Crystal Structure of *Escherichia coli* Ketopantoate Reductase at 1.7 Å Resolution and Insight into the Enzyme Mechanism†

Dijana Matak-Vinković,‡,§ Mladen Vinković,‡,|| S. Adrian Saldanha,⊥ Jennifer L. Ashurst,# Frank von Delft,‡ Tsuyoshi Inoue,‡,@ Ricardo Nunez Miguel,‡ Alison G. Smith,# Tom L. Blundell,‡ and Chris Abell*,⊥

Department of Biochemistry, University of Cambridge, 80 Tennis Court Road, Cambridge CB2 1GA, U.K., University Chemical Laboratory, Lensfield Road, Cambridge CB2 1EW, U.K., Department of Plant Sciences, University of Cambridge, Downing Street, Cambridge CB2 3EA, U.K.

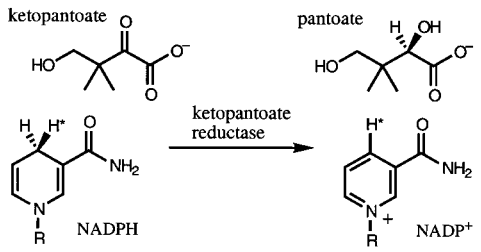
Received May 18, 2001; Revised Manuscript Received September 14, 2001

ABSTRACT: Ketopantoate reductase (KPR, EC 1.1.1.169) catalyzes the NADPH-dependent reduction of ketopantoate to pantoate on the pantothenate (vitamin B_5) biosynthetic pathway. The *Escherichia coli panE* gene encoding KPR was cloned and expressed at high levels as the native and selenomethionine-substituted (SeMet) proteins. Both native and SeMet recombinant proteins were purified by three chromatographic steps, to yield pure proteins. The wild-type enzyme was found to have a K_M(NADPH) of 20 μM, a K_M(ketopantoate) of 60 μM, and a k_{cat} of 40 s^{-1}. Regular prismatic KPR crystals were prepared using the hanging drop technique. They belonged to the tetragonal space group $P4_22_12$, with cell parameters: $a = b = 103.7$ Å and $c = 55.7$ Å, accommodating one enzyme molecule per asymmetric unit. The structure of KPR was determined by the multiwavelength anomalous dispersion method using the SeMet protein, for which data were collected to 2.3 Å resolution. The native data were collected to 1.7 Å resolution and used to refine the final structure. The secondary structure comprises 12 α-helices, three 3_{10}-helices, and 11 β-strands. The enzyme is monomeric and has two domains separated by a cleft. The N-terminal domain has an $\alpha\beta$-fold of the Rossmann type. The C-terminal domain (residues 170−291) is composed of eight α-helices. KPR is shown to be a member of the 6-phosphogluconate dehydrogenase C-terminal domain-like superfamily. A model for the ternary enzyme−NADPH−ketopantoate ternary complex provides a rationale for kinetic data reported for specific site-directed mutants.

Ketopantoate reductase (KPR,[1] EC 1.1.1.169) catalyzes the NADPH-dependent reduction of ketopantoate to pantoate as part of the pantothenate biosynthetic pathway (Scheme 1). Subsequently, pantoate is condensed with β-alanine to form pantothenate, which is the precursor to the phosphopantetheine moiety in coenzyme A and acyl carrier protein. Only plants and microorganisms possess this pathway, so animals must obtain pantothenate (vitamin B_5) in their diet.

The pathway of pantothenate biosynthesis was elucidated in *Escherichia coli* and *Salmonella typhimurium* by biochemical and genetic experiments, and three of the four genes (*panB*, *panC*, and *panD*) were mapped close together on the *E. coli* genome (*1*, *2*). The *panE* gene encoding KPR remained elusive, however, due in part to the fact that an enzyme of valine and isoleucine biosynthesis, acetohydroxy-acid isomeroreductase (AHIR, EC 1.1.1.86), encoded by *ilvC*, can catalyze the same reaction as KPR, albeit less efficiently. Mutants of *ilvC* still retain KPR activity, whereas *ilvC panE* double mutants have no measurable enzyme activity and are pantoate auxotrophs (*3*, *4*). The *panE* gene from *S. typhimurium* has recently been shown to be allelic to the *apbA* gene (*5*). This gene, required for the function of the alternative pyrimidine biosynthetic pathway, encodes an enzyme with KPR activity. Mutagenesis of the endogenous *apbA* gene in an *ilvC* background produced strains that required pantoate or pantothenate for growth and were devoid of KPR activity (*6*).

Scheme 1: NADPH-Dependent Reduction of Ketopantoate Catalyzed by KPR



† This work was supported by Astex Technology (Cambridge, U.K.), the BBSRC, the Pfizer Award for Innovation, Rhone Poulenc Agriculture, Ministerio Espanol de Educacion y Cultura, and The Oliver Gatty Trust.

* To whom correspondence should be addressed. Telephone: +44-1223-336405. Fax: +44-1223-336362. E-mail: ca26@cam.ac.uk.

‡ Department of Biochemistry, University of Cambridge.

§ Present address: Faculty of Science, University of Zagreb, Zvonimirova 8, 10000 Zagreb, Croatia.

|| Present address: PLIVA d.d., Prilaz baruna Filipovica 25, 10000 Zagreb, Croatia.

⊥ University Chemical Laboratory.

Department of Plant Sciences, University of Cambridge.

@ Present address: Department of Materials Chemistry, Graduate School of Engineering, Osaka University, Yamada-Oka 2-1, Suita, Osaka 565-871.

[1] Abbreviations: AHIR, acetohydroxy acid isomeroreductase; CENDH, *N*-(1-D-carboxylethyl)-L-norvaline dehydrogenase; ESMS, electrospray mass spectrometry; IPTG, isopropyl thio-β-D-galactoside; KPR, ketopantoate reductase; MAD, multiwavelength anomalous dispersion; SeMet, selenomethionine.

10.1021/bi011020w CCC: $20.00
Published on Web 11/06/2001

KPR has been purified and partially characterized from *Pseudomonas maltophilia* 845, and found to comprise three to five subunits with an M_r of 30 000 (*7*). In contrast, recombinant KPR from both *S. typhimurium* (*5*) and *E. coli* (*8*) are monomeric, with M_rs of 31 100 and 34 000, respectively. It has been reported that *S. typhimurium* KPR has a 3.8-fold higher specific activity with NADPH than with NADH (*5*). The recombinant *E. coli* enzyme catalyzes the transfer of the *pro-S* hydrogen from NADPH. The enzyme has an ordered bi-bi mechanism with initial binding of NADPH and, in contrast to AHIR, does not require divalent metal ions for catalysis (*8*). Several site-directed mutants have been characterized to determine their kinetic parameters, and to identify residues that are important in the catalytic mechanism (*9*). The involvement of Lys176 as a general acid or base and Glu256 in binding to the substrate was postulated. Mutations of each of these residues to alanine had significant deleterious effects on enzyme activity, and the double mutant was inactive.

In this paper, we report the overexpression and purification of both native and selenomethionine (SeMet)-substituted KPR from *E. coli*. These have been crystallized, and the X-ray structure of KPR has been determined to 1.7 Å resolution. The structure has been used to predict the binding site of NADPH and a model for the ternary complex.

MATERIALS AND METHODS

Cloning of the E. coli panE Gene. All molecular biological methods were as described (*10*), except where stated otherwise. The *E. coli panE (apbA)* gene sequence had been submitted to the EMBL database (accession number U34923), and this sequence is part of the plasmid pVJS78, which contains a 6 kb fragment of *E. coli* genomic DNA (obtained from T. Begley, Department of Chemistry, Cornell University, Ithaca, NY). A PCR product, representing the full-length *panE* gene, was generated using primers EPANE*Nde*F (5′-ATACATATGAAAATTACCGTATTGGGATG-3′) and EPANE*Bam*HIR (5′-TATGGATCCTCACCAGGGGCGAGGCAAACC-3′), which had *Nde*I and *Bam*HI restriction sites, respectively (underlined). This 915 bp PCR product was digested with an *Nde*I−*Bam*HI fragment and ligated into an *Nde*I−*Bam*HI-restricted pET24b vector (Novagen). The *panE* gene was sequenced and found to be identical to that in the database. Crude soluble protein extracts of the bacterial cells harboring pET24b-*panE* had measurable KPR activity (data not shown).

Protein Expression and Purification. A 6 L culture of *E. coli* BL21 cells containing pET24b-*panE* was grown in LB medium containing kanamycin (50 μg/mL) at 37 °C to an OD_{600} of 0.6−0.7, and expression of KPR was induced by addition of isopropyl thio-β-D-galactoside (IPTG) to a final concentration of 0.5 mM. The cells were harvested after 4 h and resuspended in extraction buffer containing 100 mM KH_2PO_4 (pH 7.5), 1 mM EDTA, and 1 mM PMSF. Cells were lysed by sonication and centrifuged. Nucleic acids were precipitated by the addition of 5 mL protamine sulfate (2% w/v) and centrifugation. The supernatant was fractionated with solid ammonium sulfate, and the precipitate occurring between 5 and 40% saturation was collected by centrifugation. The protein was redissolved in 25 mM KH_2PO_4 (pH 7.5) containing 1 mM EDTA, dialyzed against 2 L of the same buffer, concentrated to less than 10 mL of buffer containing 25 mM potassium chloride, and treated to fast phase liquid chromatography. First, the protein was applied to a Hiprep 16/10 Q XL anion exchange column and eluted with a linear 25 to 500 mM potassium chloride gradient. Fractions containing KPR were concentrated and desalted. The protein solution, in less than 10 mL of buffer containing 25 mM potassium chloride, was then applied to a Source 15Q XV 16/10 anion exchange column and eluted with a linear 25 to 500 mM potassium chloride gradient. Protein, in less than 5 mL of buffer, was then applied to a Hiload 16/60 Superdex 200 gel filtration column, and fractions containing KPR were concentrated to >5 mg/mL. The protein yield was 13 mg from 1 L of culture. Fractions were analyzed by SDS−PAGE, according to the method of Laemmli (*11*). Protein concentrations were determined using the Bio-Rad protein assay kit, using bovine serum albumin as a standard, and confirmed spectrophotometrically using A_{280} (ϵ calculated) (*12*). N-Terminal sequencing determined the first nine residues to be MKITVLGCG, identical to the sequence predicted from the *panE* gene. Amino acid analysis (data not shown) was also consistent with that predicted from the gene. Electrospray mass spectrometry (ESMS) was carried out on a VG BioQ quadrupole mass spectrometer. The mass of the protein, calculated from the predicted amino acid sequence (33 870 Da), compares well with the ESMS value (33 876 Da) determined for the recombinant protein.

SeMet-substituted KPR was prepared as described by van Duyne et al. (*13*) by methionine pathway inhibition in the presence of SeMet (*14*). *E. coli* BL21 cells harboring pET24b-*panE* were grown at 37 °C to an OD_{600} of 0.35 in M9 minimal medium containing 4% glucose, 0.5% thiamine, and 100 mg/L kanamycin. The medium was then supplemented with six amino acids (Lys, Phe, Thr, Ile, Leu, and Val) and SeMet at 50 mg/L. Cells were induced using IPTG (100 mg/L) for 2.5 h and harvested. Purification of SeMet KPR was performed as described above, with the addition of 5 mM DTT to prevent oxidation of the SeMet. One milligram of SeMet protein was obtained from 1 L of cells.

Determination of KPR Activity. KPR activity was determined at 37 °C in the forward direction by monitoring the decrease in the level of NADPH over time (as absorbance at 340 nm) by a modification of the method described in ref *7*. The assay comprised 100 mM potassium phosphate buffer (pH 7.4), 200 μM NADPH, and 30 nM KPR and was initiated with 500 μM potassium ketopantoate, in a final volume of 1 mL.

Crystallization of KPR. Regular prismatic KPR crystals were prepared by the hanging drop vapor-diffusion technique, using 25−27% PEG 4000, 0.1 M Tris (pH 9.4), and 0.20−0.25 M NaOAc as the precipitant solution. The protein concentration was 11 mg/mL in 10 mM Hepes (pH 8.0). The hanging drops, consisting of 1 μL of protein solution and 1 μL of well solution, were equilibrated at 4 °C for several days. Crystals with dimensions of 0.4 mm × 0.2 mm × 0.2 mm grew from clear drops. These belonged to the tetragonal space group $P4_22_12$ with the following cell parameters: $a = b = 103.7$ Å and $c = 55.7$ Å, accommodating one enzyme molecule per asymmetric unit, with a solvent content of 45%.

The SeMet KPR crystals, prepared in a way similar to that of native KPR crystals, were obtained using two different

Table 1: X-ray Data Collection Statistics

space group	$P4_22_12$			
cell parameters (Å)	$a = b = 103.7$ Å, $c = 55.7$ Å			
	SeMet derivative			
	edge	peak	remote	native
wavelength (Å)	0.979272	0.9790501	0.885612	0.87
resolution (Å)	2.3	2.3	2.3	1.7
no. of unique reflections	13640	11259	13720	32249
redundancy	3.77	4.5	3.95	5.4
completeness (%)	95.2	96.3	97.5	94.9
R_{ano} (%)	5.8	6.8	4.6	−
R_{merge} (%)	9.9	10.8	9.4	4.3
f', f'' (experimental)	−11.1873 4.1205	−7.5644 5.9692	−1.5 −	−
f', f'' (theoretical)	−11.0597 4.38418	−7.46687 6.29644	−1.3 3.2	−
phasing power	2.288	2.889	2.108	−
R	0.230			
R_{free}	0.289			
no. of reflections				
working set	30610			
test set	1639			
Ramachandran plot				
% residues in most favorable region	91.9			
% residues in unfavorable region	none			
rms deviation				
bond lengths	0.016			
bond angles	1.9			
planarity	0.024			

protein concentrations, 8 and 16 mg/mL. The SeMet KPR stock solution contained 15 mM Tris (pH 8) and 2 mM DTT to protect the Se atoms from oxidation. Optimized crystallization conditions include 28−38% PEG 4000, 0.1 M Tris (pH 9.4), and 0.25 M NaOAc.

Data Collection and Processing. The structure of KPR was determined by the multiwavelength anomalous dispersion (MAD) method using the SeMet derivative. Data to 2.3 Å resolution were collected at 100 K on the BW7A beamline at the EMBL outstation in DESY (Hamburg, Germany). Flash-freezing of the crystal was preceded by soaking for 30 s in a cryoprotectant solution consisting of 15% glycerol, 30% PEG 4000, 0.1 M Tris (pH 9.4), and 0.25 M sodium acetate. An X-ray fluorescence spectrum was recorded and used to select wavelengths for subsequent MAD data collection. Data were collected at the Se absorption edge λ_e = 0.979272 Å (12 661.00 eV), the absorption peak λ_p = 0.9790501 Å (12 663.77 eV), and at a remote reference wavelength λ_r = 0.885612 Å (14 000 eV). The diffraction data were indexed and integrated using the DENZO package (*15*). The three data sets were scaled to the remote data set using SCALA (*16*, *17*), and structure factor amplitudes were calculated using TRUNCATE (*16*). Statistics of the processed data are listed in Table 1. The native data set was collected to 1.7 Å resolution at the Daresbury SRS beamline 9.6. A cryoprotectant solution contained 30% PEG1500, 100 mM Tris buffer (pH 9.4), 300 mM sodium acetate, and 20% glycerol.

Structure Determination and Refinement. Six selenium sites were found with the program SHELXS96 (*18*) using direct methods and anomalous difference data of λ_p SeMet. Data were phased with SHARP (*19*) using the data sets at three wavelengths, which also revealed two additional Se sites in the residual maps. Data collected at the remote wavelength were treated as the reference data set, and resolution limits of 40−2.3 Å were imposed. Experimental values of f' and f'' estimated from fluorescence spectra were used and refined during analysis. The resulting values are very similar to the theoretical values and are given in Table 1. Experimental phases were improved by solvent flattening using SOLOMON (*16*), via the SUSHI graphical user interface (*19*) with a solvent content of 43%. The final electron density map was easily interpretable, and the whole polypeptide chain was assigned on the basis of the initial electron density map.

The polypeptide chain was fitted into the MAD electron density map using program O (*20*). Rounds of maximum-likelihood refinement with REFMAC (*21*) were alternated with visual inspection of electron density and manual rebuilding of side chains. Several rounds of simulated annealing with CNS (*22*) were included to refine the position of the main chain properly. The current model contains 291 residues and 272 water molecules.

RESULTS AND DISCUSSION

Overexpression and Purification of Recombinant E. coli KPR. The *E. coli panE* gene encoding KPR was cloned into the expression vector pET24b, and expressed to high levels (13 mg/L). Both native and SeMet-substituted recombinant proteins were purified by three chromatographic steps, to yield the pure protein. Electrophoresis in the presence of SDS revealed a protein of 33 kDa for both the wild-type and SeMet-substituted forms (data not shown), consistent with the predicted M_r, and that reported previously (*8*). The protein behaved as a monomer on gel filtration. The enzyme was found to have a K_M(NADPH) of 20 μM, a K_M(ketopantoate) of 60 μM, and a k_{cat} of 40 s^{-1}. The values reported by Zheng and Blanchard using different assay conditions (*8*) are 4 μM, 120 μM, and 8 s^{-1}, respectively.

Crystallization trials using both the hanging and sitting drop techniques were carried out with Hampton crystal screen 6, but initial crystals that were obtained were an irregular rocket shape. After optimization of the conditions as described in Materials and Methods, regular prismatic KPR crystals were prepared, which diffracted well.

Structure Solution and Refinement. Initial phases were calculated by the MAD technique using three data sets collected at wavelengths on and away from the selenium Kα absorption edge. After solvent flattening had been carried out, an easily interpretable electron density map was obtained at 2.3 Å resolution (Figure 1). Secondary structural elements were clearly defined in the electron density "bones" calculated with MAPMAN (*23*). The asymmetric unit consists of one KPR monomer. The main chain of the molecule was traced using secondary structure template building functionality in O (*20*), and the sequence was easily determined from the positions of selenium atoms from SHARP (*19*).

The model was improved by restrained isotropic, maximum-likelihood refinement with REFMAC (*21*). Crystallographic refinement was alternated with manual model rebuilding in O (*20*). Torsional simulated annealing (starting temperature of 3000 K) in CNS was used to improve difficult parts of the model. Ordered water molecules were modeled by automated cycles of water addition and removal by ARP (*24*)

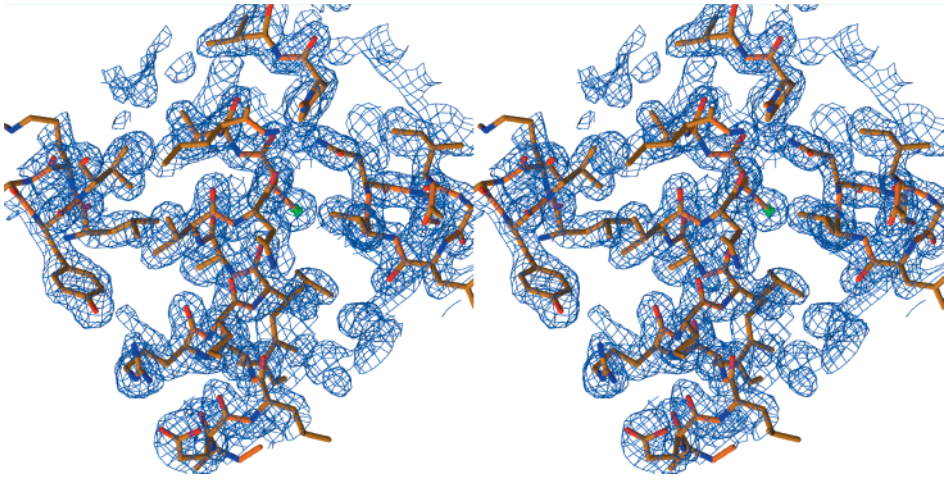


FIGURE 1: Final $2F_o - F_c$ map. The figure was prepared using MOLSCRIPT (*41*) and RASTER3D (*42*, *43*).

and refinement by REFMAC. A final model containing 291 residues and 272 water molecules was refined to an *R*-factor of 0.230 and an R_{free} of 0.289 (Table 1). The last 12 residues at the C-terminus were not visible in electron density maps. The quality of the model and its geometry were assessed by PROCHECK (*25*). There were no Ramachandran outliers, and 91.9% of the residues lie in the most favored regions of the plot.

Overall Structure of KPR. The secondary structure assignment of KPR, as analyzed using the DSSP program (*26*), shows the molecule to be constructed from 12 α-helices, three 3_{10}-helices, and 11 β-strands. The enzyme is monomeric and has two domains separated by a cleft (Figure 2). The N-terminal domain has an αβ-fold of the Rossmann type and comprises residues 1−169. It is formed from a seven-stranded mixed β-sheet with α-helices connecting adjacent parallel strands. One of the α-helices, α1, forms a core and sits in the concave face of the sheet. It is further enclosed by a four-stranded antiparallel β-sheet that is joined to the ends of the seven-stranded β-sheet by β-arches or bridges. The sequence between the end of strand β1 and helix α1 contains a glycine-rich region, including the sequence GCGALG (residues 7−12), which represents the motif GXGXXG that is very often an indicator of the nucleotide binding site in dehydrogenases (*27*). The C-terminal domain (residues 170−291) is composed of eight α-helices.

To identify other residues which might be important in the catalytic mechanism, database searches of bacterial genomes were carried out, and several putative KPRs were identified. However, except between the *E. coli* and *S. typhimurium* orthologues, which are 82% identical, the overall level of similarity between sequences was very low, at less than 30% on average. In the absence of experimental evidence that a particular gene encodes a protein with KPR activity, its assignment as a gene encoding the enzyme cannot be certain. Nevertheless, an alignment of 13 bacterial sequences identified a total of 16 completely conserved residues, compared to the 20 identified in seven bacterial KPR sequences by Zheng and Blanchard (*8*). The conserved residues, which include the GXGXXG motif, are all located in or close to the cleft between the two domains (Figure 2), with the notable exception of Glu210.

Structural and Functional Comparison of KPR with Other Enzymes. A search with the Dali server (*29*, *30*) for proteins

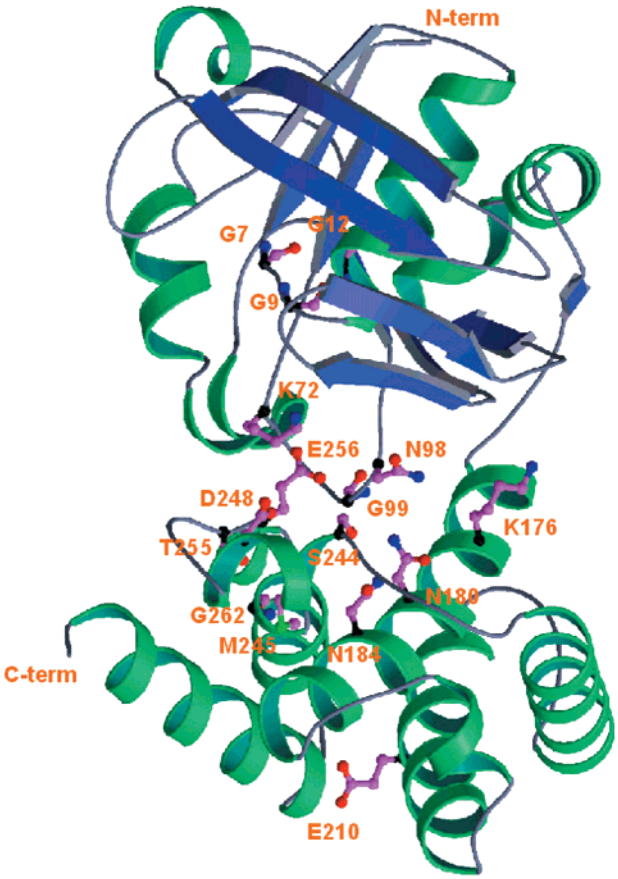


FIGURE 2: Structure of KPR, highlighting the positions of the 16 strictly conserved residues.

with a structure similar to that of KPR has identified five distant homologues: AHIR from spinach (*31*, *32*), *Trypanosoma brucei* 6-phosphogluconate dehydrogenase (6PGDH) (*33*), UDP-glucose dehydrogenase (UDPGDH) from *Streptococcus pyrogenes* (*34*), human short chain L-3-hydroxyacyl CoA dehydrogenase (SCHAD) (*35*), and *N*-(1-D-carboxyethyl)-L-norvaline dehydrogenase (CENDH) from *Arthrobacter* sp. (*36*). Among these proteins, only the nucleotide binding site residues are conserved, and the overall level of sequence identity is ≤20%. All these proteins belong to the 6-phosphogluconate dehydrogenase C-terminal domain-like superfamily from the SCOP database (*37*). SCOP defines the fold of the superfamily as a Rossmann fold N-terminal

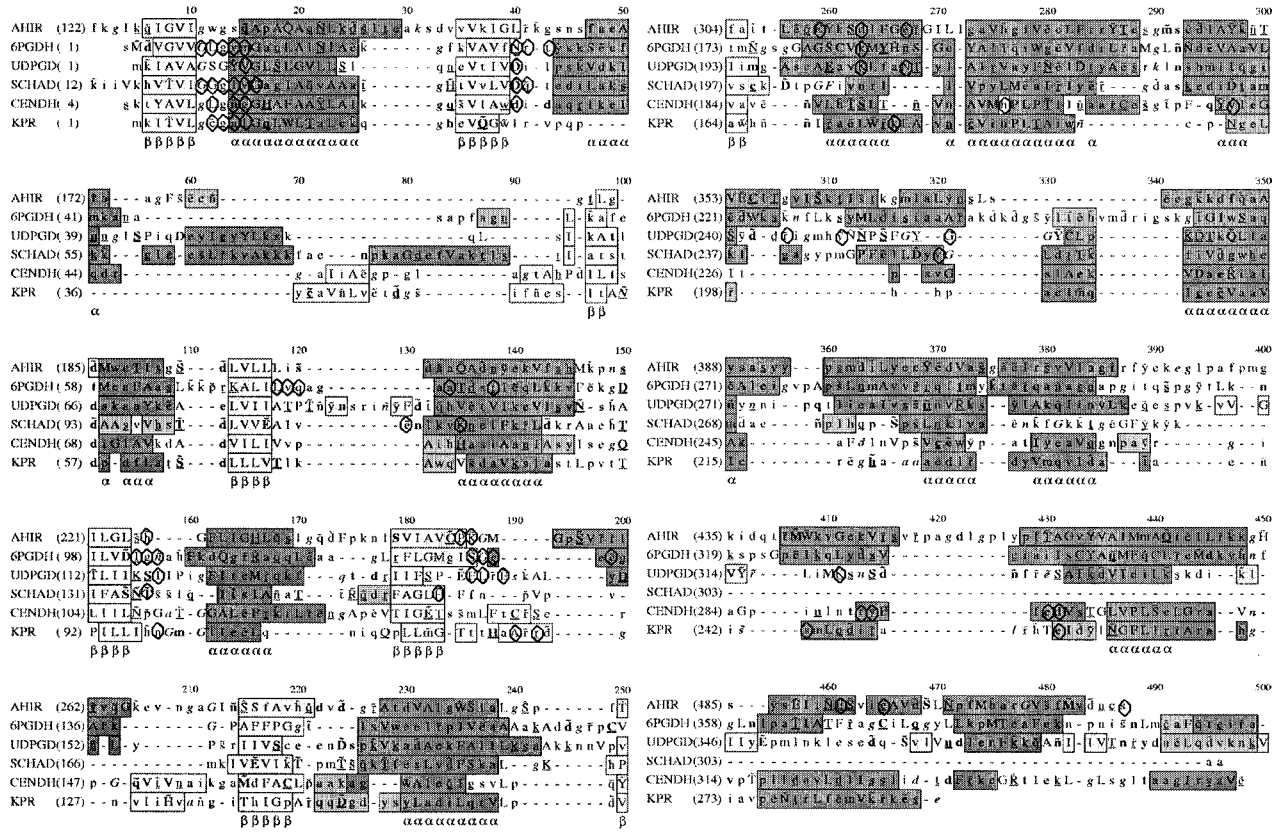


FIGURE 3: Sequence alignment of the 6-phosphogluconate dehydrogenase C-terminal domain-like superfamily formatted using the program JOY (*39*). The sequences are AHIR from spinach (*31, 32*), 6PGDH from *T. brucei* (*33*), UDPGDH from *St. pyrogenes* (*34*), SCHAD from humans (*35*), CENDH from *Arthrobacter* sp. (*36*), and KPR from *E. coli*. The formatting convention of JOY is as follows: darkest shading, α-helices; lightest shading, *β*-strands; uppercase letters, solvent inaccessible; lowercase letters, solvent accessible; bold type, hydrogen bonds to main chain amides; underlining, hydrogen bonds to main chain carbonyls; ∼, hydrogen bonds to other side chain and/or heterogen groups; italic, positive main chain torsion angles *ϕ*.

Table 2: Rmsd and Percentage Sequence Identity Values between KPR and Superfamily Members

superfamily member	rmsd	% identity of the whole molecule	% identity of N-termini	% identity of C-termini
AHIR	4.5	12.2	15.8	7.6
6PGDH	4.4	11.8	13.0	5.9
UDPGDH	4.1	11.4	17.0	9.6
SCHAD	4.4	9.0	12.5	2.5
CENDH	3.1	15.2	17.2	11.1

domain with a family-specific C-terminal extension with a common core formed around two long antiparallel helices related by (pseudo) 2-fold rotation.

A structural alignment of KPR and the five above members of the superfamily using the program COMPARER (*38*) and formatted using the JOY program (*39*) is shown in Figure 3. It reveals a relatively high degree of similarity in the N-terminal domain for the six proteins and a very low degree of similarity in the C-terminal domain. The level of identity and rmsd between KPR and AHIR, 6PGDH, UDPGDH, SCHAD, and CENDH for the whole molecule, the N-terminal domain, and the C-terminal domain are in Table 2. The alignment shows that CENDH is the protein most similar to KPR. The schematic representation of the secondary structure of SCHAD, KPR, and CENDH (Figure 4a), made with the help of the program HERA (*40*), shows the Rossmann fold in the N-terminal domain of all three proteins, but reveals that there are some differences. KPR has one fewer *β*-strands than the others at one edge of the sheet, whereas KPR and CENDH have an additional *β*-sheet.

The C-terminal domain in KPR is similar only to that in CENDH, where long helices in KPR correspond to several small helices in CENDH. In fact, a Dali search with this domain revealed no other proteins with a similar fold. Despite this high degree of structural similarity, KPR is a monomer whereas CENDH is a homodimer (subunit molecular weight of approximately 36 000) (*30*, *31*), with the α-helical C-terminal domain providing the dimerization interface. Both enzymes bind α-keto acids, and the residues that are proposed to bind the substrate in KPR (see below), Ser244 and Glu256, are very close to Arg292 and Glu296, the residues that bind the substrate in CENDH (Figure 4b). In Figure 3, the active site residues in the five superfamily members and the proposed active site residues in KPR have been highlighted with circles. The nucleotide binding site GXGXX[GA] is present in all the proteins, and proposed active site residues of KPR are in positions similar to those of active site residues in the other proteins; e.g., Asn98, Ala122, and Arg124 are aligned in all the proteins except CENDH.

Nucleotide Binding Site. The structural similarity between the N-terminal domains of KPR and human SCHAD enabled us to carry out molecular modeling to identify the cofactor binding site, since the structure of the binary complex of SCHAD with NAD^+ has been determined (PDB entry 2hdh). The SCHAD structure was fitted to the KPR structure and

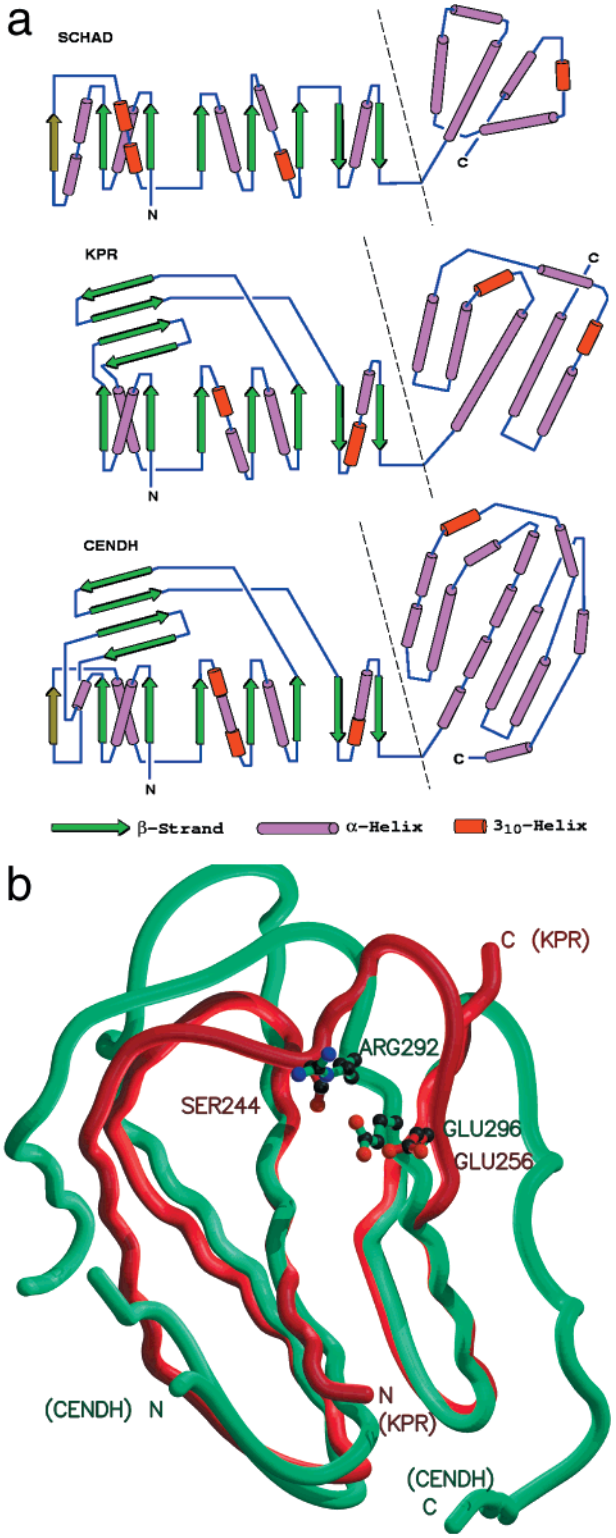


FIGURE 4: (a) Schematic representation of the secondary structure of SCHAD, KPR, and CENDH. β-Strands are represented as arrows and helices as cylinders. The secondary structure elements were defined using the program HERA (*40*). (b) Backbone representation of the C-terminal domain of KPR (red) and CENDH (green) showing the two active site residues.

the NAD^+ cofactor extracted from SCHAD to KPR. Water molecules clashing with the cofactor were removed and hydrogen atoms added to the model with the SYBYL Biopolymer module (Tripos, 1998) for force field optimization purposes. The NAD^+ molecule was edited to NADPH and optimized as a substructure using the Tripos force field.

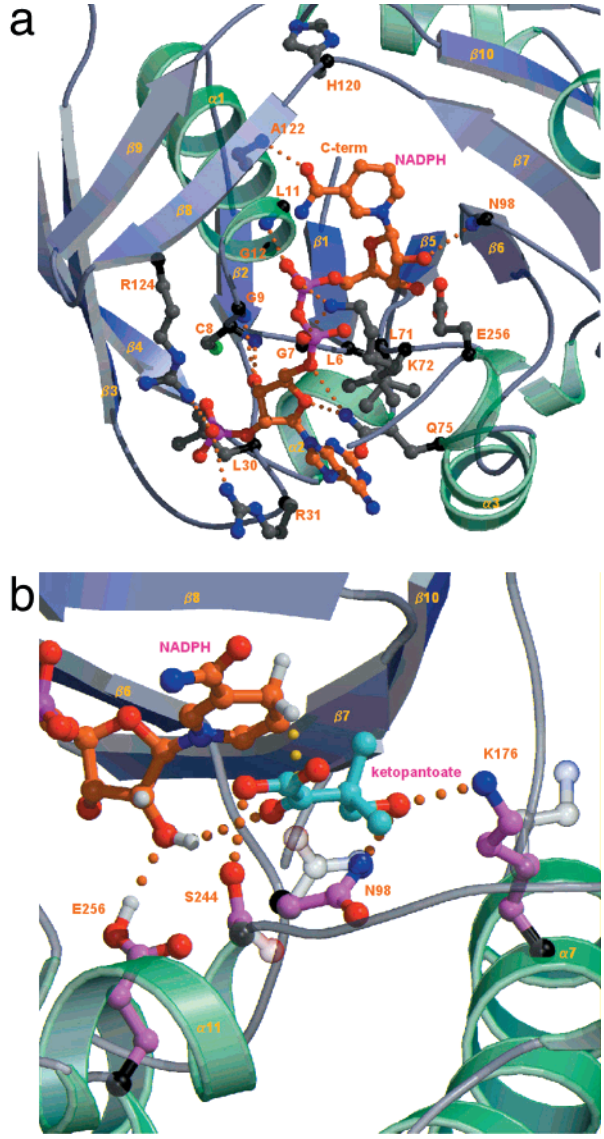


FIGURE 5: (a) Proposed model for interaction of the NADPH cofactor with KPR. Hydrogen bonding contacts are highlighted and the amino acids that are involved drawn in ball-and-stick form. (b) Proposed model for the interaction of ketopantoate with KPR. Hydrogen bonding contacts are highlighted and the amino acids that are involved drawn in ball-and-stick form. Where a side chain position has been adjusted, the original position is shown faintly.

As a result of this process, the conformation of the pyrophosphate group of NADPH changed so that it now fitted better into the interdomain cleft (Figure 5a). Only side chains involved in interactions with the cofactor moved during geometry optimization. In the resulting model, NADPH is largely buried between two domains, with only the adenine moiety of NADPH partially exposed, and the nicotinamide portion of the cofactor anchored in place by hydrogen bonds to its ribose hydroxyl by the conserved residues, the Glu256 Oϵ2 atom and the Asn98 main chain nitrogen atom (Figure 5b). In this model, the conserved residue Lys72 interacts with the pyrophosphate moiety through three hydrogen bonds, suggesting that this residue plays a major role in cofactor recognition. This observation is interesting, since Zheng and Blanchard (*9*) reported that a Lys72Ala mutant had wild-type enzyme activity in a crude cell free preparation, comparable to their Asp250Ala mutation, which is remote from the active site.

Substrate Binding Site and Reaction Mechanism. Although the apoenzyme is in a more open conformation than would be expected for catalysis, it is still possible to identify key catalytic residues and those involved in binding ketopantoate. In particular, Lys176 and Glu256, which have previously been identified as being important in catalysis by site-directed mutagenesis (*9*), are located near the putative NADPH site, some distance apart, and could accommodate ketopantoate between them. From knowledge of the stereochemical course of the reaction, in which the *proS* hydrogen of NADPH is delivered to the *si* face of ketopantoate (*8*), and when the expected geometry for hydride transfer is considered, it is possible to construct two models for the ternary complex. In one, the carbonyl group of ketopantoate points away from the nitrogen of the dihydronicotinamide ring, but this introduces unfavorable steric interactions between the gem dimethyl group on ketopantoate and the C-3 primary amide on the dehydronicotinamide ring. In the alternative structure (Figure 5b) (after some slight adjustment of the side chains of these residues and a larger movement of Lys176), there is the potential for good hydrogen bonding of Lys176 and Asn98 to the hydroxyl group of ketopantoate, and of Ser244 to the carboxyl. Furthermore, Glu256 could act as the acid to protonate the developing negative charge on ketopantoate as hydride is delivered from NADPH, albeit relayed through the C-3 hydroxyl of the ribose. All four of these residues are conserved.

CONCLUSION

KPR plays a key step in the biosynthesis of pantothenate, and as such is a potential herbicide and antimicrobial target. From the structure of the apoenzyme, it has been possible to construct models of the holoenzyme and the ternary complex, and suggest a catalytic mechanism. While the proposed structures appear to be reasonable, there are inherent uncertainties in any such models, and consequently, attempts to obtain crystal structures of these complexes are presently underway. Our results confirm the importance of Lys176 and Glu256 in binding ketopantoate and in the catalytic mechanism, although that the precise roles differ from those suggested by Zheng and Blanchard (*9*).

ACKNOWLEDGMENT

We thank Dr. T. Begley for the gift of the plasmid pVJS78 containing a region of *E. coli* genomic DNA encoding the *panE* gene.

REFERENCES

1. Cronan, J. E., Jr. (1980) *J. Bacteriol. 141*, 1291−1297.
2. Cronan, J. E., Jr., Little, K. J., and Jackowski, S. (1982) *J. Bacteriol. 149*, 916−922.
3. Manch, J. N. (1981) *Can. J. Microbiol. 27*, 1231−1233.
4. Primerano, D., and Burns, R. O. (1993) *J. Bacteriol. 153*, 259−269.
5. Frodyma, M. E., and Downs, D. (1998a) *J. Biol. Chem. 273*, 5572−5576.
6. Frodyma, M. E., and Downs, D. (1998b) *J. Bacteriol. 180*, 4757−4759.
7. Shimizu, S., Kataoka, M., Chung, M. C., and Yamada, H. (1988) *J. Biol. Chem. 263*, 12077−12084.
8. Zheng, R., and Blanchard, J. S. (2000) *Biochemistry 39*, 3708−37617.
9. Zheng, R., and Blanchard, J. S. (2000) *Biochemistry 39*, 16244−16251.
10. Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989) *Molecular cloning: a laboratory manual*, 2nd ed., Cold Spring Harbor Laboratory Press, Plainview, NY.
11. Laemmli, U. K. (1970) *Nature 227*, 680−685.
12. Gill, S. C., and von Hippel, P. H. (1989) *Anal. Biochem. 182*, 319−326.
13. van Duyne, G. D., Standaert, R. F., Karplus, P. A., Schreiber, S. L., and Clardy, Y. (1993) *J. Mol. Biol. 229*, 105−124.
14. Doublie, S. (1997) *Methods Enzymol. 278*, 523−530.
15. Otwinowski, Z., and Minor, W. (1997) *Methods Enzymol. 276*, 307−326.
16. Collaborative Computational Project, Number 4 (1994) *Acta Crystallogr. D50*, 760−763.
17. Evans, P. R. (1997) in *Recent Advances in Phasing* (Wilson, K. S., Davies, G., Ashton, A. W., and Bailey, S., Eds.) pp 97−102, Council for the Central Laboratory of the Research Councils, Daresbury Laboratory, Daresbury, U.K.
18. Sheldrick, M. (1996) *SHELXS*96. *Program for the Solution of Crystal Structures*, University of Göttingen, Göttingen, Germany.
19. La Fortelle, E. D., Irwin, J. J., and Bricogne, G. (1997) in *CCP4 study weekend: Recent Advances in Phasing* (Wilson, K. S., Davies, G., Ashton, A. W., and Bailey, S., Eds.) Council for the Central Laboratory of the Research Councils, Daresbury Laboratory, Daresbury, U.K.
20. Jones, T. A., Zou, J. Y., Cowan, S. W., and Kjeldgaard, M. (1991) *Acta Crystallogr. A47*, 110−119.
21. Murshudov, G. N., Vagin, A. A., and Dodson, E. J. (1997) *Acta Crystallogr. D53*, 240−255.
22. Brunger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J. S., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T., and Warren, G. L. (1998) *Acta Crystallogr. D54*, 905−921.
23. Kleywegt, G. J., and Jones, T. A. (1996) *Acta Crystallogr. D52*, 826−828.
24. Perrakis, A., Morris, R., and Lamzin, V. S. (1999) *Nat. Struct. Biol. 6*, 458−463.
25. Laskowski, R. A., Macarthur, M. W., Moss, D. S., and Thornton, J. M. (1993) *J. Appl. Crystallogr. 26*, 283−291.
26. Kabsch, W., and Sander, C. (1983) *Biopolymers 22*, 2577−2637.
27. Branden, C., and Tooze, J. (1991) *Introduction to protein structure*, Garland Publishing, New York.
28. The following sequences were used (with Genbank accession numbers): *Aquifex aeolicus* (O67619), *Archaeoglobus fulgidus* (AAB89553), *Bacillus halodurans* (BAB06298), *Bacillus subtilis* (O34661), *E. coli* (AAC73528), *Halobacterium* sp. NRC-1 (AAG19208), *Lactococcus lactis* (AAK05421), *Pseudomonas aeruginosa* (E83096), *Pyrococcus abyssi* (H75119), *S. typhimurium* (P37402), *Staphylococcus epidermis* (AR269609), *St. pyogenes* (AAK33780), and *Vibrio cholerae* (D82092).
29. Holm, L., and Sander, C. (1998) *Nucleic Acids Res. 26*, 316−319.
30. Holm, L., de Daruvar, A., Sander, C., and Dodge, C. (1995) The Dali Server (http://www2.ebi.ac.uk/dali/).
31. Biou, V., Dumas, R., Cohen-Addad, C., Douce, R., Job, D., and Pebay-Peyroula, E. (1997) *EMBO J. 16*, 3405−3415.
32. Thomazeau, K., Dumas, R., Halgand, F., Forest, E., Douce, R., and Biou, V. (2000) *Acta Crystallogr. D56*, 389−397.
33. Phillips, C., Dohnalek, J., Gover, S., Barrett, M. P., and Adams, M. J. (1998) *J. Mol. Biol. 282*, 667−681.

34. Campbell, R. E., Mosimann, S. C., van de Rijn, I., Tanner, M. E., and Strynadka, N. C. J. (1994) *Acta Crystallogr. D50*, 760−763.
35. Barycki, J. J., O'Brien, L. K., Bratt, J. M., Zhang, R., Sanishvili, R., Strauss, A. W., and Banaszak, L. J. (1999) *Biochemistry 38*, 5786−5798.
36. Britton, K. L., Asano, Y., and Rice, D. W. (1998) *Nat. Struct. Biol. 5*, 593−601.
37. Murzin, A. G., Brenner, S. E., Hubbard, T., and Chothia, C. (1995) *J. Mol. Biol. 247*, 536−540.
38. Sali, A., and Blundell, T. L. (1990) *J. Mol. Biol. 212*, 203−428.
39. Mizuguchi, K., Deane, C. M., Blundell, T. L., Johnson, M. S., and Overington, J. P. (1998) *Bioinformatics 14*, 617−623.
40. Hutchinson, E. G., and Thornton, J. M. (1990) *Proteins 8*, 203−212.
41. Kraulis, P. J. (1991) *J. Appl. Crystallogr. 24*, 946−950.
42. Bacon, D., and Anderson, W. F. (1988) *J. Mol. Graphics 6*, 219−220.
43. Merritt, E. A., and Murphy, M. (1994) *Acta Crystallogr. D50*, 869−873.

BI011020W