Structure, Vol. 11, 985–996, August, 2003, ©2003 Elsevier Science Ltd. All rights reserved. DOI 10.1016/S0969-2126(03)00158-8

Structure of *E. coli* Ketopantoate Hydroxymethyl Transferase Complexed with Ketopantoate and Mg^{2+}, Solved by Locating 160 Selenomethionine Sites

Frank von Delft,[1,4] Tsuyoshi Inoue,[1,5] S. Adrian Saldanha,[2] Harald H. Ottenhof,[3] Florian Schmitzberger,[1] Louise M. Birch,[2] Venugopal Dhanaraj,[1] Michael Witty,[3] Alison G. Smith,[3] Tom L. Blundell,[1] and Chris Abell[2,*]

[1]Department of Biochemistry
80 Tennis Court Road
Cambridge CB2 1GA
United Kingdom

[2]University Chemical Laboratory
Lensfield Road
Cambridge CB2 1EW
United Kingdom

[3]Department of Plant Sciences
Downing Street
Cambridge CB2 3EA
United Kingdom

Summary

We report the crystal structure of *E. coli* ketopantoate hydroxymethyltransferase (KPHMT) at 1.9 Å resolution, in complex with its product, ketopantoate. KPHMT catalyzes the first step in the biosynthesis of pantothenate (vitamin B_5), the precursor of coenzyme A and the acyl carrier protein cofactor. The structure of the decameric enzyme was solved by multiwavelength anomalous dispersion to locate 160 selenomethionine sites and phase 560 kDa of protein, making it the largest structure solved by this approach. KPHMT adopts the $(\beta\alpha)_8$ barrel fold and is a member of the phosphoenolpyruvate/pyruvate superfamily. The active site contains a ketopantoate bidentately coordinated to Mg^{2+}. Similar binding is likely for the substrate, α-ketoisovalerate, orienting the C3 for deprotonation.

Introduction

Pantothenate (vitamin B_5) is the precursor for the synthesis of the 4′-phosphopantetheine moiety found in coenzyme A, a molecule central to energy metabolism in all organisms. Additionally, the phosphopantetheine group is found in acyl carrier protein involved in fatty acid synthesis and as an essential prosthetic group for enzymes involved in nonribosomal peptide biosynthesis (Kleinkauf, 2000). Whereas most primary producers, including microorganisms and plants, can synthesize pantothenate de novo, animals require vitamin B_5 as an essential dietary nutrient. This makes the enzymes of pantothenate biosynthesis attractive targets for novel herbicides and antimicrobials.

The first committed step in the biosynthesis of pantothenate is the formation of ketopantoate (**5**) from α-ketoisovalerate (**1**; α-KIVA), catalyzed by ketopantoate hydroxymethyltransferase (KPHMT; 5,10-methylene-5,6,7,8-tetrahydrofolate: α-ketoisovalerate hydroxymethyltransferase; EC 2.1.2.11). Powers, Snell, and coworkers purified the enzyme and showed that KPHMT is a class II aldolase (i.e., metal requiring) that utilizes 5,10-methylene-5,6,7,8-tetrahydrofolate (**2**; methylene-THF) to transfer a hydroxymethyl group to α-KIVA, as shown in Figure 1 (Teller et al., 1976; Powers and Snell, 1976). The purified enzyme resisted denaturation by 4 M urea or heating up to 80°C (Powers and Snell, 1976). It was demonstrated to be a decamer by sedimentation equilibrium experiments, gel filtration chromatography, and polyacrylamide gel electrophoresis under native conditions (Powers and Snell, 1976).

The *panB* gene encoding KPHMT was cloned from *E. coli* by functional complementation of a *panB* mutant (Jones et al., 1993). Subsequently, a region of the *E. coli* K12 genome encoding the *panB*, *panC*, and *panD* genes was identified by transposon mutagenesis (Merkel and Nichols, 1996). The *panB* gene in this case encoded a protein that differed from the first by six amino acids. The protein encoded by the K12 *panB* gene (Swiss-Prot ID P31057) has 264 amino acids, corresponding to a molecular weight of 28.2 kDa. More recently, the *panB* gene from the filamentous fungus *Aspergillus nidulans* was isolated, again by functional complementation of the corresponding mutant (Kurtov et al., 1999). KPHMT from this organism is 58% similar at the amino acid level to the *E. coli* K12 enzyme but comprises 349 amino acids (37.7 kDa), the extra residues being located mainly as extensions at both the N and C termini. The recombinant protein expressed in *E. coli* behaved as an octamer on gel filtration (Kurtov et al., 1999).

Little is known about the mode of action of KPHMT, except that the addition of the hydroxymethyl group proceeds with retention of configuration (Aberhart and Russell, 1984). Additionally, Mg^{2+} has been shown to be an activator of the enzyme (Powers and Snell, 1976). In the absence of added Mg^{2+}, the activity of the enzyme was reduced by greater than 10-fold. Mg^{2+} can be replaced by Mn^{2+}, Co^{2+}, Ni^{2+}, or Zn^{2+}, but this results in progressively lower activity.

Here we describe the X-ray crystal structure of product-bound *E. coli* KPHMT. The structure was successfully phased by selenomethionine multiwavelength anomalous dispersion (SeMet MAD). This required the 160-atom selenium substructure to be solved by Shake-and-Bake (SnB) direct methods (Weeks and Miller, 1999). The structure adopts a $(\beta\alpha)_8$ (triose-phosphate isomerase, or TIM) barrel fold and is shown to be a member of the phosphoenolpyruvate/pyruvate superfamily. The bound product delineates the active site, and substrate binding and activation are discussed.

*Correspondence: ca26@cam.ac.uk
[4]Present address: Scripps Research Institute, 10550 North Torrey Pines Road, La Jolla, California 92037.
[5]Present address: Department of Materials Chemistry, Graduate School of Engineering, Osaka University, Yamada-Oka 2-1, Suita, Osaka 565-871, Japan.

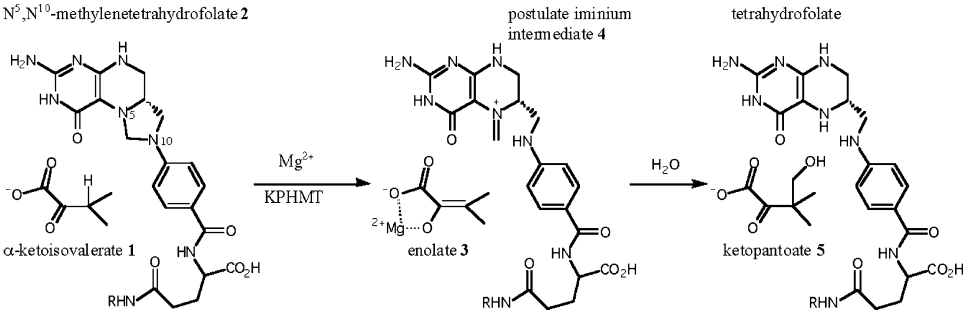


Figure 1. Proposed Mechanism for the Conversion of α-KIVA and Methylene-THF into Ketopantoate and THF Catalyzed by KPHMT

Results and Discussion

Crystallization of Recombinant KPHMT

The *panB* gene was overexpressed in the *panB*-deficient *E. coli* strain YA139 (Cronan, 1980) to yield recombinant native KPHMT, which was purified in a three-step chromatographic procedure. Similarly, SeMet-substituted KPHMT was prepared after growth of *E. coli* in the presence of SeMet (van Duyne et al., 1993). The recombinant proteins both had a subunit molecular weight of 28 kDa on SDS-PAGE, close to that predicted from the gene sequence, and were judged by analytical ultracentrifugation to be decameric in solution (data not shown), as reported for the nonrecombinant protein (Powers and Snell, 1976). The extent of SeMet incorporation into the protein was determined by electrospray mass spectrometry. The native and SeMet proteins had observed masses of 28,234.7 ± 16.3 and 28,663.2 ± 11.6, respectively (predicted masses, 28,237 and 28,660). The mass difference, 429 ± 20.0, corresponds within error to that expected (423) if all nine methionine residues were replaced with SeMet.

Data Collection

Data were collected on two different crystal forms, both of which belonged to space group $P2_1$, with related unit cell dimensions (Table 1), form A having a c axis twice as long as that of form B. A two-wavelength MAD experiment with a SeMet crystal of form A yielded solvent-flattened 3 Å electron density maps that confirmed the presence of two decamers in the asymmetric unit (ASU). These allowed the subunit backbone to be traced, even before applying 20-fold noncrystallographic symmetry (NCS) averaging. A subunit was built into averaged maps and used to reconstruct a decamer, which was positioned in the crystal form B by molecular replacement with data from a form B native crystal and refined to a final resolution of 1.8 Å (effective resolution of 1.9 Å).

The quality of the final, deposited model is summarized in Table 1. All expected residues are present, except the N-terminal two residues of each chain. Residues 138–160 and 220–240 have high B factors in all subunits and, in subunits E and J, could only be built by comparison with NCS-related subunits. There are 42 residues with side chains in multiple conformations, and 30 have unstructured side chains that were omitted. Bound product (ketopantoate) and Mg^{2+} are present.

The NCS agreements between the ten subunits of the decamer are summarized in Table 2. NCS restraints were eventually relaxed during refinement, improving R_{free} significantly. In addition, NCS is broken in two residue stretches that adopt distinct conformations. Differing crystal contacts around residues 218–228 cause three conformations. Residues 235–239 lie on either side of the noncrystallographic 2-fold, and strict symmetry would cause steric clashes, which are prevented by adopting complementary conformations.

There are large differences in disorder between the five dimers of the decamer, as shown in Table 2. As a result, the assignment of subunits to separate TLS groups led to an >4% drop in R_{free} and to average B factors that are virtually identical in all subunits, as is reasonable for a homomultimer.

SeMet Structure Solution

The SeMet MAD experiment is noteworthy, in that the 160-selenium atom substructure is, to our knowledge, the largest solved to date without the aid of prior phase information, while the 560 kDa ASU is the largest phased by SeMet MAD only. An emphasis on high observational redundancy (specifically at the "peak" wavelength; Table 1) and reliable measurement of low-resolution reflections (only first order and some second order reflections remained unobserved) were sufficient for SnB (Weeks and Miller, 1999) to locate at least 120 sites correctly by direct methods, allowing the remainder to be located by SHARP (de la Fortelle and Bricogne, 1997). Because of the size of both the unit cell and the substructure involved, substantial computational time was required. More than 2000 hr of extensive parameter exploration in SnB (Weeks and Miller, 1999) preceded the first correct substructure. For substructure refinement in SHARP, only B factors and nonisomorphism parameters were refined in order to keep computations to within acceptable time limits

The Structure of KPHMT

E.coli KPHMT is a homodecamer with dimensions of 100 × 105 × 80 Å, made up of roughly spherical subunits measuring 50 × 50 × 40 Å, arranged in 522-point symmetry (Figures 2A and 2B). It appears to be a pentamer of dimers, rather than the converse, with the dimer being the functional unit. With SURFACE in CCP4 (CCP4, 1994), the dimer interface was found to bury more surface than the pentamer interface (1140 versus 760 $Å^2$) and involve four times as many interactions: 88 total (20

Table 1. Crystallographic Data Quality, Phasing, Refinement, and Model Quality

Lattice	SeMet Large		Native Small
Space group	$P2_1$		$P2_1$
Cell parameters (rmsd)			
a (Å)	87.8 (0.05)		86.1 (0.05)
b (Å)	155.6 (0.05)		157.2 (0.05)
c (Å)	209.9 (0.05)		100.2 (0.05)
β (°)	99.4° (0.1)		97.4 (0.1)
$V_{unit\ cell}$ ($Å^3$)	2.79×10^6		1.34×10^6
Data Quality			
Data set	Peak	Remote	Native
Wavelength (Å)	0.97927	0.9393	0.979
Limiting resolution (Å)	2.8	3.0	1.8
R_{meas}[a] (high resolution[b])	0.120 (0.60)	0.113 (0.60)	0.093 (0.62)
$< I/\sigma I >$ (high resolution[b])	25.6 (6.0)	12.6 (2.1)	12.9 (2.4)
Completeness (high resolution[b])	0.99 (0.99)	0.99 (0.99)	0.94 (0.75)
Number of unique reflections (multiplicity[c])	136,609 (21.2)	111,181 (6.6)	229,086 (4.5)
Number of reflections unobserved below 12 Å	8/1,808 possible	40/1,808	13/866
Wilson B factor ($Å^2$)	63.6	72.4	20.7
Experimental f′/f″ (electrons)[d]	−8.6/5.4	−1.3/3.2	—
Phasing			
Phasing power[e]			
Acentrics (iso/ano)	—/3.04	0.59/1.77	
Centrics	—	0.43	
R_{cullis}[f] for acentrics (iso/ano)	—/0.54	0.93/0.85	
Figure of Merit[g] (centrics/acentrics)	0.32/0.56		
Refinement (75–1.8 Å)			
R_{cryst}[h]	0.158 (high resolution[b], 0.22)		
R_{free}[i]	0.196 (high resolution[b], 0.25)		
Number of reflections in working/test[i] set	224,262/4,812 (in thin resolution shells)		
Number of restraints	316,553		
Number of parameters	90,440		
Model Contents and Quality			
Residues	2,620		
Hetero groups	10 ketopantoate, 10 Mg^{2+}		
Nonhydrogen atoms	22,629		
Solvent molecules	2,563		
Average B factor[j] ($Å^2$)			
Subunits A–J	22.1, 23.8, 24.9, 27.6, 39.4, 25.6, 19.6, 25.9, 26.4, 43.8		
Solvent	34.2		
Ramachandran plot[k]			
Percentage of residues in favourable regions	98.8		
Percentage of residues in allowed regions	100		
Rms deviations (percentage of outliers[l])			
Bond lengths (Å)	0.017 (0.2)		
Bond angles (°)	1.56 (0.02)		
Planarity (Å)	0.008 (0)		

[a] $R_{meas} = [\Sigma_h w\Sigma_i|\langle I_h\rangle - I_{h,i}|]/\Sigma_h\Sigma_i I_{h,i}$, where $w = [n_h/(n_h - 1)]^{1/2}$ and $\langle I_h\rangle = [\Sigma_i{}^n I_{h,i}]/n_h$. This is the multiplicity-weighted R_{symm} (Diederichs and Karplus, 1997).

[b] Highest-resolution shell: SeMet large, 2.95–2.80 Å; native small, 1.9–1.8 Å.

[c] Multiplicity for Friedel mates considered equivalent.

[d] Estimates from fluorescence scans.

[e] Phasing power = $\langle|F_H|/\epsilon\rangle$; $\epsilon = |F_{PH} - |F_P + F_H||$, the residual lack of closure (LOC); F_H is the calculated heavy-atom structure factor amplitude; F_{PH} and F_P are the trial structure factors with and without heavy atoms. Numbers are as output by SHARP (de la Fortelle and Bricogne, 1997).

[f] $R_{cullis} = \langle LOC\rangle/\langle|F_{PH} - F_P|\rangle$.

[g] FOM = $\langle\cos\Delta\alpha_j\rangle$; $\Delta\alpha_j$ is the phase angle error for phase angle j.

[h] $R_{cryst} = \Sigma||F_{obs}| - |F_{calc}||/\Sigma|F_{obs}|$; F_{obs} and F_{calc} are the observed and calculated structure factor amplitudes.

[i] R_{free} is the crossvalidation of R_{cryst}, i.e., calculated using randomly selected test data not used in refinement.

[j] Average B factor includes contribution from TLS definition.

[k] As defined in Molprobity (http://kinemage.biochem.duke.edu/molprobity/index.html).

[l] Percentage parameters flagged as outliers by PROCHECK (Laskowski et al., 1993).

Table 2. NCS Agreement between the Ten Subunits

Subunit	Positional Rmsd (Å)	Rmsd $B_{residual}$ (Å^2)	Rmsd B_{total} (Å^2)	Average $B_{residual}$ (Å^2)	Average B_{total} (Å^2)
A	—	—	—	12.6	22.1
B	0.43	3.86	7.4	13.2	23.8
C	0.17	3.01	6.2	13.3	24.9
D	0.19	3.42	8.0	13.7	27.6
E	0.32	4.28	20.2	14.9	39.4
F	0.17	2.48	6.8	13.2	25.9
G	0.30	4.38	5.6	13.0	19.6
H	0.21	2.39	6.1	13.1	25.9
I	0.22	2.71	6.1	13.1	26.4
J	0.44	6.05	24.7	15.9	34.2

Rmsd, root-mean-square difference with arbitrary reference to subunit A.
Positional rmsd's are for all main chain atoms and exclude residues that do not obey NCS (218–228 and 235–239; see text).
$B_{residual}$, residual temperature factors excluding contribution from TLS parameters.
B_{total}, sum of residual temperature factors and contribution from TLS (as generated by TLSANL).

hydrophobic, 68 polar) versus 20 (14 hydrophobic, 6 polar).

The decamer as a whole is tightly packed, since the combined surface area of all subunits is 108,000 Å^2, whereas only 83,200 Å^2 are accessible, leaving nearly a quarter buried in subunit interfaces. The KPHMT subunit

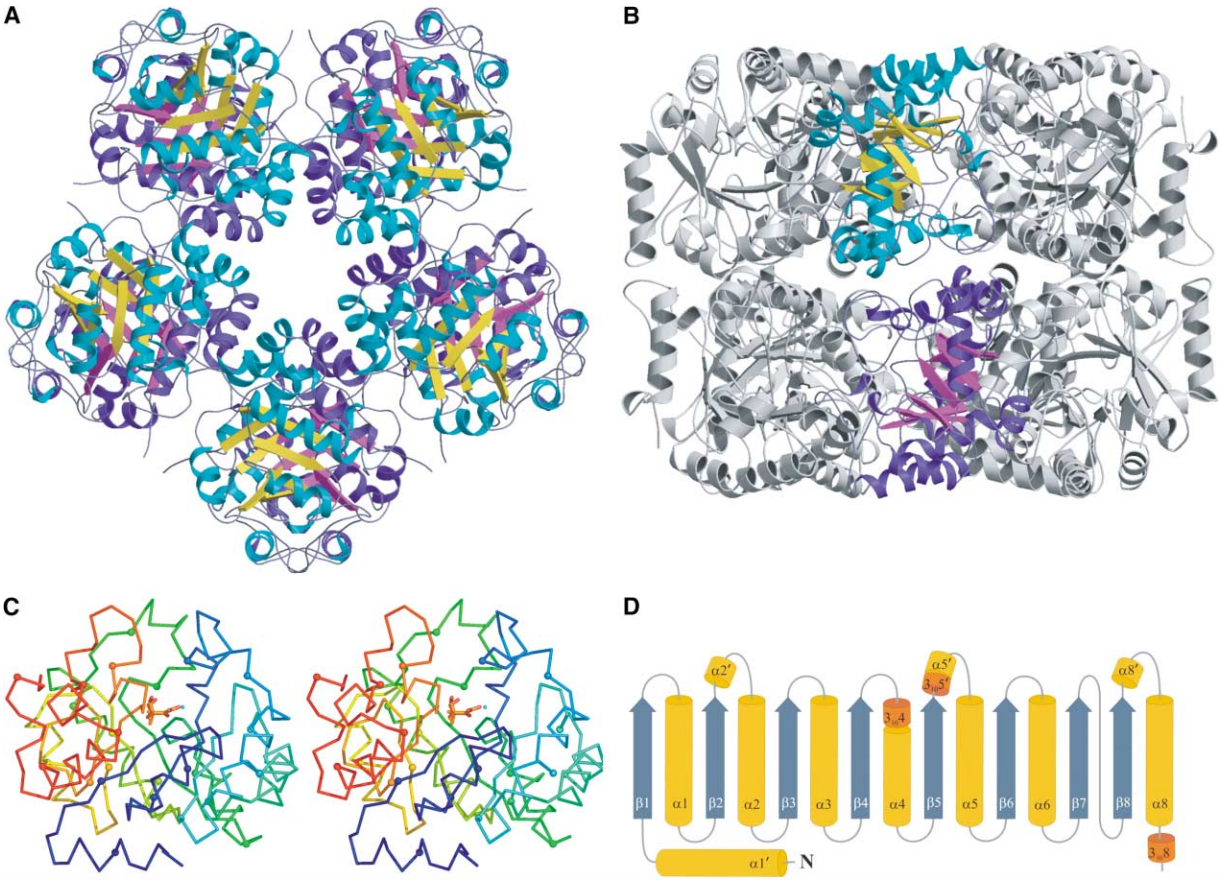


Figure 2. Overall Structure of Decameric KPHMT

(A) View down the noncrystallographic 5-fold axis.

(B) View from the side of the noncrystallographic 2-fold axis. One of the pentamers is colored in cyan for the helices and in yellow for the strands. The other pentamer is shown in blue and purple for helices and strands, respectively.

(C) Stereo Cα trace of the KPHMT protomer, including the product, ketopantoate. The structure is colored by position in the sequence, and every tenth residue in the structure is highlighted by a ball.

(D) Schematic diagram of the secondary structure elements of the KPHMT protomer. β strands, gray-blue; α helices, yellow; 3_{10} helices, orange.

Figures (A)–(C) were prepared with MOLSCRIPT (Kraulis, 1991) and rendered in RASTER3D (Merritt and Murphy, 1994).

adopts the classic $(\beta\alpha)_8$ (TIM) barrel fold: a barrel of parallel β strands that alternate with α helices, although the usual helix between strands 7 and 8 is here replaced by a loop (Figures 2C and 2D). An additional N-terminal α helix is located at the base (N-terminal end) of the β barrel. The axis of the β barrel lies in the plane perpendicular to the 5-fold axis, the base forming an interaction with the adjacent subunit of the pentamer (Figure 2A).

KPHMT Sequence Alignment

At least 48 protein or translated gene sequences with sufficient similarity to be classified as KPHMT orthologs can be identified in the published databases to date, with examples from bacteria, fungi, and plants, but not from animals. ClustalW analysis of the sequences has enabled the identification of conserved residues. The correlation between primary structure among bacterial and fungal KPHMTs and the secondary structure of the *E. coli* enzyme is shown in Figure 3. From the consensus sequence it is clear that, of the 264 residues in the *E. coli* protein, 99 residues are conserved (20 invariant and 79 conserved). Eight regions (Figure 3, boxed) can be identified in which there are four or more contiguous conserved residues. Five of these stretches contain residues involved in binding the product or metal ion (see below). The eukaryotic sequences are longer than those from bacteria, and the presence of an extension at the N terminus suggests that this might act as a targeting signal to direct import into mitochondria.

Ketopantoate auxotrophs have been reported from *E. coli* (Cronan, 1980; Cronan et al., 1982), *Salmonella typhimurium* (Cronan et al., 1982), *A. nidulans* (Kurtov et al., 1999), and the higher plant *Datura innoxia* (thornapple) (Sahi et al., 1988). These mutants can grow if supplemented with ketopantoate or pantothenate, but not α-KIVA, and the bacterial (Cronan, 1980; Cronan et al., 1982) and *A. nidulans* (Kurtov et al., 1999) mutants have been shown to be deficient in KPHMT activity. The *panB100* auxotroph from *A. nidulans* has been fully characterized and found to be a deletion of Gly268 (Kurtov et al., 1999), corresponding to conserved Gly205 in *E. coli* KPHMT (see Figure 3), the last residue of one of the conserved regions mentioned above. This residue is invariant in 43 out of the 48 KPHMT sequences and a serine in the other 5. The residue lies in a tight loop connecting strands β7 and β8, replacing helix α7 of the classical TIM barrel. It has a positive ϕ torsion angle and is inaccessible to solvent, an environment where it is difficult to make a deletion. The nature of the mutations in the three bacterial *panB* mutants, *E. coli* YA139 (Cronan, 1980), and *S. typhimurium* pan-4 and pan-6 (Cronan et al., 1982) were investigated. The genes were amplified by PCR from genomic DNA extracted from the mutants and sequenced. The results of two independent amplifications found single-base changes in each of the *panB* genes, all of which resulted in amino acid changes in the protein (see Figure 3). In the *E. coli* mutant, the invariant Val72 was altered to Glu, which would disrupt the hydrophobic core between strand β2 and helix α2. The *S. typimurium* pan-4 mutation is Gly200Ser (Gly201 in *E. coli*). This mutation would be expected to disrupt the same β7-loop-β8 motif affected by the *A. nidulans* *panB100* mutation (Kurtov et al., 1999), but, in this case, by the inability of the core of the motif to accommodate the bigger side chain. The pan-6 mutation, Thr92Pro, would disrupt the initiation of helix α3: the Thr92 side chain initiates the helix through its hydrogen bond with the N-terminal backbone amide group of the helix; though not conserved, this residue is H bonding (or glycine) in all sequences.

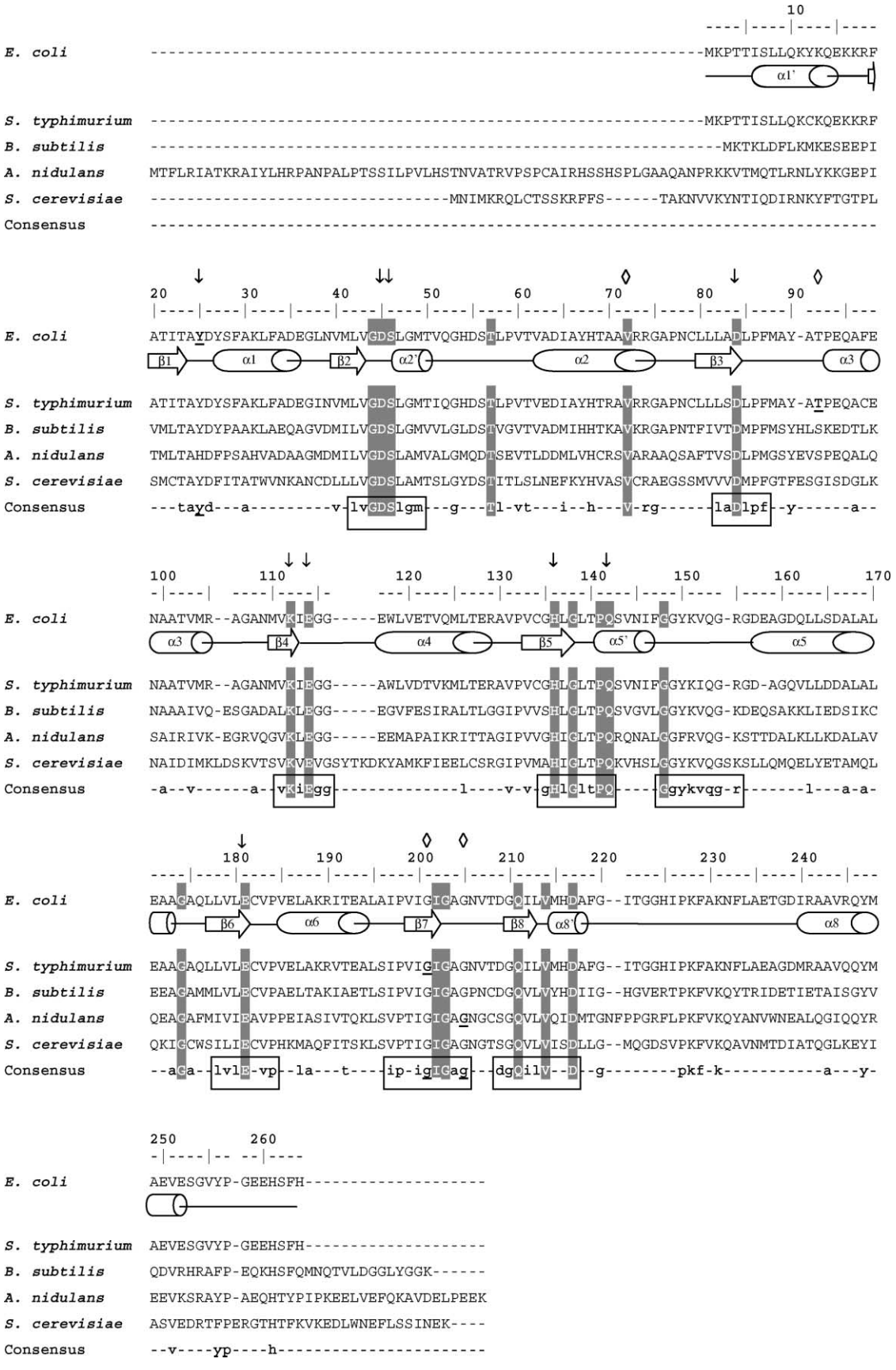
The Substrate Binding Site

The active site, as indicated by the bound product and as is usual for $(\beta\alpha)_8$ enzymes, lies in a deep pocket formed between the C termini of the β strands. The funnel-shaped pocket (Figure 4A) extends more than a third into the protein, about 15 Å deep, but has a wider opening (20 $\times$ 10 Å). Nonprotein electron density (Figure 4B) at the bottom of the pocket was interpreted to be cocrystallized product (ketopantoate) bidentately coordinated to a Mg^{2+} ion, the preferred divalent cation that is required for activity (Powers and Snell, 1976); binding residues are strongly conserved (Figures 4C and 3).

The putative Mg^{2+} has difference density peak heights similar to those of surrounding, ordered solvent molecules, which are isoelectronic to Mg^{2+}. The site also has octahedral coordination, and observed coordination distances are typical for Mg^{2+} (Figure 4C). Three side chain carboxylate groups, Asp45, Asp84, and Glu114, occupy axial, equatorial, and equatorial positions, respectively, the latter indirectly via a fully buried water molecule. The keto and carboxyl groups of ketopantoate coordinate axially and equatorially, respectively, and a second solvent molecule binds equatorially.

Unexpectedly, the carboxyl and carbonyl C=O bonds are not coplanar; instead, there is a torsion of 30° around C1-C2 (Figures 4B and 4C), implying poor π conjugation between the groups. Apart from coordinating Mg^{2+}, the carboxyl-carbonyl groups of the product are bound by H bonds to S46, the main chain amide of S46, and the amino group of K112. In contrast, the tertiary C3 dimethyl-hydroxymethyl group lies in a fully conserved, hydrophobic pocket, accommodating the two methyl groups. The hydroxymethyl group therefore points toward the mouth of the active site and occupies alternate conformations: one H bonded to conserved E181 and the other H bonded to two water molecules.

The substrate α-KIVA **1** and product **5** are structurally very similar, and it is likely therefore that they bind in a similar manner. The carbonyl and carboxyl groups of α-KIVA would coordinate to Mg^{2+}, and the binding of the two methyls in the hydrophobic pocket would present the hydrogen on C3 toward Glu181. The acidity of C3 of α-KIVA would be enhanced by coordination of the keto group to the magnesium ion. Glu181 is well placed to act as the base for the abstraction of the C3 proton from α-KIVA to form enolate **3**. Its basicity is enhanced by a network of hydrogen bonds connecting it with His136 and Lys112, constituting an invariant triad. Glu181 is then available to protonate N10 of methylene-THF, which is expected to bind above the substrate, facilitating the formation of an iminium intermediate **4** on THF after cleavage of the C11-N10 bond. The reaction is then assumed to proceed by attack of enolate **3** on


E. coli
S. typhimurium
B. subtilis
A. nidulans
S. cerevisiae
Consensus
α1'
β1
α1
β2
α2'
α2
β3
α3
β4
α4
β5
α5'
α5
β6
α6
β7
β8
α8'
α8

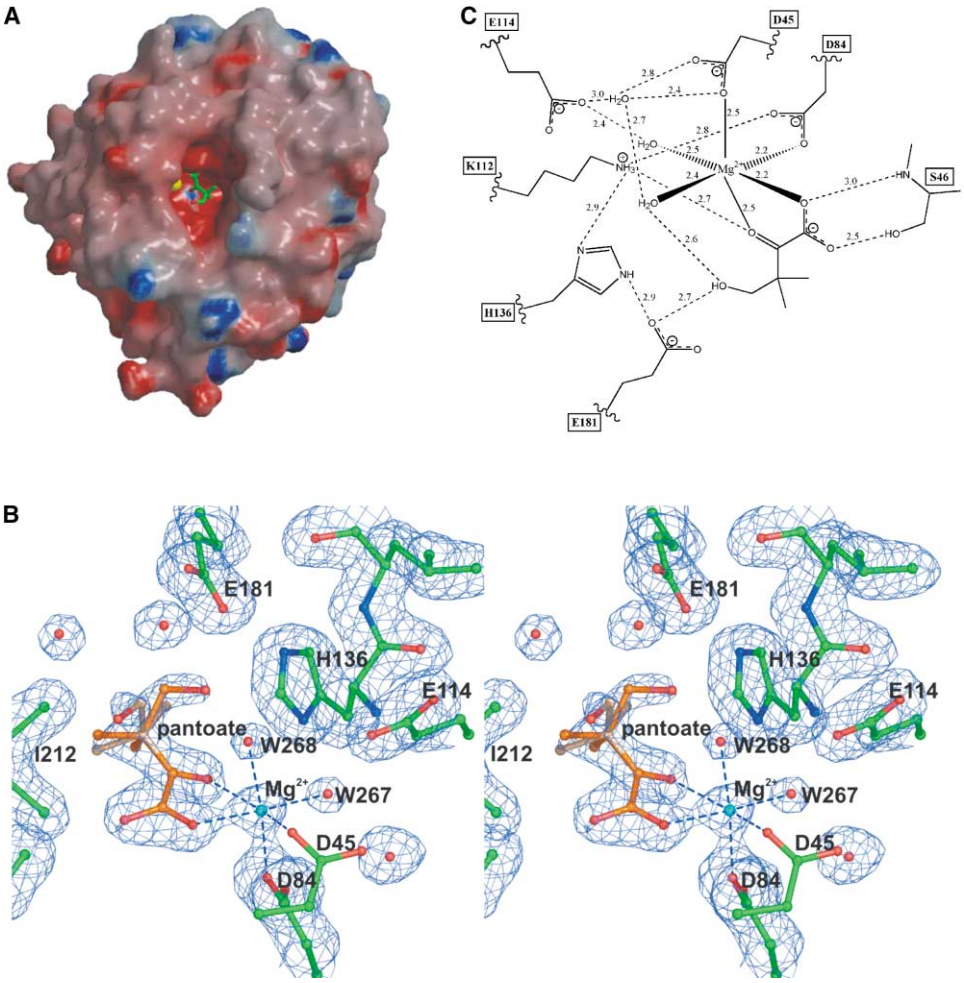


Figure 4. Active Site of *E. Coli* KPHMT

(A) View down into the metal and substrate binding site in a protomer of KPHMT. The electrostatic potential mapping on the molecular surface was calculated with GRASP (Nicholls et al., 1991). Blue, electropositive potential ($>$15 kT); red, electronegative potential ($<-$15 kT).

(B) σ_A-weighted $2F_o - 2F_c$ omit electron density maps (contoured at 1.2 σ) of residues in the active site, showing a bound ketopantoate and Mg^{2+}.

(C) A schematic drawing of the active site structure, showing the hydrogen bond networks around ketopantoate and Mg^{2+}. The top of the figure corresponds to the top of the $(\beta\alpha)_8$ barrel. All the residues shown are completely conserved. Distances shown are in angstroms.

the iminium species. A study of the enolisation step in KPHMT with a range of keto acids has recently been reported (Sugantino et al., 2003).

Comparison with Other $(\beta\alpha)_8$ Barrel Enzymes

A search of protein folds with the complete coordinates of the KPHMT protomer and the protein coordinates in the Protein Data Bank is complicated by the fact that a large number of superfamilies adopt the $(\beta\alpha)_8$ fold. The program DALI (http://www2.ebi.ac.uk/dali/) (Holm and Sander, 1993) indicated closest structural similarity of KPHMT with phosphoenolpyruvate mutase (PEPM; Protein Data Bank code 1pym; Huang et al., 1999) and isocitrate lyase (ICL; 1dqu and 1igw; Britton et al. 2001) (Table 3). Both of these enzymes were also identified as close homologs by a KPHMT sequence to structure search with FUGUE, a homology recognition server (Shi et al., 2001). PEPM from *Mytilus edulis* is a homotetra-

Figure 3. Alignment of the Amino Acid Sequences of KPHMT from Several Organisms

The predicted amino acid sequences of 48 KPHMTs present in the databases, including those from bacteria, archaea, fungi, and plants, were aligned with each other with ClustalW, to generate a consensus of conserved residues. The figure shows this consensus, together with the sequences of the *E. coli* KPHMT (Merkel and Nichols, 1996) (P31057, BAB33561), *S. typhimurium* (Q8ZRR0), *B. subtilis* (P52996), *A. nidulans* (Q9Y7B6) (Kurtov et al., 1999), and *Saccharomyces cerevisiae* (P38122). In the consensus sequence residues in uppercase are invariant, while those in lower case are conserved in at least 25 out of the 48 sequences. Residues marked with a diamond show the position of mutations in the four ketopantoate auxotrophs. Residues marked with a downward-pointing arrow interact with Mg^{2+} or ketopantoate in the crystal structure.

Table 3. Comparison of Sequence Identity, Z Score, and Number of Aligned Residues, as Found in DALI, and the Corresponding Rmsd for Aligned Residues of the PEPM, ICL, and DDGA, Assigned to the Phosphoenolpyruvate/Pyruvate Superfamily in SCOP

Enzyme	PID (%)	Z Score DALI	Number of Aligned Residues	RMSD (Å)	Z Score FUGUE
PEPM	13.5	17.1	198	3.2	14.47
ICL	16.5	14.6	201	3.8	3.35
DDGA	11.4	12.3	173	3.0	—

All comparisons are made with respect to KPHMT. Sequence identities were calculated from structure-based COMPARER alignments. Additionally, the Z scores for PEPM and ICL, which could both be identified with FUGUE, are indicated. For the comparison with ICL, the *E.coli* structure was used. From this comparison it is clear that the closest structural homolog of KPHMT is PEPM. PEPM is also the closest structural homologue to ICL and vice versa, as identified with DALI.

meric enzyme that catalyses the conversion of phosphoenolpyruvate to phosphonopyruvate, via cleavage of an O-P bond and formation of a C-P bond (Huang et al., 1999). ICL, which, like PEPM, is a homotetramer, catalyzes the first step in the glyoxylate-bypass pathway, the conversion of isocitrate into glyoxylate and succinate. The first part of this mechanism involves a deprotonation, as does the KPHMT reaction (Britton et al., 2001). The SCOP database (Murzin et al., 1995) currently distinguishes 24 superfamilies of $(\beta\alpha)_8$ barrel proteins where the members of the superfamily have a probable common evolutionary origin. Both ICL and PEPM are assigned to the phosphoenolpyruvate/pyruvate superfamily. 2-dehydro-2-deoxy-galactarate aldolase (DDGA) is a class 2 aldolase with a single domain $(\beta\alpha)_8$ fold (Izard and Blackwell, 2000). After ICL and PEPM, DDGA shows the highest structural similarity to KPHMT within the superfamily.

Structure-based sequence alignment between all four enzymes is shown in Figure 5A. All of the structures have the N-terminal α helix that blocks the end of the barrel where the N termini of the strands are found. The insertion in the loop region between β4 and α4 in PEPM and ICL is thought to be involved in conformational change upon substrate binding and blocks the other end of the barrel where the active site is found. There is no similar functional region in KPHMT or DDGA. All the enzymes except KPHMT have C-terminal helices that protrude from one protomer to pack against a neighboring protomer, a phenomenon known as helix swapping. Structural conservation between KPHMT, ICL, and PEPM is most obvious in β3, α5, and the loop connecting α5 and β6.

There are striking resemblances in the coordination spheres of the magnesium in the four enzymes. PEPM and ICL (Figure 5B) are the most similar to KPHMT (Figure 4C). Asp84 in β3 and Glu114 at the C terminus of β4 of KPHMT correspond to the invariant residues Asp85 in β3 and Glu114 at the C terminus of β4 in PEPM. The carboxyl groups of the oxalate residue in PEPM coordinate to axial and equatorial positions of Mg^{2+}, as do the carboxyl and keto groups of ketopantoate. In both proteins there is also a hydrogen bond from the equatorially liganded carboxyl group to Ser46. There are also some differences. Asp45 in KPHMT is replaced in PEPM by a water molecule, which is held in place by hydrogen bonding to Asp58 and Asp87. Asp58 also hydrogen bonds to a second Mg^{2+}-coordinated water, which is also hydrogen-bonded to Lys120, a residue for which there is no direct equivalent in KPHMT. In contrast, the absolute position of the metal ion-coordinating residues is different in DDGA, such that the metal ion has moved from being between β3 and β4 to a position between β5 and β6, which also inverts the orientation of the α-ketoacid substrate. On this basis there appears to be a subdivision within the superfamily, with ICL and PEPM in one group and DDGA, PEP carboxylase, pyruvate kinase, and pyruvate phosphate dikinase in the other.

The other class 2 aldolases for which structures are available are all from *E. coli*: fructose-1,6-bisphosphate adolase (Protein Data Bank code 1dos) and tagatose-1,6-bisphosphate aldolase (1gvf), which are assigned to the adolase superfamily, and L-fuculose-1-phosphate aldolase (4fua) and L-ribulose-5-phosphate 4-epimerase (1jdi). The latter two enzymes do not adopt the $(\beta\alpha)_8$ fold and belong to the class 2 superfamily within the α and β (α/β) fold (SCOP database; Murzin et al., 1995). In each of these four structures, the environment of the catalytic divalent cation is different from the magnesium coordination site found in KPHMT. As the class 2 aldolases are spread out across folds and superfamilies, membership of this functional group is not a particularly helpful criterion for understanding structure and function in KPHMT. Nor is the identity of the second substrate or electrophilic donor a basis for classification, as folate binding is not a trait of any existing $(\beta\alpha)_8$ superfamily.

The structural comparison of KPHMT with PEPM and ICL shows exact equivalence and conservation of functional sites. We therefore propose that KPHMT should be assigned to the phosphoenolpyruvate/pyruvate superfamily and should form a family on its own within this superfamily. Though functionally distinct and mechanistically diverse, members of the phosphoenolpyruvate/pyruvate superfamily use a catalytic Mg^{2+} as an electron sink and to position and orient the α-ketoacid substrate.

Biological Implications

This structure of *E. coli* KPHMT completes our structural analysis of pantothenate synthesis enzymes in this organism, together with the structures of aspartate decarboxylase (Albert et al., 1998), ketopantoate reductase (Matak-Vinkovic et al., 2001), and pantothenate synthetase (von Delft et al., 2001). We were able to identify the active sites in all cases. The enzymes are structurally unrelated, but, interestingly, all were seen to adopt previously known folds.

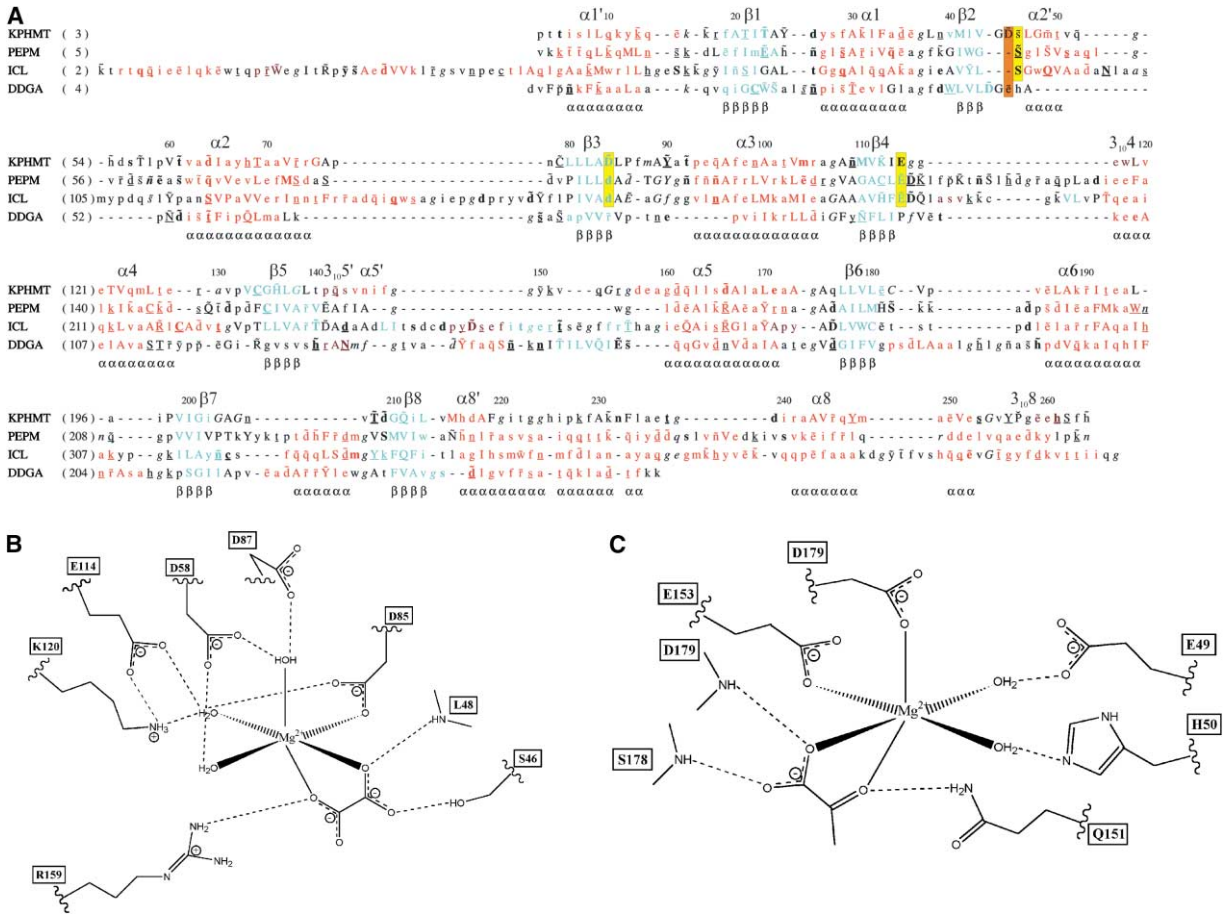


Figure 5. Comparison of KPHMT with $(\beta\alpha)_8$ Homologs

(A) Structure-based sequence alignment of KPHMT, PEPM, ICL, and DDGA. The 3D structures were aligned with COMPARER (Sali and Blundell, 1990) and formatted in JOY (Mizuguchi et al., 1998), indicating the local structural environment. The secondary structural elements are also denoted according to KPHMT. Every tenth residue in KPHMT is numbered. The formatting convention of JOY is as follows: red, α helices; maroon, 3_{10} helices, blue, β strands; uppercase letters, solvent inaccessible; lower case letters, solvent accessible; bold type, hydrogen bonds to main chain amides; underlined, hydrogen bonds to main chain carbonyls; $\sim$, hydrogen bonds to other side chain and/or heterogeneous groups; italic, positive main chain torsion angles (ϕ). A consensus of the secondary structure is shown beneath the sequences. Residues Ser46, Asp84, and Glu114 in KPHMT together with the equivalent residues in PEPM and ICL are highlighted in yellow. Residue Asp45 in KPHMT together with the aligned Glu49 in DDGA is highlighted in brown.

(B) Schematic diagram of the hydrogen-bonding network in the active sites of PEPM and DDGA, in the same orientation as that for KPHMT in Figure 3C.

The structure of KPHMT adopts a $(\beta\alpha)_8$ barrel fold and reveals that it is part of the PEP/pyruvate superfamily. SeMet MAD phasing enabled the 160 atoms of the selenium atom substructure to be located by SnB (Weeks and Miller, 1999) with anomalous differences measured at a single wavelength. Although this is the largest substructure solution by direct methods reported to date, the procedure was quite general, so that this is unlikely to represent the upper achievable limit, and, indeed, we believe that much larger structures would also succumb.

The product, ketopantoate, is observed in the active site and indicates the mode of substrate binding: the two carbonyl groups of α-KIVA and three enzyme carboxylates coordinate octahedrally with Mg^{2+}, which orients C3 of the substrate for abstraction of its acidic proton to produce the nucleophilic enolate reaction intermediate. KPHMT is the first enzyme in the pantothenate pathway and is inhibited by later intermediates, namely, pantoate, pantothenate, and CoA (Powers and Snell, 1976). All three effectors exhibit negative feedback, decreasing V_{max}, increasing K_m, and enhancing cooperativity for the substrate (Powers and Snell, 1976). The regulation of enzyme activity by later pathway intermediates may be linked to the decameric architecture of the enzyme, through subunit interactions. The crystal structure should enable this hypothesis to be explored more systematically.

Experimental Procedures

Preparation of Recombinant Protein

The *panB* gene from *E. coli* K12 was amplified by PCR with primers designed against the sequence of the gene reported by Merkel and Nichols (1996). The gene was cloned into pUC18, so that it was under the control of the *lacZ* promoter, and the plasmid was transformed into *E. coli* strain Hfr3000 YA139, which is a *panB* mutant (Cronan, 1980). The native recombinant protein was overexpressed

overnight at 37°C on LB medium by induction with IPTG for the duration of the culture and then purified to homogeneity, in a modification of the procedure described by Jones et al. (1993), by ammonium sulfate fractionation (0%–60% w/v), double anion exchange (Pharmacia Hiprep Q XL and MonoQ columns), and gel filtration chromatography (Pharmacia HiLoad Superdex 200 pg column). The reported heating step was omitted.

Preparation of SeMet-Labeled Protein

Selenium-labeled protein was produced as described by van Duyne et al. (1993) by methionine pathway feedback inhibition in the presence of 50 mg/l SeMet. Transformed Hfr3000 Y139 cells were grown in minimal medium containing ampicillin (100 μg/ml); starter cultures lacked selenomethionine and inhibitory amino acids. A 1 ml starter culture was used to inoculate 250 ml of culture medium, and overexpression was continuously induced with IPTG at 37°C. The labeled protein was purified as described for the native enzyme, without added antioxidants.

Electrospray Mass Spectrometry

To assess the incorporation of SeMet, we injected (at 4 μl/min) 10 μl of either native or SeMet-substituted protein (0.1 mg/ml) in aqueous formic acid (2.5%) and acetonitrile (50%) into an electrospray quadrupole mass spectrometer (Micromass Quattro LC) calibrated with horse heart myoglobin. Fifty averaged quadrupole scans were baseline corrected and smoothed, and at least 20 peaks were transformed to calculate the main component mass.

Crystallization and Cryomounting

Both native and SeMet-labeled protein crystallized under similar conditions, by hanging drop vapor diffusion. Reservoir solutions contained PEG 8000 (9%–11% v/v), NaCl (100–200 mM), sodium acetate (50–100 mM), and sodium citrate buffer (50 mM [pH 6.8]). Reservoir solution was mixed in equal volumes with protein-ketopantoate solution (final concentrations of 24–32 mg/ml and 40 mM, respectively; 50 mM HEPES [pH 7.4]) to form hanging drops of 1–4 μl. Crystallization trials were kept at 4°C, and large, irregular plate-like crystals appeared within 2 hr. Trials with numerous small crystals occasionally produced large, regular crystals when subsequently reequilibrated to 19°C.

Two crystal forms were used for structure solution. Crystals of form A were irregular plates, generally with layered or jagged appearance, up to 500 μm long, but no thicker than 50 μm (and half this size for the SeMet protein). They were fragile, with an absolute requirement for product, ketopantoate, during cryoprotection, and highly sensitive to the time spent soaking in each concentration of cryobuffer (15 min) and the use of mounting loops significantly larger than the crystal itself. A second type, form B, was observed in hanging drops of the native protein after 4 months. They were smooth, tear drop-shaped crystals with dimensions of 100 × 100 × 200 μm. These were not observed for the SeMet protein. Crystals were cryosoaked for a minimum of 15 min successively in 5%, 10%, 15%, and 20% (v/v) PEG 400-enriched mother liquours containing 2 mM ketopantoate (and 10 mM DTT for SeMet-labeled protein) before cryogenic mounting in cryoloops that were at least three times larger than the diameter of the crystal to ensure low crystal mosaicity.

Data Collection

A two-wavelength MAD data set was collected at the SBC-CAT beamline ID-19 at the APS (Argonne National Laboratory) from a SeMet-containing crystal (form A), selected after extensive screening for diffraction strength, low mosaicity, low background scatter, and size. The wavelength of peak absorbance (peak) was obtained from near-identical experimental fluorescence spectra from three different crystals. The peak data set consisted of three 360° ϕ oscillation segments; the crystal was reset by 20° (κ) after each segment and translated after the second, when radiation decay was apparent, to expose fresh crystal volume. The "remote" consisted of a 360° oscillation. The beam was attenuated and exposure times were brief (5 s /0.5° oscillation), resulting in weak images and a lower effective resolution limit, but enabling crystal survival for 3.5 hr of effective exposure and, hence, highly redundant measurements and eliminating detector saturation by strong reflexions; care was taken to measure low-resolution data. A native data set of the form B crystal was collected in two passes, to allow measurement of high- and low-resolution data.

Images were corrected for distortion and dark current with D*TREK (Pflugrath, 1999), integrated with MOSFLM, and scaled and merged with SCALA with the "smooth" scaling option (CCP4, 1994). MAD data sets were treated separately and scaled together with FHSCAL (CCP4, 1994).

Substructure Solution and Phasing

Anomalous differences of the peak data set were normalized with DREAR (Blessing and Smith, 1999) with default settings and used to solve the substructure with SnB (Weeks and Miller, 1999). Parameters were extensively explored, and successful trials could be reliably identified from an R_{min} significantly lower than average. The top 120 sites of the first successful trial were entered into SHARP (de la Fortelle and Bricogne, 1997), and the heavy-atom model was refined initially against the peak data only; the remaining 40 sites were located in log-likelihood gradient residual maps (Bricogne, 1997) after the first iteration. Occupancies were fixed to unity throughout and coordinates were refined only in the first iteration, in order to speed computation. The "remote" data were only added for final phase calculation to 3.0 Å, and anomalous scattering factors were obtained from fluorescence spectra with CHOOCH (Evans, 1999).

The 20 NCS operators of the form A crystal were generated from the substructure, and phases from SHARP were improved by NCS averaging and solvent flattening with DM. With a preliminary model of the decamer, derived from a skeleton of the excellent maps, the decamer was located in form B with AMORE, and phases were extended to the resolution of the form B native data set by multicrystal averaging with DMMULTI (CCP4, 1994).

Model Building and Refinement

The initial model was built into electron density maps from averaging with O (Jones et al., 1991) and iteratively improved by alternating refinement of torsional angles in CNS (Brünger et al., 1998) with simulated annealing, bulk solvent correction, and a randomly assigned free set for R_{free} crossvalidation (Brünger, 1992), with manual rebuilding in O into σ_A-weighted (Read, 1986) $2F_o - F_c$ maps. Ordered solvent was modeled into peaks of 1.2 SD above background. Because of the high NCS, the test reflection set was reassigned (in 38 resolution shells of 0.0003 $Å^{-1}$ width; 2% of total), and atoms were randomized before a second stage of iterative model building in XFIT (McRee, 1999) and restrained, TLS refinement with REFMAC (CCP4, 1994). This included two cycles of automated solvent building with ARP/WARP. Each subunit, together with its ligands and active site waters, was defined as a separate TLS group, and loose NCS restraints were maintained throughout. The ligand target geometry was derived from expected bond energies; however, no planarity was imposed on the carboxyl-carbonyl group.

Model quality was assessed with PROCHECK (Laskowski et al., 1993), MOLPROBITY (http://kinemage.biochem.duke.edu/molprobity/), and WHATCHECK (Hooft et al., 1996).

Analysis of *panB* Mutants

Mutant strain *E. coli* Hfr3000 YA139 (Cronan, 1980) was obtained from the *E. coli* Genetic Stock Center (CGSC), Yale University, Connecticut. Both *S. typhimurium* mutant strains *pan-4* and *pan-6* (Cronan et al., 1982) were kindly provided by Prof. John E. Cronan, Jr., University of Illinois. All three mutants exhibited measurable rates of reversion to the wild-type, so, to select auxotrophs, we streaked colonies on GB1 minimal media and GB1 minimal media supplemented with pantothenate (100 mg/l). Colonies that grew only on the supplemented media were genuine *panB* mutants.

The *panB* genes were amplified directly from auxotrophs by colony PCR with the proofreading polymerase, Pfu (Promega, UK). The *E. coli panB* gene was amplified with a forward primer (5′-CTGCTACGGATCCATCATTGCCCC-3′) that primed 57 bp upstream from the *panB* start codon and a reverse primer (5′-GCCACGAATTC GCCTTCCATACGC-3′) that primed 61 bp downstream of the respective stop codon. Similarly, the *S. typhimurium panB* gene was

amplified with primers that annealed 7 bp upstream of the start codon (5′-GCCGGGATCCATAACTTTTACCATCAGG-3′) and 31 bp downstream from the stop codon (5′-GAGGAATTCTAGGCCGGAT AAGACGC-3′). The PCR products were sequenced at the DNA Sequencing Facility, Department of Genetics, University of Cambridge. Identical mutations were found in two independently amplified PCR products for each mutant.

Model and Sequence Analysis

Surface areas were calculated with the CCP4 (CCP4, 1994) program SURFACE, without taking into account hydrogen atoms. Structural homologs were identified with DALI (Britton et al., 2001). Structural alignments were generated with COMPARER (Sali and Blundell, 1990) and formatted in JOY (Mizuguchi et al., 1998).

Acknowledgments

We thank Martyn Symmons for help with mass spectrometry, Sue Bethel for help with analytical ultracentrifugation, and John E. Cronan, Jr., for *S. typhimurium* mutant strains. Use of the Argonne National Laboratory Structural Biology Center beamline ID19 at the Advanced Photon Source was supported by the U.S. Department of Energy, Office of Biological and Environmental Research, under contract number W-31-109-ENG-38. Funding from the Oliver Gatty Foundation, BBSRC, EPSRC, Aventis Crop Science (UK), Cambridge European Trust, and Astex Technology is gratefully acknowledged.

Received: December 20, 2002
Revised: April 17, 2003
Accepted: May 12, 2003
Published: August 5, 2003

References

Aberhart, D.J., and Russell, D.J. (1984). Steric course of ketopantoate hydroxymethyltransferase in *Escherichia coli*. J. Am. Chem. Soc. *106*, 4902–4906.

Albert, A., Dhanaraj, V., Genschel, U., Khan, G., Ramjee, M.K., Pulido, R., Sibanda, B.L., von Delft, F., Witty, M., Blundell, T.L., et al. (1998). Crystal structure of aspartate decarboxylase at 2.2 Å resolution provides evidence for an ester in protein self-processing. Nat. Struct. Biol. *5*, 289–293.

CCP4 (Collaborative Computational Project, Number 4) (1994). The CCP4 suite—programs for protein crystallography. Acta Crystallogr. D Biol Crystallogr. *50*, 760–763.

Blessing, R.H., and Smith, G.D. (1999). Difference structure-factor normalization for heavy-atom or anomalous-scattering substructure determinations. J. Appl. Crystallogr. *32*, 664–670.

Bricogne, G. (1997). Bayesian statistical viewpoint on structure determination: basic concepts and examples. Methods Enzymol *276*, 361–423.

Britton, L., Abeysinghe, I.S.B., Baker, P.J., Barynin, V., Diehl, P., Langridge, S.J., McFadden, B.A., Sedelnikova, S.E., Stillman, T.J., Weeradechapona, K., et al. (2001). The structure and domain organization of *Escherichia coli* isocitrate lyase. Acta Crystallogr. D Biol. Crystallogr. *57*, 1209–1218.

Brünger, A.T. (1992). Free R value—a novel statistical quantity for assessing the accuracy of crystal structures. Nature *355*, 472–475.

Brünger, A.T., Adams, P.D., Clore, G.M., DeLano, W.L., Gros, P., Grosse-Kunstleve, R.W., Jiang, J.S., Kuszewski, J., Nilges, M., Pannu, N.S., et al. (1998). Crystallography and NMR system: a new software suite for macromolecular structure determination. Acta Crystallogr. D Biol. Crystallogr. *54*, 905–921.

Cronan, J.E., Jr. (1980). β-alanine synthesis in *Escherichia coli*. J. Bacteriol. *141*, 1291–1297.

Cronan, J.E., Jr., Little, K.J., and Jackowski, S. (1982). Genetic and biochemical analysies of pantothenate biosynthesis in *Escherichia coli* and *Salmonella typhimurium*. J. Bacteriol. *149*, 916–922.

de la Fortelle, E., and Bricogne, G. (1997). Maximum-likelihood heavy-atom parameter refinement for multiple isomorphous replacement and multiwavelength anomalous diffraction methods. Methods Enzymol. *276*, 472–494.

Diederichs, K., and Karplus, A. (1997). Improved R factors for diffraction data analysis in macromolecular crystallography. Nat. Struct. Biol. *4*, 269–275.

Evans, Q.G., and Pettifer, R.F. (2001). J. Appl. Crystallogr. *34*, 82–86.

Holm, L., and Sander, C. (1993). Protein structure comparison by alignment of distance matrices. J. Mol. Biol. *233*, 123–128.

Hooft, R.W., Vriend, G., Sander, C., and Abola, E.E. (1996). Errors in protein structures. Nature *381*, 272.

Huang, K., Li, Z., Jia, Y., Dunaway-Mariano, D., and Herzberg, O. (1999). Helix swapping between two alpha/beta barrels: crystal structure of phosphoenolpyruvate mutase with bound Mg(2+)-oxalate. Structure *7*, 539–548.

Izard, T., and Blackwell, N.C. (2000). Crystal structures of the metal-dependent 2-dehydro-3-deoxy-galactarate aldolase suggest a novel reaction mechanism. EMBO J. *19*, 3849–3856.

Jones, C.E., Brook, J., Buck, D., Abell, C., and Smith, A.G. (1993). Cloning and sequencing of the *Escherichia coli panB* gene, which encodes ketopantoate hydroxymethyltransferase, and overexpression of the enzyme. J. Bacteriol. *175*, 2125–2130.

Jones, T.A., Zou, J.Y., Cowan, S.W., and Kjeldgaard, M. (1991). Improved methods for building protein models in electron density maps and the location of errors in these models. Acta Crystallogr. A *47*, 110–119.

Kleinkauf, H. (2000). The role of 4′-phosphopantetheine in the biosynthesis of fatty acids, polyketides and peptides. Biofactors *11*, 91–92.

Kraulis, P.J. (1991). MOLSCRIPT: a program to produce both detailed and schematic plots of protein structures. J. Appl. Crystallogr. *24*, 946–950.

Kurtov, D., Kinghorn, J.R., and Unkles, S.E. (1999). The *Aspergillus nidulans panB* gene encodes ketopantoate hydroxymethyltransferase, required for biosynthesis of pantothenate and coenzyme A. Mol. Gen. Genet. *262*, 115–120.

Laskowski, R.A., MacArthur, M.W., Moss, D.S., and Thornton, J.M. (1993). PROCHECK—a program to check the stereochemical quality of protein structures. J. Appl. Crystallogr. *26*, 283–291.

Matak-Vinkovic, D., Vinkovic, M., Saldanha, S.A., Ashurst, J.L., von Delft, F., Inoue, T., Miguel, R.N., Smith, A.G., Blundell, T.L., and Abell, C. (2001). Crystal structure of *Escherichia coli* ketopantoate reductase at 1.7 Å resolution and insight into the enzyme mechanism. Biochemistry *40*, 14493–14500.

McRee, D.E. (1999). XtalView/Xfit—a versatile program for manipulating atomic coordinates and electron density. J. Struct. Biol. *125*, 156–165.

Merkel, W.K., and Nichols, B.P. (1996). Characterization and sequence of the *Escherichia coli panBCD* gene cluster. FEMS Microbiol. Lett. *143*, 247–252.

Merritt, E.A., and Murphy, M. (1994). Raster3D version 2.0: a program for photorealistic molecular graphics. Acta Crystallogr. D Biol. Crystallogr. *50*, 869–873.

Mizuguchi, K., Deane, C.M., Blundell, T.L., Johnson, M.S., and Overington, J.P. (1998). JOY: protein sequence-structure representation and analysis. Bioinformatics *14*, 617–623.

Murzin, A.G., Brenner, S.E., Hubbard, T., and Chothia, C. (1995). SCOP: a structural classification of proteins database for the investigation of sequences and structures. J. Mol. Biol. *247*, 536–540.

Nicholls, A., Sharp, K.A., and Honig, B. (1991). Protein folding and association: insights from the interfacial and thermodynamic properties of hydrocarbons. Proteins *11*, 281–296.

Pflugrath, J.W. (1999). The finer things in X-ray diffraction data collection. Acta Crystallogr. D Biol. Crystallogr. *55*, 1718–1725.

Powers, S.G., and Snell, E.E. (1976). Ketopantoate hydroxymethyltransferase. II. Physical, catalytic, and regulatory properties. J. Biol. Chem. *251*, 3786–3793.

Read, R.J. (1986). Improved Fourier coefficients for maps using

phases from partial structures with errors. Acta Crystallogr. A *42*, 140–149.

Sahi, S.V., Saxena, P.K., Abrahams, G.D., and King, J. (1988). Identification of the biochemical lesion in pantothenate-requiring auxotroph of *Datura innoxia* P mill. J. Plant Physiol. *133*, 277–280.

Sali, A., and Blundell, T.L. (1990). Definition of general topological equivalence in protein structures. A procedure involving comparison of properties and relationships through simulated annealing and dynamic programming. J. Mol. Biol. *212*, 403–428.

Shi, J., Blundell, T.L., and Mizuguchi, K. (2001). FUGUE: sequence-structure homology recognition using environment-specific substitution tables and structure-dependent gap penalties. J. Mol. Biol. *310*, 243–257.

Sugantino, M., Zheng, R., Yu, M., and Blanchard, J.S. (2003). *Mycobacterium tuberculosis* ketopantoate hydroxymethyltransferase: tetrahydrofolate-independent hydroxymethyltransferase and enolization reactions with alpha-keto acids. Biochemistry *42*, 191–199.

Teller, J.H., Powers, S.G., and Snell, E.E. (1976). Ketopantoate hydroxymethyltransferase. I. Purification and role in pantothenate biosynthesis. J. Biol. Chem. *251*, 3780–3785.

Van Duyne, G.D., Standaert, R.F., Karplus, P.A., Schreiber, S.L., and Clardy, J. (1993). Atomic structures of the human immunophilin FKBP-12 complexes with FK506 and rapamycin. J. Mol. Biol. *229*, 105–124.

von Delft, F., Lewendon, A., Dhanaraj, V., Blundell, T.L., Abell, C., and Smith, A.G. (2001). The crystal structure of *E. coli* pantothenate synthetase confirms it as a member of the cytidylyltransferase superfamily. Structure *9*, 439–450.

Weeks, C.M., and Miller, R. (1999). The design and implementation of SnB version 2.0. J. Appl. Crystallogr. *32*, 120–124.

Accession Numbers

Coordinates have been deposited in the Protein Data Bank with accession number 1m3u.