coordinates. [c] Isotropic temperature factor. [d] Arbitrary scale. [e] No obvious ligand.

of chilled air. Data were collected on an R-AXIS IIC image plate detector using 0.3 mm collimated monochromatized Cu Kα radiation from a Rigaku RU-200 rotating anode generator (50 kV × 150 mA). Each data set was collected from a single crystal. The diffraction images were reduced to integrated indexed intensities with the R-AXIS processing software PROCESS (T. Higashi, Rigaku Corporation, 1990) and MOSFLM (A. J. Wonocott, Imperial College, 1980). Data for which $[I/\sigma(I)] \geq 1$ were used for further analysis.

Structure Determination. The structure of benzoylformate decarboxylase was solved by the method of multiple isomorphous replacement. Heavy atoms were introduced to the crystals either by soaking or cocrystallization, with the latter method being more effective. Four derivatives were obtained. The derivative HGI was obtained by the standard crystallization method with the addition of 0.1 mM K_2HgI_4 to the well solution before setting up the drop. The derivative APMA was obtained by the standard crystallization method with the addition of 0.1 mM aminophenylmercuric acetate to the well solution before setting up the drop. The derivative PTCL6 was obtained by soaking a native crystal overnight in a solution of 26% PEG 400, 0.5% MPD, 0.15 M $CaCl_2$, 0.1 M NaHEPES (pH 7.5), and 20 mM K_2PtCl_4. The derivative PTCL8 was obtained by soaking a native crystal in 26% PEG 400, 0.5% MPD, 0.01 M $CaCl_2$, and 0.1 M NaHEPES (pH 7.5) for 3 h, and the same buffer with the addition of 10 mM K_2PtCl_4 overnight. Data collected on native and derivative crystals are described in Table 1. Unless otherwise noted, the PROTSYS package of programs was used for further calculations in the structure solution.

The positions of the mercury atoms in the derivative HGI were determined from inspection of an isomorphous difference Patterson map and confirmed by the direct-methods program SHELX (*35*). The positions of the heavy atoms in the remaining three derivatives were determined by inspection of difference Fourier maps using phases derived from the refined mercury sites in HGI, and confirmed by difference Patterson maps. The anomalous scattering contribution of the derivative HGI was used to determine the correct hand of the data and was used in all subsequent calculations. The positions of the heavy atom sites were refined (Table 2) with the programs HEAVYREF (T. C. Terwilliger) and TENEYCK (L. F. Teneyck; both programs modified by G. Petsko), and protein phases were calculated to 2.5 Å resolution, with an overall mean figure of merit of 47%. After several rounds of solvent flattening using the programs BCWANG (*36*), modified by G. Petsko, and SQUASH (*37*), the figure of merit improved to 89%. The electron density map calculated from these phases was quite good.

A model of the protein was first built as a polyalanine chain into interpretable regions of the electron density map using the graphics program O (*38*). Side chains were also built in as they became identifiable from comparison of the map with the known protein sequence. After approximately 50 residues or side chains were built, positional least-squares refinement of the model was carried out using the program XPLOR (*39*). The phases were improved by combining the solvent-flattened heavy atom phases with the calculated phases from the partial model using the program COMBINE (W. Kabsch; modified by G. Petsko). After 11 rounds of model building and phase improvement, the partial model (approximately 437 residues) was refined by simulated annealing in XPLOR (*39*) against the native data from 15 to 1.6 Å resolution. Inclusion of the higher resolution data improved the electron density maps calculated from the model and allowed the correction of several important sequence errors (see below), the placing of water molecules, and the detection of alternate conformations for several

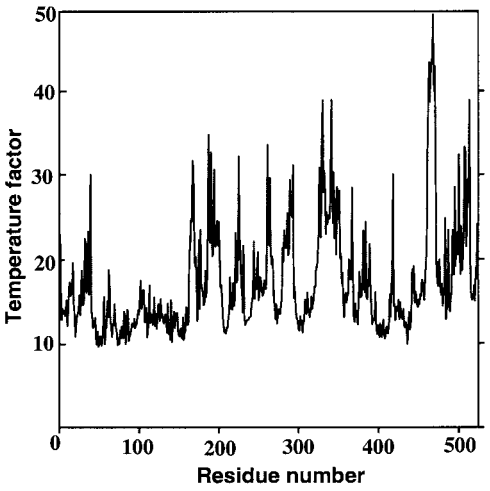


FIGURE 1: Plot of mean temperature factor, as a function of residue number, for all nonhydrogen atoms.

residues. Geometric analysis of the model was carried out with the programs PROCHECK (*40*) and OOPS (part of the O package). The entire structure was then checked with a series of simulated-annealing omit maps (*41*) with 60 contiguous residues omitted in the calculation of each map. Individual isotropic *B*-factors were refined with restraints. Thirteen additional cycles of refinement and manual rebuilding yielded the final model, which contains residues 2−524 of the protein, the cofactor thiamin diphosphate, two Ca^{2+} ions, one Mg^{2+} ion, and 345 water molecules (Table 3). The final *R*-factor is 15.0%, and the final R_{free} (calculated using 5% of the total reflections) is 18.6% for all data with $[I/\sigma(I)] \geq 1$ from 5 to 1.6 Å resolution. Root-mean-square (rms) deviations from ideality for bond lengths is 0.009 Å and for bond angles is 1.8°. No residues are in the disallowed regions of a Ramachandran plot, and only two residues, Ser 71 and Pro 121, are in generously allowed regions. The average temperature factor (*B*-factor) for each residue is shown in Figure 1. Areas of high *B*-factor are found exclusively in external loops.

Superposition of Structures. The BFD structure was aligned with structures of related enzymes using the LSQ facility in the graphics program O. For each pairwise comparison, the initial superposition was based on the conserved metal-binding sequence GDG (residues 427−430 in BFD). First, the sum of the squares of the distances between corresponding α-carbons in this region of two structures was minimized. The structural alignment was subsequently improved, again in O, by linking proximate secondary structural elements of the two proteins. The least-squares minimization was repeated using the expanded list of corresponding α-carbon atoms.

DNA Sequencing and Mass Spectrometry. Using appropriate primers, the *mdlC* structural gene was sequenced in both directions by the Biomolecular Resource Center DNA Sequencing Facility at the University of California, San Francisco. Mass spectrometric analysis of BFD was carried out by Dr. David King at the Howard Hughes Medical Institute, University of California, Berkeley, using a Hewlett-Packard HP5989A electrospray mass spectrometer.

Generation of Figures. Figures portraying protein models were rendered using RAYSCRIPT (E. Fontano, D. Peisach, and E. Peisach; http://www.sb.fsu.edu/software/rayscript.html), which takes input suitable for MOLSCRIPT (*42*) and converts it to input for the program RAYSHADE (L. Coffin and D. Debry; http://www-graphics.stanford.edu/~cek/rayshade/rayshade.html). Figures 3 and 12 were rendered in MOLSCRIPT.

Table 3: Refinement Statistics

resolution	5.0−1.6 Å
unique reflections with $[I/\sigma(I)] \geq 1$	59 179
completeness	81%
final *R*-factor	15.0%
final free *R*-factor	18.6%
rms deviation of bond lengths	0.009 Å
rms deviation of bond angles	1.8°
avg main-chain *B*-factor	16 Å^2
avg side-chain *B*-factor	19 Å^2
avg water *B*-factor	30 Å^2
avg ThDP *B*-factor	14 Å^2
avg metal *B*-factor	16 Å^2
total number of non-hydrogen protein atoms	3949
total number of water molecules	346
total number of non-hydrogen ThDP atoms	26
total number of Mg^{2+} ions	1
total number of Ca^{2+} ions	2

RESULTS

Description of the Overall Structure. BFD is composed of three domains, each built around a central β-sheet. The α− and γ-domains are topologically equivalent, while the β-domain has a slightly different fold. Residues that make contact with the cofactor ThDP are located at the C-terminal ends of strands in the α- and γ-domain sheets. A detailed schematic diagram of the structure of BFD is presented in Figure 2, and a ribbon diagram of the monomer is shown in Figure 3.

The architecture of BFD is shared by the other enzymes of known structure that utilize the cofactor ThDP: pyruvate decarboxylase (*7*), pyruvate oxidase (*10*), and transketolase (*12*). The structure of BFD was superimposed on each of the other three ThDP-containing structures. The final root-mean-square distances between α-carbon pairs for the superpositions are as follows: for BFD/pyruvate oxidase, 1.8 Å for 404 α-carbon pairs; for BFD/PDC, 1.9 Å for 352 α-carbon pairs; and for BFD/transketolase, 2.0 Å for 248 α-carbon pairs. The overall arrangement of elements of secondary structure is almost identical in BFD and PDC.

Quaternary Structure. The enzymes BFD, PDC, and pyruvate oxidase are tetramers, while transketolase is a dimer. ThDP is bound at a dimer interface in BFD, as it is in the other three enzymes; the binding site is formed by the α-domain of one monomer and the γ-domain of another. Thus, it is useful to consider the BFD tetramer as a dimer of dimers (Figure 4). The tetramer shows exact 222 symmetry, as required by the fact that the asymmetric unit contains one monomer and the tetramer is built up from crystallographic symmetry elements in the space group *I*222.

The two dimers that make up the BFD tetramer are much more intimately associated than those that make up the PDC tetramer (*8*, *43*); in this sense, BFD is much more similar to pyruvate oxidase (*9*, *10*). The major difference in topology between BFD and PDC is in the second half of the β-domain β-sheet, which is involved in tetramer contacts (Figure 5).

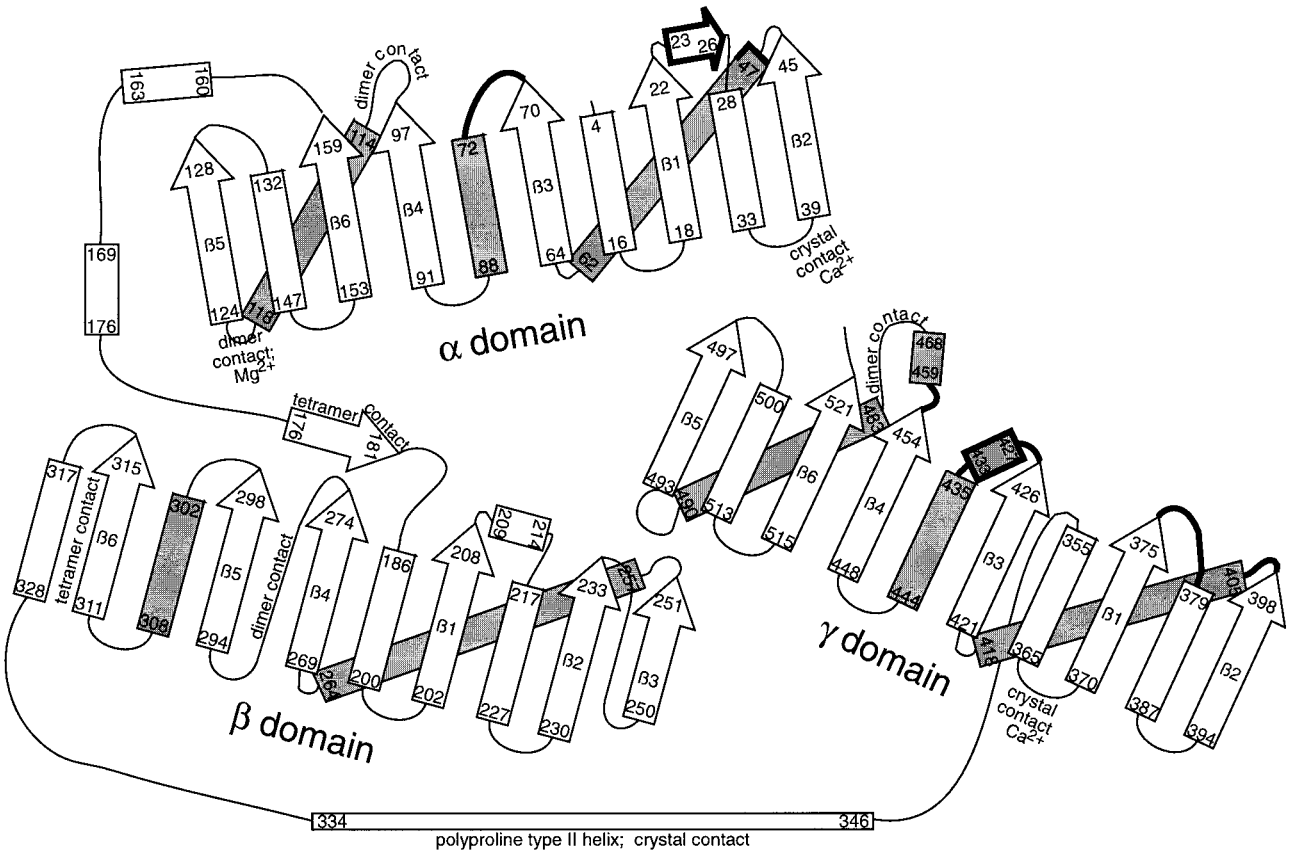


FIGURE 2: Schematic diagram of the BFD structure. Arrows indicate β-strands; rectangles indicate helices. Beginning and ending residues for each element of secondary structure are indicated. Major contacts (between monomers in the dimer and tetramer and crystal contacts) are indicated. Contacts with the cofactor ThDP are indicated in bold. Helices in white are in front of the β-sheets; shaded helices are behind the sheet.

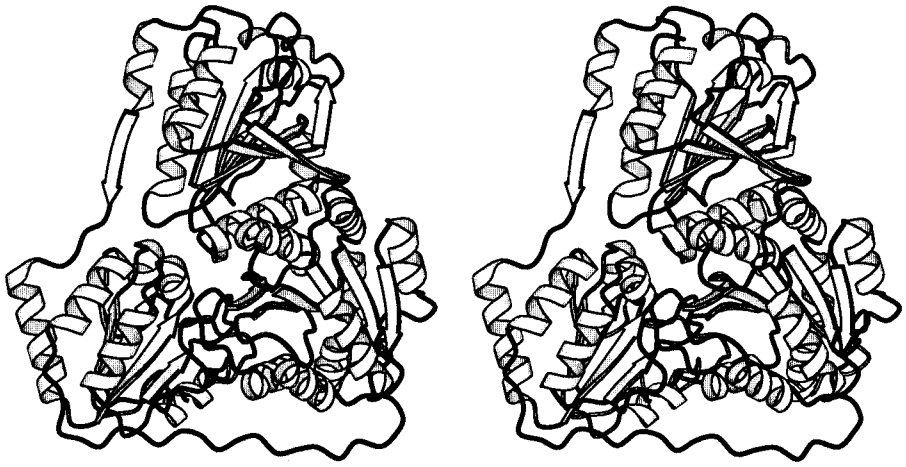


FIGURE 3: Ribbon diagram of the structure of a monomer of BFD. This stereo diagram (defocused representation) is in approximately the same orientation as the schematic diagram in Figure 2, with the α-domain on top, the β-domain at lower left, and the γ-domain at lower right.

The topology of the β-domain is significantly different from the α- and γ-domains in PDC (*8*), while in BFD, the three domains are more similar. In PDC, the last three strands in the β-domain sheet are antiparallel. Contact between dimers in a tetramer of PDC is mediated by rather tenuous contacts between the last strand in the β-domain of each dimer. In contrast, in BFD, there are no antiparallel strands in the β-sheets of any of the three domains, and the β-domain of BFD has one strand fewer than the β-domain of PDC. An extra β-strand is supplied by the other dimer in the tetramer, where it is part of the connection between the α- and β-domains (Figures 2 and 5). Thus, two contacts in the PDC tetramer (the two β-domains of one dimer with the two β-domains of the other) are replaced by four more intimate connections in the BFD tetramer (the two β-domains of each dimer with the α/β-connections of the other). The disposition of dimers in the tetramer is very similar in BFD and pyruvate oxidase, except that in pyruvate oxidase, the β-strand from

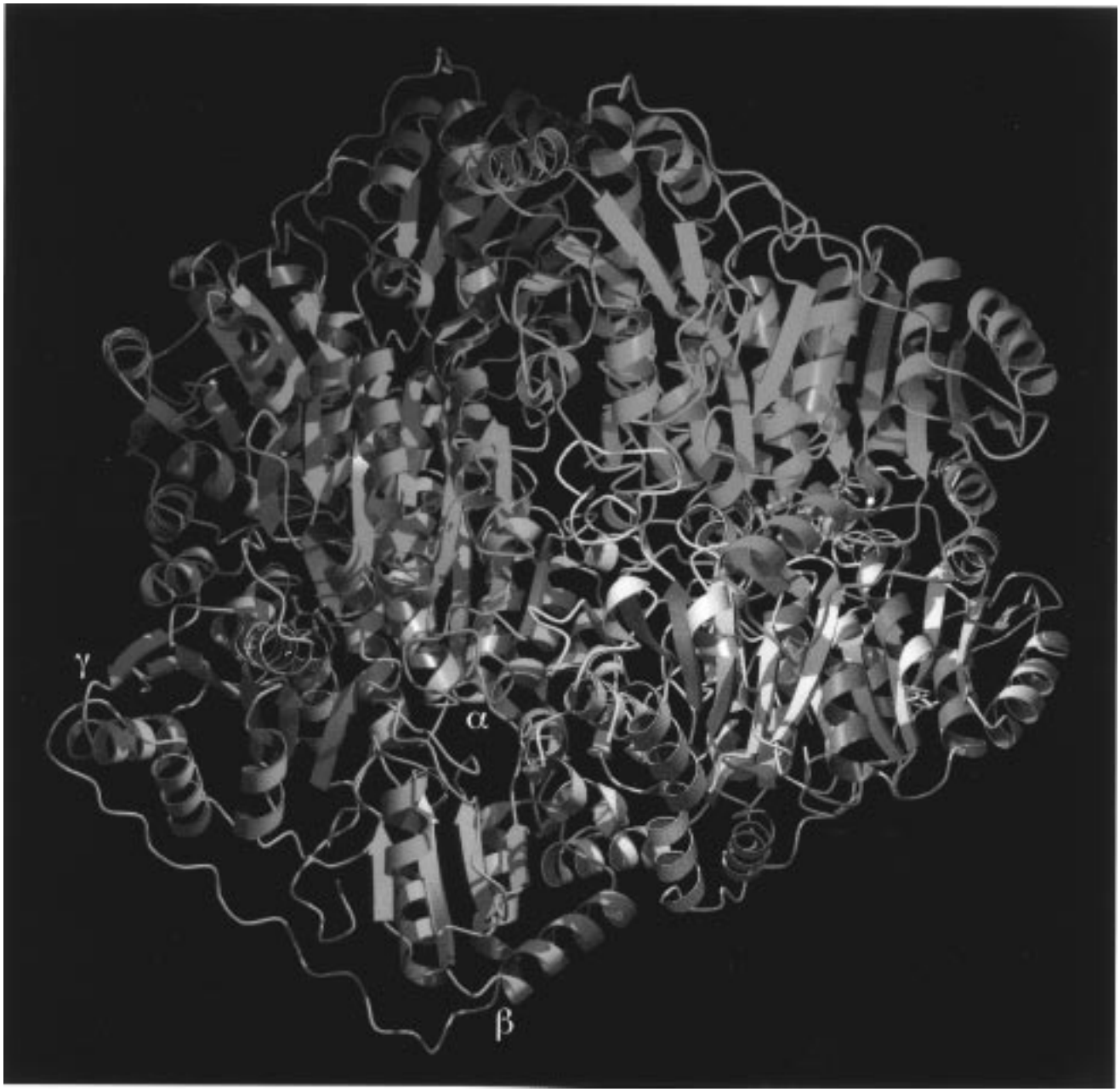


FIGURE 4: Ribbon diagram of the structure of a tetramer of BFD. Two active sites are at the interface between the red and purple monomers, and two are at the interface between the yellow and green monomers. These interfaces are referred to in the text as "dimer interfaces"; other contacts are referred to as "tetramer contacts." The cofactor ThDP and its associated metal ion are shown as a ball-and-stick representation. Domains are labeled in the red monomer.

the second dimer does not form the requisite hydrogen bonds to be part of the β-sheet of the β-domain in the first dimer.

Correction of the Sequence of BFD. During the course of building and refining the model of BFD, it became quite clear that there were discrepancies between the electron-density map and the protein sequence predicted from the original DNA sequence of the *mdlC* gene, which encodes BFD. Analysis of the high-quality electron-density map allowed the prediction of several errors in the published sequence of the *mdlC* gene (*6*), including a frame shift that changes the last 41 residues of the predicted sequence (of 499 residues) of BFD, and adds another 28 residues to its C-terminus. An additional residue was also inserted in the region 382−386 and the sequence modified. To confirm these changes, the *mdlC* gene was resequenced. The new protein sequence, consistent with both the electron density map and the *mdlC* gene sequence, is shown in Figure 6. Analysis of BFD by ion-spray mass spectrometry gives a molecular weight of 56 236 ± 7 (data not shown), which is consistent with the new sequence (predicted molecular weight of 56 224 without the N-terminal methionine) but not with the originally published sequence (predicted molecular weight of 53 484). The corrected BFD sequence is 25% identical to the sequence of pyruvate oxidase and 21% identical to the sequence of PDC.

Active Site. Binding of the cofactor is similar in BFD (Figure 7) and other ThDP-dependent enzymes. The ligands to the essential metal ion (Figure 10A), the glutamate (Glu 47 in BFD) that binds to N1′ of the ThDP, the main-chain carbonyl (Gly 401 in BFD) that interacts with the N4′ amino group, and the general hydrophobic character of the site are conserved. ThDP is bound in the "V" conformation (*44*) in

FIGURE 5: The major difference between the secondary structures of BFD and PDC contributes to the difference in the quaternary structures of the tetramers formed by the two enzymes. PDC (blue) contains an extra strand at the end of the β-sheet in its β-domain. The strand in BFD (orange) is replaced by a strand (between the α- and β-domain) from a neighboring monomer in the tetramer (yellow), resulting in a more intimate association between dimers in the tetramer.

BFD, as in the other enzymes, with similar values for the two torsion angles defining the relative orientation of the thiazolium and pyrimidine rings: $\phi_t = 95.4°$ and $\phi_t = -64.7°$.

Inspection of the structure of the active site of BFD suggests the identity of residues that may be involved in catalysis (Figure 8). A covalent bond forms between the substrate Cα atom and the C2 atom of ThDP during the course of the reaction (*45*). A water molecule bound in the active site, 3.2 Å from the C2 atom of ThDP, may be near the position of the Cα atom of the substrate upon binding to the enzyme. This water molecule forms hydrogen bonds with the hydroxyl group of Ser 26 and the imidazole ring of His 70. The ability of His 70 to act as a catalytic base is probably increased by the contact between the other side of the imidazole ring and the carboxylate group of Glu 28. His 281 is 5.3 Å from the water molecule bound to the cofactor C2, and may also be involved in catalysis. Surprisingly, none of the residues that are likely to be involved in catalysis (aside from residues directly bound to the cofactor) are conserved between any pair of the three enzymes BFD, PDC, and pyruvate oxidase (Figure 9, Table 4).

Binding Sites for Divalent Cations in BFD. As in other ThDP-dependent enzymes, a divalent metal ion in the active site of BFD binds to the two phosphate groups of the cofactor. The metal ion is bound in an octahedral coordination by two phosphate oxygens of the ThDP, the side chains of Asp 428 and Gln 455, the main-chain carbonyl of Thr 457, and a water molecule (Figure 10A). This coordination is very well conserved in pyruvate oxidase, PDC, transketolase, and BFD.

Two additional divalent cations were identified in the BFD structure. The first is at a crystal contact, has seven ligands (Figure 10C), and was tentatively identified as a Ca^{2+} ion because of the number of ligands and because of the requirement of Ca^{2+} for crystal growth and stability (data not shown). The second divalent cation is on a crystallographic 2-fold axis, at the dimer interface (the same interface that forms the active site), and has six backbone carbonyl ligands in octahedral coordination. The identity of this atom was uncertain, as the crystallization mix contains two divalent cations, Mg^{2+} (approximately 0.1 mM on achieving equilibrium between drop and well) and Ca^{2+} (approximately 150 mM on achieving equilibrium).

An anomalous difference electron-density map was analyzed to assist in the identification of the three divalent cations as Ca^{2+} or Mg^{2+}, exploiting the fact that Ca^{2+} is a much stronger anomalous scatterer than Mg^{2+}. The map was calculated as a Fourier transform in XPLOR, using phases derived from the protein model (shifted by 90°) and amplitudes derived from the anomalous differences in native data scaled to retain anomalous information. The largest peak in the map, 5.2σ (standard deviations above the mean density of the map), corresponds to the cation bound to ThDP (Figure 10A). The peak corresponding to the cation at the crystal contact (Figure 10C) is 5.0σ. Thus, these two cations are almost certainly Ca^{2+}. The peak corresponding to the cation at the dimer contact (Figure 10B) is considerably lower, at 4.2σ, and thus may be at least partially occupied by a Mg^{2+} ion. For comparison, the sulfur atoms of the six cysteinyl residues in the protein correspond to peaks of 4.2−4.5σ, the sulfur atoms of 11 of the 13 methionyl residues correspond to peaks of 4.2−4.5σ (the other two are more mobile, with the sulfur atom in Met 79 having two distinct conformations, and the sulfur atom of Met 222 having a *B*-factor of 36 Å^2), the sulfur atom of the ThDP corresponds to a peak of 4.3σ, and the phosphorus atoms of the ThDP correspond to peaks of 4.3 and 4.4σ. Two water molecules (709 and 793) were found to correspond to 4.4σ peaks in the anomalous difference map; these sites may also be partially occupied by Ca^{2+} ions.

Our identification of the major atom bound to each of the three metal-binding sites as Ca^{2+} or Mg^{2+} is supported by analysis of their behavior during refinement. First, the refined temperature factors of the metals are close to that of their ligands. For the Ca^{2+} bound to ThDP, $B = 15.6$ Å^2 for the metal and 14.8 ± 1.3 Å^2 for the ligand atoms. For the Mg^{2+} at the dimer interface, $B = 10.6$ Å^2 for the metal and 11.4 ± 0.8 Å^2 for the ligand atoms. For the Ca^{2+} at the crystal contact, $B = 21.2$ Å^2 for the metal and 25 ± 10 Å^2 for the ligand atoms. As an additional confirmation, the *B*-factor of each metal was set to the average *B*-factor of its ligand atoms and fixed, and the metal occupancies were refined. An occupancy of 1 is expected if the atom has been correctly identified. The final occupancy values were 1.0 for the Mg^{2+} and Ca^{2+} within the tetramer and 1.2 for the Ca^{2+} at the crystal contact, indicating correct identification.

An Unusual cis-Peptide Bond. A cis-peptide bond that does not involve a proline or a glycine is observed in the BFD structure between Val 277 and Phe 278 (Figure 11). Non-proline cis-peptide bonds are rare in proteins and are almost exclusively found in the active sites of enzymes (*46*); this general rule is also supported by more recent examples (*47*−*61*). This bias may represent the reality of what is present in protein structure or, alternatively, the fact that investigators are likely to spend more time and care on

```
  1  MASVHGTTYE LLRRQGIDTV FGNPGSNELP FLKDFPEDFR YILALQEACV

 51  VGIADGYAQA SRKPAFINLH SAAGTGNAMG ALSNAWNSHS PLIVTAGQQT

101  RAMIGVEALL TNVDAANLPR PLVKWSYEPA SAAEVPHAMS RAIHMASMAP

151  QGPVYLSVPY DDWDKDADPQ SHHLFDRHVS SSVRLNDQDL DILVKALNSA

201  SNPAIVLGPD VDAANANADC VMLAERLKAP VWVAPSAPRC PFPTRHPCFR

251  GLMPAGIAAI SQLLEGHDVV LVIGAPVFRY HQYDPGQYLK PGTRLISVTC

301  DPLEAARAPM GDAIVADIGA MASALANLVE ESSRQLPTAA PEPAKVDQDA

                       old sequence  GKC-AA
351  GRLHPETVFD TLNDMAPENA IYLNESTSTT AQMWQRLNMR NPGSYYFCAA

401  GGLGFALPAA IGVQLAEPER QVIAVIGDGS ANYSISALWT AAQYNIPTIF

old sequence  R FAMVCRRSRS RKRSWAGCAR DRLPRTRQGL WCPSAESRQP*
451  VIMNNGTYGA LRWFAGVLEA ENVPGLDVPG IDFRALAKGY GVQALKADNL

501  EQLKGSLQEA LSAKGPVLIE VSTVSPVK
```

FIGURE 6: Corrected amino acid sequence of BFD from *P. putida*. The original sequence predicted from nucleotide sequencing (*6*) is shown in italics where it differs from the corrected sequence. The C-terminus of each sequence is indicated with an asterisk.

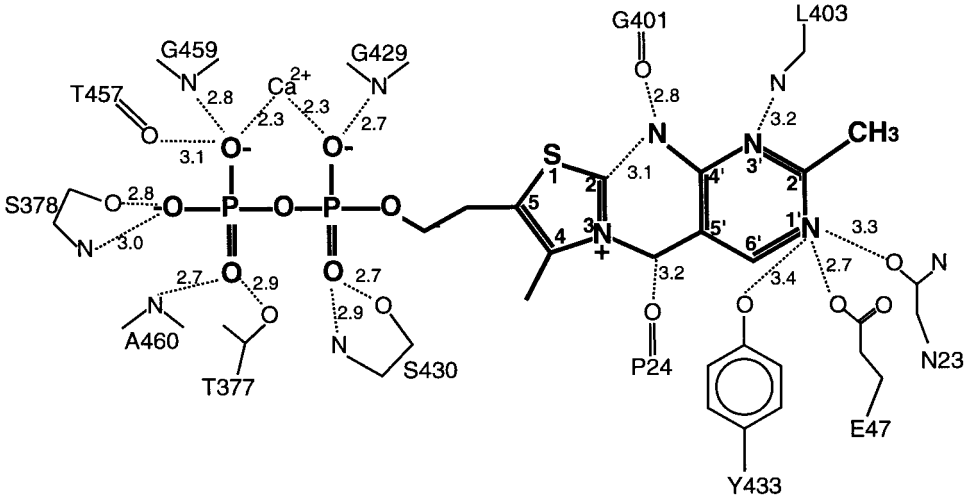


FIGURE 7: Schematic diagram of cofactor binding in BFD (not to scale). All residues that are within 3.2 Å of the ThDP, and selected residues within 3.5 Å, are shown. Distances of closest approach to the cofactor are given in angstroms. Atoms in the thiazolium and pyrimidine rings of the ThDP are numbered. Hydrogen atoms and charges on protein side chains are not shown. Residues numbered less than 100 belong to one monomer, and those numbered more than 100 belong to another.

building the parts of their model in which they are most interested, such as the active site, and so are more likely to find unexpected conformations there. In BFD, the *cis*-peptide bond is not found in the active site. The significance of this unusual backbone conformation is unknown in this case. However, its presence allows both Val 277 and Phe278 to pack in the same hydrophobic core.

A Long, Regular Polyproline Type II Helix. An unusually long and regular example of a left-handed polyproline type II (ppII) helix connects the β- and γ-domains of BFD (Figures 2, 3, and 12). The 13-residue helix lies flat against the surface of the protein and is accessible to solvent. Interestingly, the major crystal contact between tetramers in the crystal is mediated by the ppII helix (Figure 13). Most of the specific contacts in this region are, in turn, mediated by bound water molecules. There is a single salt bridge, between Lys 345 in the polyproline helix and Asp18. The non-active-site Ca^{2+} ion is also bound at this interface (Figure 10C).

DISCUSSION

Conservation of chemical strategies for catalysis by related enzymes is clearly a powerful evolutionary force. Consequently, one would expect that enzymes that retain 20−25% sequence identity and catalyze similar overall reactions, albeit on different substrates, would share a nearly identical array of catalytic residues in their active sites. The determination of the structure of benzoylformate decarboxylase presents a case where this expectation is not fulfilled. None of the residues in the active site, besides those bound directly to the cofactor, are conserved between BFD and PDC, or between any pair of the four ThDP-dependent enzymes of known structure (a group which includes pyruvate oxidase and transketolase; Table 4).

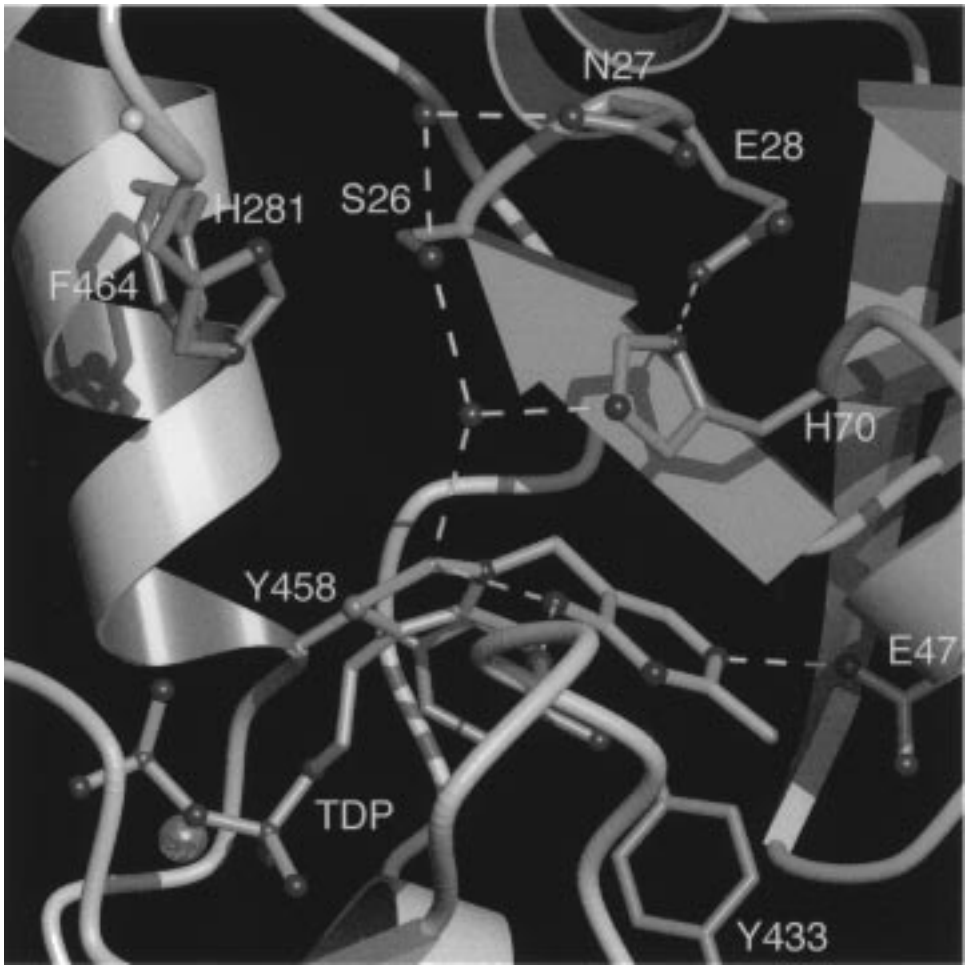


FIGURE 8: The active site of BFD. The two monomers that form the active site are shown as ribbon diagrams, one in pink and one in yellow. Atomic models are shown for the ThDP, for the active-site metal ion and for residues whose position in the active site imply possible involvement in catalysis. Contacts ranging from 2.7 to 3.2 Å are indicated by dashed yellow lines.

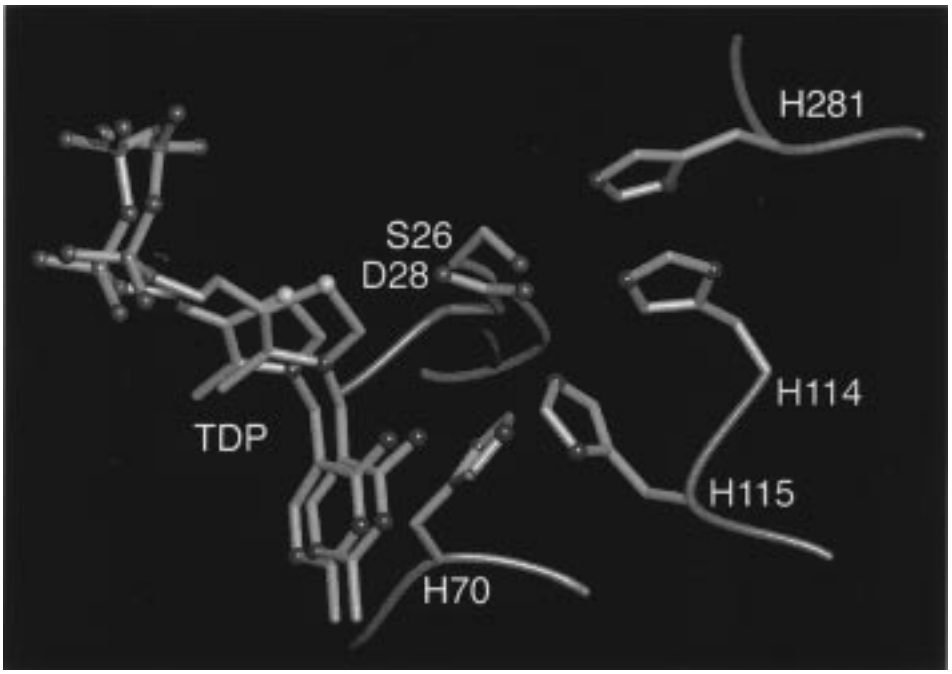


FIGURE 9: Comparison of the active sites of BFD and PDC. PDC (Brookhaven Protein Data Bank entry 1pvd) is shown in blue, BFD in pink. The two proteins were superimposed by least-squares minimization of the distances between 352 paired C_α atoms.

Two possible explanations for such a pattern of conservation are (i) the mechanisms of this family of enzymes are dominated by the chemistry of the cofactor and (ii) the mechanisms differ substantially because of differences in the

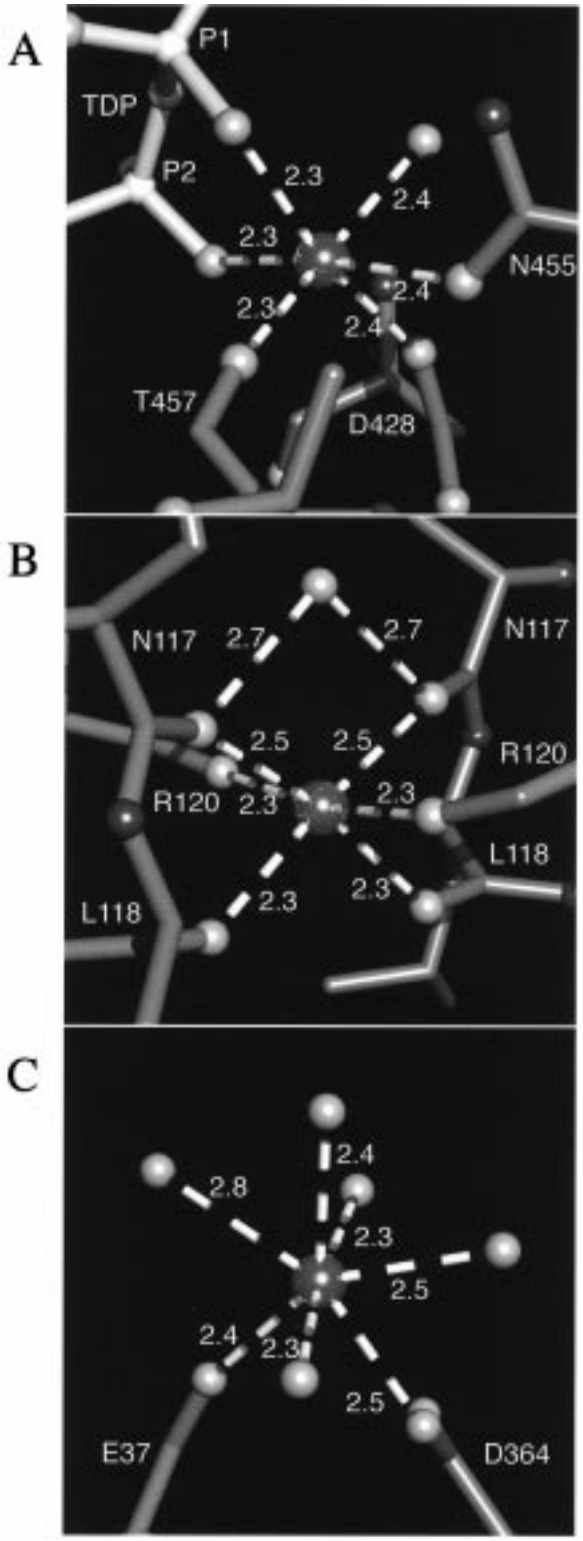


FIGURE 10: Metal-binding sites in BFD. Panel A shows the metal ion (Ca^{2+}) bound in the active site. Panel B shows a metal ion (Mg^{2+}) at the dimer interface. Protein atoms to the right of the metal ion belong to one monomer, while protein atoms to the left of the metal ion belong to a second monomer related by a crystallographic 2-fold axis. Panel C shows a metal ion (Ca^{2+}) at a crystal contact. Metal−ligand distances are indicated in angstroms. Oxygen atoms are light gray, nitrogen atoms are dark gray, phosphorus atoms are white, and metal ions are reflective.

chemical properties of the substrates, intermediates, or transition states of the reactions catalyzed by this family of enzymes. In this case, both factors are almost certainly at work. First, the prime importance of the cofactor in catalysis is clear from the fact that the initial step of the reaction, abstraction of the proton from the C2 carbon of ThDP, is catalyzed by the N4′ amino group of the cofactor itself (*20*). Second, due to the electron-withdrawing nature of the phenyl group of benzoylformate, the chemical characteristics of the intermediates in the reactions catalyzed by BFD and PDC differ markedly. The compound 2-(1-hydroxybenzyl)thiamin, very similar to the intermediate in the BFD-catalyzed reaction, is subject to an apparent internal redox process in neutral or acidic solutions that results in cleavage between the thiazole and pyrimidine portions of thiamin (*62*, *63*). This type of fragmentation is not observed in the analogous compound for the PDC-catalyzed reaction, 2-(1-hydroxyethyl)thiamin. The active-site environment in BFD may prevent destructive fragmentation of the cofactor by lowering the p*K* of the conjugate acid of the pyrimidine (*17*). This may be accomplished partially through conserved interactions (between N1′ and Glu 47 and between N1′ and the backbone N of Leu 403; Figure 7) and partially through interactions that are unique to BFD. Tyr 433 and Asn 23 of BFD are both within 3.5 Å of the pyrimidine ring (Figure 7), while analogous interactions are not found in the PDC active site. Thus, similarities in the overall reaction catalyzed by two enzymes may hide crucial differences in the particular aspects of the reaction chemistry that the enzymes must control during catalysis.

Table 4: Correspondence of Active-Site Residues in BFD, PDC, and Pyruvate Oxidase (POX)

	BFD	PYD	POX
Conserved Residues			
metal binding	N455	N471	N474
	D428	D444	D447
ThDP binding	Y458	Y474	Y477
phosphate binding	G427	G433	G446
	G429	G445	G448
interaction with N1′	E47	E51	E59
Nonconserved Residues			
	S26	D28	G35
	N27	F29	S36
	E28	N30	I37
	H70	T73	S82
	L109	H114	F121
	L110	H115	Q122
	Y433	L449	M452
	L461	E477	I480

Although none of the potential catalytic residues, besides those involved in cofactor binding, are identical in BFD and PDC, a more detailed comparison of the two active sites reveals two aspects of the arrangement of catalytic residues in the active site that may in fact be conserved (Figure 9), as follows.

(i) *Positional Conservation of Functional Groups*. Two histidinyl residues are present in each active site: His 70 and His 281 in BFD and His 114 and His 115 in PDC. Although the residues originate in very different places in the primary sequence of the two proteins, the functional groups on the imidazole rings are in nearly the same positions (as close as 3 Å apart) in the active sites relative to the protein backbone and the cofactor.

(ii) *Positional Conservation of Critical Residues in the Protein Scaffold.* Ser 26 of BFD and Asp 28 of PDC are both in position to perform some role in catalysis. Although

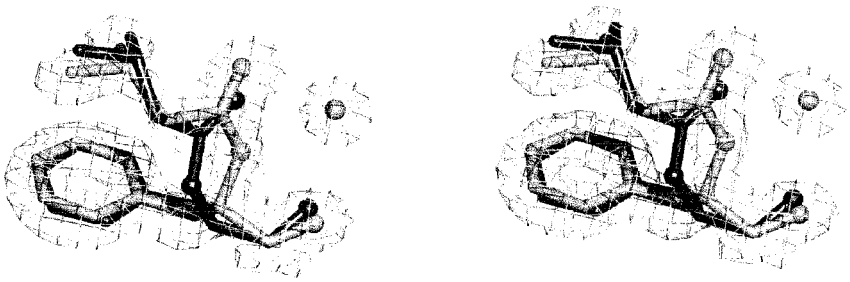


FIGURE 11: Stereo diagram (defocused) of a *cis*-peptide bond that does not involve a proline. The electron density map was calculated with the coefficients $2f_o - f_c$ using the model of the normal *trans*-peptide bond (black). A model constructed with a *cis*-peptide bond, accommodated by the electron-density map, is shown in gray.

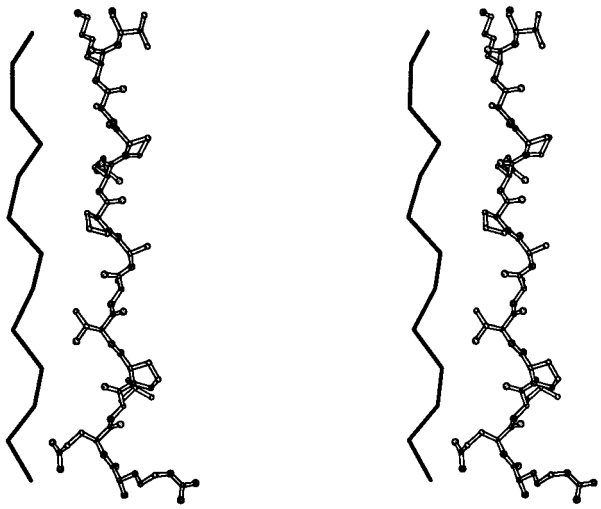


FIGURE 12: Stereo diagram (defocused) of an extended polyproline type II helix in BFD. A tracing of C_α atoms in residues 334−346 of BFD is shown on the right. A model of all nonhydrogen atoms in the same region is shown on the left.

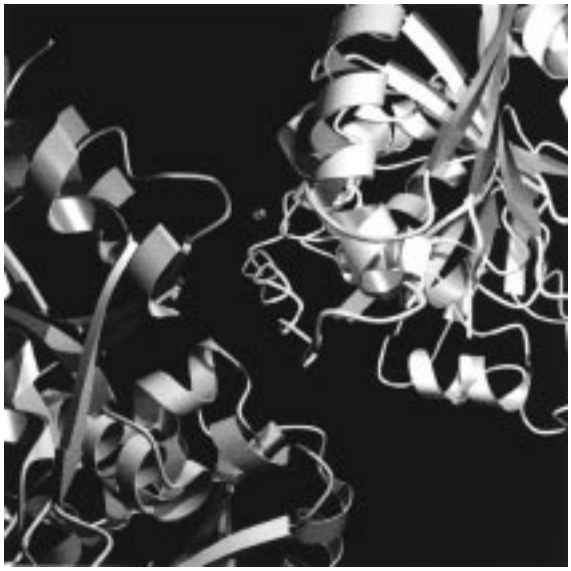


FIGURE 13: The major contact site between tetramers in the BFD crystals. The polyproline type II helix is shown in the center. The Ca^{2+} at the crystal contact is shown as a ball just above center (for a more detailed view of the metal binding, see Figure 9C).

these two residues have different identities and presumably also have different roles, they are conserved in one sense: they are at the identical positions in the protein sequence and fold. Examples of both of these types of positional conservation have been noted in the enolase superfamily (*64*; Hasson, M. S., Kenyon, G. L., Babbitt, P. C., Gerlt, J. A., Petsko, G. A., and Ringe, D., unpublished observations), and the first has been noted in the thioredoxin superfamily (*65*). Further examples of positional conservation without total identity of catalytic residues may become apparent, and perhaps widespread, as the structures of diverse members of other families of enzymes become available. Assignment of residues to a particular function based on sequence comparisons alone is apt to be dangerous if this phenomenon is common at low levels of sequence identity.

The initial step in the reaction, proton abstraction from the C2 position of ThDP, is probably accomplished by the 4′ amino nitrogen of the cofactor itself, as has been postulated for other ThDP-dependent enzymes (*7*, *20*−*23*). This intramolecular catalysis is facilitated by similar physical and chemical influences on the cofactor exerted by each ThDP-dependent enzyme. In all cases, the cofactor is forced into a V conformation by the enzyme, forcing N4′ and C2 into close proximity. In BFD, as in the other ThDP-dependent enzymes, a bulky hydrophobic residue present at a conserved position (Leu 403) probably helps to constrain the cofactor conformation. Deprotonation of C2 by the 4′ amino group is presumably facilitated by two conserved interactions between the enzyme and the cofactor: (i) proximity of N1′ to a glutamic acid (Glu 47 in BFD), which, using the pyrimidine ring as a proton relay, serves to drive the 4′ amino group to an imino tautomer, and (ii) a hydrogen bond between the resulting imino group and a glycine carbonyl (Gly 401 in BFD), which, among other effects, may serve to position the lone pair of electrons on the imino nitrogen for abstraction of the C2 proton of the cofactor. In the active site of BFD, as in PDC (*7*), the distance between N4′ and C2 is not long enough to accommodate two protons. That is, if C2 is protonated in the resting state of the enzyme, N4′ must exist in the singly protonated imino form, poised to accept the C2 proton.

Residues that, from inspection of the structure of the active site (Figure 8), seem likely to play a role in catalysis by BFD include the His 70/Glu 28 pair, Ser 26, and His 281. These residues might be involved in any of several steps in catalysis following cofactor activation, including donation of a proton to the benzoylformate-ThDP adduct after decarboxylation (*66*) and reprotonation of the cofactor C2 position to complete the reaction cycle (Scheme 2). The pH dependence of V/K for the BFD-catalyzed reaction suggests a requirement for the participation of at least two functional groups in catalysis, one protonated and one

unprotonated (*19*). Efforts to identify the roles of the various active-site residues in catalysis through biochemical and structural studies are currently underway.

Substrate specificity in BFD may be provided by the aromatic residues Phe 464, Trp 463, or Phe 397, which are well suited for binding the phenyl ring of benzoylformate. In fact, the region including the helix that contains Trp 463 and Phe 464 is the most disordered area of the protein, with main-chain *B*-factors between 20 and 46 Å^2. Twenty-one main-chain *B*-factors in this region are above 40 Å^2, while the *B*-factors for all other main-chain atoms in the protein are below 33 Å^2. Perhaps this region becomes more ordered upon binding of substrate; this would be consistent with direct contacts between these residues and substrate.

The relationship of dimers in a tetramer of BFD is much more similar to that in pyruvate oxidase than to that in PDC. Contact between the two dimers is made between analogous elements of secondary structure in pyruvate oxidase and BFD. In contrast, the packing of dimers is radically different in PDC, in which dimers are loosely associated. In fact, the relationship between dimers in a PDC tetramer differs between different crystal forms (*7*, *43*). These differences in quaternary structure between the enzymes seem to be physiologically relevant. In PDC, changes in tetramer conformation are linked to activation (*67*, *68*), whereas in BFD and pyruvate oxidase, no activation and no alterations in tetramer conformation have been observed.

Yeast PDC is activated by substrate and substrate analogues (*69*) through covalent modification of Cys 221 (*14*), a residue in the β-domain which is removed from the active site (*8*, *14*). The side chain of Cys 221 is located in a cavity between the three domains. Activation is mediated in part by a residue in the α-domain, His 92 (*24*, *25*), which is close to Cys 221 and at a position between it and the nearest active site. In contrast, neither BFD nor pyruvate oxidase are known to require an activating factor for maximal activity. Nevertheless, a histidine residue is conserved in both enzymes at the site of His 92 in PDC (His 89 in BFD and His 101 in pyruvate oxidase). In BFD, the imidazole ring of His 89 forms a hydrogen bond with Asp 210, which is homologous with Cys 221 in PDC. In pyruvate oxidase, His 101 forms a hydrogen bonded with the second enzyme-bound cofactor, FAD, which is bound by the β-domain. Thus, it seems that similar structural connections between the active site and the β-domain are present in the three enzymes and are modulated by yeast PDC to effect second-site activation by substrate.

Three classes of metal ions have been identified in the crystals described here: Ca^{2+} bound to the cofactor diphosphate in the active site, Mg^{2+} bound at the dimer interface, and Ca^{2+} bound at a crystal contact. This observation illuminates some aspects of the dynamics of the BFD tetramer. Apparently, a native BFD tetramer has a Mg^{2+} ion bound in each active site, plus two Mg^{2+} ions holding the tetramer together. When BFD is put into the crystallization solution with a high Ca^{2+} concentration (150 mM), the Mg^{2+} ion in the active site is replaced with a Ca^{2+} ion as the cofactor exchanges. In contrast, the tetramers come apart only very rarely, so the Mg^{2+} ions at the dimer interface are not replaced by Ca^{2+} ions. Interestingly, pyruvate oxidase also binds a cation (probably Na^{2+}) at the dimer interface (*10*), although the precise location and the ligand structure are different from BFD.

BFD contains an unusually long and regular example of a polyproline type II (ppII) helix. This element of secondary structure is quite common in globular proteins but it is unusual for the helix to be composed of more than five residues (*70*). Remarkably, the ppII helix in BFD is 13 residues long. The connection between the β- and γ-domains of PDC also forms a ppII helix, while in pyruvate oxidase the corresponding residues form an α-helix. Two other well-known examples of conserved ppII helices of 10−13 residues in length occur in (i) molecules that interact with SH3 domains (*71*, *72*) and (ii) peptides bound to major histocompatibility complex molecules for presentation to T cells (*73*−*76*). In both of these cases, the ppII helix mediates an important protein−protein interaction. In the case of BFD, the ppII helix mediates the main contact between tetramers in the extremely well-ordered, stable crystals described here. In what sort of physiological protein−protein interaction might the ppII helix be involved? In vivo, BFD is evidently part of a multienzyme complex that includes other members of the mandelate pathway (*2*). Perhaps the long, regular ppII helix, exposed on the surface of BFD, is a convenient handle by which BFD in held in the complex.

The results of the comparison of the structure of BFD to those of related enzymes are intriguing, showing several striking patterns of conservation. There is a pronounced similarity in the mode of cofactor activation, but a lack of strict conservation of any residue in the active site that is not directly bound to cofactor. There is, however, a ghostly positional conservation of other active-site residues and catalytic atoms. These structural comparisons suggest that cofactor chemistry, the nature of reaction intermediates, and architectural considerations have been dominant forces in the evolution of this enzyme family. Studies to define the roles of active-site residues in BFD, to allow a deeper appreciation of the implications of the active-site comparisons, are currently in progress.

ACKNOWLEDGMENT

We would like to thank Thomas K. Harris for his thorough review of the manuscript and useful discussions. M.S.H. thanks David A. Sanders and Ronald Kluger for their thoughtful comments.

REFERENCES

1. Stanier, R. Y. (1947) *J. Bacteriol. 54*, 339−348.
2. Halpin, R. A., Hegeman, G. D., and Kenyon, G. L. (1981) *Biochemistry 20*, 1525−1533.
3. Petsko, G. A., Kenyon, G. L., Gerlt, J. A., Ringe, D., and Kozarich, J. W. (1993) *Trends Biochem. Sci. 18*, 372−6.
4. Neidhart, D. C., Howell, P. L., Petsko, G. A., Gerlt, J. A., Kozarich, J. W., Powers, V. M., and Kenyon, G. L. (1990) *Biochem. Soc. Symp. 57*, 135−141.
5. Robinson, B. H., and Chun, K. (1993) *Febs Lett. 328*, 99−102.
6. Tsou, A. Y., Ransom, S. C., Gerlt, J. A., Buechter, D. D., Babbitt, P. C., and Kenyon, G. L. (1990) *Biochemistry 29*, 9856−9862.
7. Arjunan, P., Umland, T., Dyda, F., Swaminathan, S., Furey, W., Sax, M., Farrenkopf, B., Gao, Y., Zhang, D., and Jordan, F. (1996) *J. Mol. Biol. 256*, 590−600.
8. Dyda, F., Furey, W., Swaminathan, S., Sax, M., Farrenkopf, B., and Jordan, F. (1993) *Biochemistry 32*, 6165−6170.

9. Muller, Y. A., and Schulz, G. E. (1993) *Science 259*, 965−967.
10. Muller, Y. A., Schumacher, G., Rudolph, R., and Schulz, G. E. (1994) *J. Mol. Biol. 237*, 315−35.
11. Lindqvist, Y., Schneider, G., Ermler, U., and Sundstrom, M. (1992) *EMBO J. 11*, 2373−2379.
12. Nikkola, M. N., Lindquist, Y., and Schneider, G. (1994) *J. Mol. Biol. 238*, 387−404.
13. Muller, Y. A., Lindqvist, Y., Furey, W., Schulz, G. E., Jordan, F., and Schneider, G. (1993) *Structure 1*, 95−103.
14. Baburina, I., Gao, Y., Hu, Z., Jordan, F., Hohmann, S., and Furey, W. (1994) *Biochemistry 33*, 5630−5635.
15. Bisswanger, H., and Schellenberger, A., Eds. (1996) *Biochemistry and Physiology of Thiamin Diphosphate Enzymes*, A. u. C. Intemann Verlag, Prien, Germany.
16. Kluger, R. (1987) *Chem. Rev. 87*, 863−876.
17. Kluger, R. (1997) *Pure Appl. Chem. 69*, 1957−1967.
18. Schowen, R. L. (1998) in *Comprehensive Biological Catalysis* (Sinnott, M., Ed.) pp 217−266, Academic Press Ltd., San Diego.
19. Weiss, P. M., Garcia, G. A., Kenyon, G. L., Cleland, W. W., and Cook, P. F. (1988) *Biochemistry 27*, 2197−2205.
20. Kern, D., Kern, G., Neef, H., Tittmann, K., Killenbergjabs, M., Wikner, C., Schneider, G., and Hubner, G. (1997) *Science 275*, 67−70.
21. Jordan, F., and Mariam, Y. H. (1978) *J. Am. Chem. Soc. 100*, 2534−2541.
22. Jordan, F., Chen, G., Nishikawa, S., and Wu, B. S. (1982) in *Thiamin: Twenty Years of Progress* (Sabel, H. Z., and Gubler, C. J., Eds.) pp 14−31, New York Academy of Sciences, New York.
23. Schellenberger, A. (1982) *Ann. N. Y. Acad. Sci. 378*, 51−62.
24. Baburina, I., Li, H. J., Bennion, B., Furey, W., and Jordan, F. (1998) *Biochemistry 37*, 1235−1244.
25. Baburina, I., Dikdan, G., Guo, F. S., Tous, G. I., Root, B., and Jordan, F. (1998) *Biochemistry 37*, 1245−1255.
26. Meshalkina, L., Nilsson, U., Wikner, C., Kostikowa, T., and Schneider, G. (1997) *Eur. J. Biochem. 244*, 646−652.
27. Killenbergjabs, M., Konig, S., Eberhardt, I., Hohmann, S., and Hubner, G. (1997) *Biochemistry 36*, 1900−1905.
28. Ibdah, M., Barilan, A., Livnah, O., Schloss, J. V., Barak, Z., and Chipman, D. M. (1996) *Biochemistry 35*, 16282−16291.
29. Wikner, C., Nilsson, U., Meshalkina, L., Udekwu, C., Lindqvist, Y., and Schneider, G. (1997) *Biochemistry 36*, 15643−15649.
30. Wang, J., Martin, P. R., and Singleton, C. K. (1997) *Biochim. Biophys. Acta 1341*, 165−172.
31. Schenk, G., Leeper, F. J., England, R., Nixon, P. F., and Duggleby, R. G. (1997) *Eur. J. Biochem. 248*, 63−71.
32. Nilsson, U., Meshalkina, L., Lindqvist, Y., and Schneider, G. (1997) *J. Biol. Chem. 272*, 1864−1869.
33. Singleton, C. K., Wang, J., Shan, L., and Martin, P. R. (1996) *Biochemistry 35*, 15865−15869.
34. Hasson, M. S., Muscate, A., Henehan, G. T., Guidinger, P. F., Petsko, G. A., Ringe, D., and Kenyon, G. L. (1995) *Protein Sci. 4*, 955−959.
35. Sheldrick, G. M. (1985) in *Crystallographic Computing 3* (Sheldrick, G. M., Krüger, C., and Goddard, R., Eds.) p 175, Oxford University Press, Oxford, England.
36. Wang, B.-C. (1985) *Methods Enzymol. 115*, 90−112.
37. Zhang, K. Y. J. (1993) *Acta Crystallogr., Sect. D 49*, 213−222.
38. Jones, T. A. (1991) *Acta Crystallogr., Sect. A 47*, 110−119.
39. Brünger, A. T. (1992) *XPLOR: A system for X-ray crystallography and NMR*, Yale University Press, New Haven, CT.
40. Laskowski, R. A., MacArthur, M. W., Moss, D. S., and Thornton, J. M. (1993) *J. Appl. Crystallogr. 26*, 283−291.
41. Hodel, A., Kim, S. H., and Brünger, A. T. (1992) *Acta Crystallogr., Sect. A 48*, 851−858.
42. Kraulis, P. J. (1991) *J. Appl. Crystallogr. 24*, 946−950.
43. Lu, G. G., Dobritzsch, D., Konig, S., and Schneider, G. (1997) *FEBS Lett. 403*, 249−253.
44. Shin, W., Pletcher, J., Blank, G., and Sax, M. (1977) *J. Am. Chem. Soc. 99*, 3491−3499.
45. Breslow, R. (1958) *J. Am. Chem. Soc. 80*, 3719−3726.
46. Herzberg, O., and Moult, J. (1991) *Proteins 11*, 223−229.
47. Blake, C. C., Ghosh, M., Harlos, K., Avezoux, A., and Anthony, C. (1994) *Nat. Struct. Biol. 1*, 102−105.
48. Dominguez, R., Souchon, H., Lascombe, M., and Alzari, P. M. (1996) *J. Mol. Biol. 257*, 1042−1051.
49. Fisher, A. J., Thompson, T. B., Thoden, J. B., Baldwin, T. O., and Rayment, I. (1996) *J. Biol. Chem. 271*, 21956−21968.
50. Ghosh, M., Avezoux, A., Anthony, C., Harlos, K., and Blake, C. C. (1994) *EXS 71*, 251−260.
51. Hamelryck, T. W., Poortmans, F., Goossens, A., Angenon, G., Van Montagu, M., Wyns, L., and Loris, R. (1996) *J. Biol. Chem. 271*, 32796−32802.
52. Harris, G. W., Jenkins, J. A., Connerton, I., and Pickersgill, R. W. (1996) *Acta Crystallogr., Sect. D 52*, 393−401.
53. Hazes, B., Magnus, K. A., Bonaventura, C., Bonaventura, J., Dauter, Z., Kalk, K. H., and Hol, W. G. (1993) *Protein Sci. 2*, 597−619.
54. Hennig, M., Jansonius, J. N., Terwisscha van Scheltinga, A. C., Dijkstra, B. W., and Schlesier, B. (1995) *J. Mol. Biol. 254*, 237−246.
55. Hohenester, E., Maurer, P., Hohenadl, C., Timpl, R., Jansonius, J. N., and Engel, J. (1996) *Nat. Struct. Biol. 3*, 67−73.
56. Koellner, G., Luic, M., Shugar, D., Saenger, W., and Bzowska, A. (1997) *J. Mol. Biol. 265*, 202−216.
57. Lang, D., Hofmann, B., Haalck, L., Hecht, H. J., Spener, F., Schmid, R. D., and Schomburg, D. (1996) *J. Mol. Biol. 259*, 704−717.
58. Leesong, M., Henderson, B. S., Gillig, J. R., Schwab, J. M., and Smith, J. L. (1996) *Structure 4*, 253−264.
59. Vanscheltinga, A., Hennig, M., and Dijkstra, B. W. (1996) *J. Mol. Biol. 262*, 243−257.
60. Vos, S., Dejersey, J., and Martin, J. L. (1997) *Biochemistry 36*, 4125−4134.
61. Zajc, A., Romao, M. J., Turk, B., and Huber, R. (1996) *J. Mol. Biol. 263*, 269−283.
62. Kluger, R., Lam, J. F., and Kim, C.-S. (1993) *Bioorgan. Chem. 21*, 275−283.
63. Kluger, R., Lam, J. F., Pezacki, J. P., and Yang, C. M. (1995) *J. Am. Chem. Soc. 117*, 11383−11389.
64. Babbitt, P. C., Hasson, M. S., Wedekind, J. E., Palmer, D. R., Barrett, W. C., Reed, G. H., Rayment, I., Ringe, D., Kenyon, G. L., and Gerlt, J. A. (1996) *Biochemistry 35*, 16489−16501.
65. Martin, J. L. (1995) *Structure 3*, 245−250.
66. Barletta, G. L., Zou, Y., Huskey, W. P., and Jordan, F. (1997) *J. Am. Chem. Soc. 119*, 2356−2362.
67. Konig, S., Svergun, D., Koch, M. H., Hubner, G., and Schellenberger, A. (1993) *Eur. Biophys. J. 22*, 185−194.
68. Hübner, G., König, S., Schellenberger, A., and Koch, M. H. (1990) *FEBS Lett. 266*, 17−20.
69. Hübner, G., Weidhase, R., and Schellenberger, A. (1978) *Eur. J. Biochem. 92*, 175−181.
70. Adzhubei, A. A., and Sternberg, M. J. (1994) *Protein Sci. 3*, 2395−2410.
71. Cohen, G. B., Ren, R., and Baltimore, D. (1995) *Cell 80*, 237−248.
72. Pawson, T. (1995) *Nature 373*, 573−580.
73. Jardetzky, T. S., Brown, J. H., Gorga, J. C., Stern, L. J., Urban, R. G., Strominger, J. L., and Wiley, D. C. (1996) *Proc. Natl. Acad. Sci. U.S.A. 93*, 734−738.
74. Stern, L. J., Brown, J. H., Jardetzky, T. S., Gorga, J. C., Urban, R. G., Strominger, J. L., and Wiley, D. C. (1994) *Nature 368*, 215−221.
75. Madden, D. R., Gorga, J. C., Strominger, J. L., and Wiley, D. C. (1992) *Cell 70*, 1035−1048.
76. Stern, L. J., and Wiley, D. C. (1994) *Structure 2*, 245−251.

BI973047E