Crystal Structures of Human Pantothenate Kinases

INSIGHTS INTO ALLOSTERIC REGULATION AND MUTATIONS LINKED TO A NEURODEGENERATION DISORDER

Received for publication, March 5, 2007, and in revised form, July 12, 2007 Published, JBC Papers in Press, July 13, 2007, DOI 10.1074/jbc.M701915200

Bum Soo Hong§, Guillermo Senisterra³, Wael M. Rabeh⁷, Masoud Vedadi⁴, Roberta Leonardi⁵, Yong-Mei Zhang⁵, Charles O. Rock⁵, Suzanne Jackowski⁵, and Hee-Won Park*†

From the ⁴Structural Genomics Consortium and ⁵Department of Pharmacology, University of Toronto, Toronto, Ontario M5G 1L5, Canada and the ⁶Department of Infectious Diseases, St. Jude Children’s Research Hospital, Memphis, Tennessee 38105-2794

Pantothenate kinase (PanK) catalyzes the first step in CoA biosynthesis and there are three human genes that express four isoforms with highly conserved catalytic core domains. Here we report the homodimeric structures of the catalytic cores of PanK1α and PanK3 in complex with acetyl-CoA, a feedback inhibitor. Each monomer adopts a fold of the actin kinase superfamily and the inhibitor-bound structures explain the basis for the allosteric regulation by CoA thioesters. These structures also provide an opportunity to investigate the structural effects of the PanK2 mutations that have been implicated in neurodegeneration. Biochemical and thermodynamic analyses of the PanK3 mutant proteins corresponding to PanK2 mutations show that mutant proteins with compromised activities and/or stabilities correlate with a higher incidence of the early onset of disease.

Pantothenate kinase (PanK)² is the first and rate-limiting enzyme in the CoA biosynthetic pathway, which catalyzes the phosphorylation of vitamin B₅, pantothenate (1). In human, there are three genes that express four catalytically active PanK isoforms. PanK1α and PanK1β arise from the PANK1 gene through the use of alternate initiating exons (2). PANK2 encodes a protein that is processed twice and localizes to the mitochondria (3, 4) and PANK3 encodes a single cytosolic isoform (5). All of these PanK isoforms share a common, highly homologous catalytic core that is composed of about 355 residues (e.g. the identity of ~83%). PanK1β and PanK3 consist of the catalytic core alone, whereas PanK1α has a feedback regulatory domain in an amino-terminal extension (2). The NH₂-terminal domain of PanK2 includes two mitochondrial targeting signals, which are removed proteolytically concomitant with its mitochondrial localization (3, 6). In addition, PANK4 encodes a protein with sequence similarity to the other PanKs, but it lacks an essential catalytic glutamate for enzyme activity and has a long carboxyl-terminal extension of unknown function (3).

Mutations in PanK2 have been linked to the inherited pantothenate kinase-associated neurodegeneration (PKAN) (7), which constitutes a subset of neurodegeneration with brain iron accumulation, formerly known as Hallervorden-Spatz syndrome (8). Patients with PKAN exhibit motor symptoms, such as dystonia, dysarthria, intellectual impairment, and gait disturbance (9) and early onset patients have rapidly progressive disease, whereas later onset patients have a slowly progressive, atypical disease (9, 10). The reason for brain iron accumulation in PKAN is still unknown. A current hypothesis of PKAN pathogenesis states that PANK2 mutations lead to a loss of PanK2 function and reduced phosphopantothenate, resulting in the accumulation of cysteine-containing molecules and CoA deficiency (6, 7). These cysteine-containing substrates may undergo rapid autoxidation in the presence of iron, leading to free radical generation and cell damage (6, 7). PANK2 mutations are clustered in the catalytic core domain, and insertion, deletion, and frameshift mutations in the PANK2 gene are linked clinically to early onset disease (10), indicating that null mutants of PanK2 lead to rapidly progressive PKAN. In contrast, the missense mutations have mixed early and late onset patterns (10), and a set of missense mutant proteins expressed in transiently transfected cells are active (4, 11), suggesting that defects unrelated to catalysis contribute to PKAN disease.

A common feature of the PanK isoforms is that they are feedback inhibited by CoA and its thioesters. This regulation is a primary mechanism that controls PanK activity and hence the intracellular CoA level (2, 5, 12). Recent biochemical data show that acetyl-CoA is a more potent feedback inhibitor than free CoA for mouse PanKs (5, 11). Structural studies on PanK have been reported only in bacterial orthologs (13–17). The crystal structures of Escherichia coli PanK reveal why its activity,
Structures of Human Pantothenate Kinases

unlike mouse PanKs, is more potently inhibited by free CoA, (13, 14). Staphylococcus aureus PanK (SaPanK) is the most similar in sequence to the mammalian PanK catalytic core (e.g. the identity of 23%), but is refractory to feedback regulation by CoA (17, 18). Due to the low sequence homology, the structural interpretation of the disease-causing human PanK2 mutations using the SaPanK model is limited to three missense mutations that directly affect the active site (17).

We have determined the homodimeric structures of the catalytic cores of human PanK1α and PanK3 in complex with acetyl-CoA. These crystallographic data provide the first look at the structures of metazoan PanKs and their allosteric regulation. The high degree of similarity between human PanK3 and PanK2 allows an analysis of the mutations associated with PKAN. PanK2 missense mutations mapped onto the active site, dimerization interface, and protein interior tend to be associated with the early onset of PKAN phenotypes and their corresponding PanK3 mutant proteins show PanK inactivation and destabilization. On the other hand, mutations of surface residues have less severe effects on enzyme function and correlate with the mixed early and late onsets of disease.

EXPERIMENTAL PROCEDURES

Protein Expression and Purification—The genes encoding the catalytic core domains of human PanK1α (residues 231–597, abbreviated PanK1 below) and PanK3 (residues 12–368) were amplified using PCR and cloned into the expression vector pET28a to obtain N-terminal His6 tag fusion proteins. E. coli BL21(DE3) (Stratagene) was transformed with the resulting plasmids and both recombinant PanK1 and PanK3 proteins were overexpressed upon 1.0 mM isopropyl 1-thio-β-D-galactopyranoside induction for 18 h at 18 °C. The proteins were purified using nickel-nitrilotriacetic acid affinity and Superdex 200 (GE Healthcare) gel-filtration columns. Purified PanK1 and PanK3 samples were concentrated to 25 and 30 mg/ml respectively, and stored at ~80 °C in the gel filtration buffer (20 mM Tris-HCl, pH 8.0, 200 mM NaCl, 10 mM dithiothreitol).

Crystallization—PanK proteins were crystallized using the sitting drop vapor diffusion method at 18 °C. To achieve full occupancy of bound ligands, the proteins were mixed with 5-fold molar excess of an ATP analogue AMP-PNP or acetyl-CoA. The mixtures were incubated overnight at 4 °C. Equal volumes of protein solution and reservoir solution were combined for crystallization. Crystals of PanK1 were grown at 0.1M Tris-HCl (pH 8.5), 16% PEG3350, 0.2M CaCl2, and 0.02M Tris(2-carboxyethyl) phosphine hydrochloride-HCl, and PanK3 was crystallized in 0.1M Tris-HCl, pH 8.5, 28% PEG400, 0.2M CaCl2, and 0.02M tris(2-carboxyethyl) phosphine hydrochloride-HCl, and PanK3 was crystallized in 0.1M Tris-HCl (pH 8.5), 16% PEG3350, 0.2M ammonium citrate, and 0.02M octaethyleneglycol dodecyl ether (C12E8). The mercury derivatives were obtained by soaking PanK1 crystals in the reservoir solution with 5 mM ethylmercury thiosalicylate overnight. All crystals were cryoprotected in the 50:50 mixture of Paratone-N and mineral oil, and flash-frozen in liquid nitrogen. The diffraction quality crystals of PanK1 were obtained in the presence of acetyl-CoA. Although AMP-PNP was added for the crystallization of PanK3, we found acetyl-CoA molecules in the structure that co-purified with PanK3.

Data Collection and Structure Determination—Diffraction data were collected on beamline 19-BM at the Advanced Photon Source, running at 100 K and processed using the HKL2000 program package (19). The program SOLVE (20) was used to locate four mercury sites and to calculate initial phases of PanK1 and the program RESOLVE (21) to improve the initial phases. The model of PanK1 was completed in several repetitions of manual building with the program O (22). The structure of PanK3 was solved by the molecular replacement method using the program MOLREP in the CCP4 package (23) and the PanK1 structure as a search model, and built using the program O (22). Finally, both PanK structures were refined with the program REFMAC5 (24). The crystals of PanK1 contained a dimer in the asymmetric unit and the two monomers were related by noncrystallographic 2-fold symmetry. In PanK3, two dimers were in the asymmetric unit. A dimer is the most common functional organization for PanK proteins (4, 5, 25). Data collection and refinement statistics are given in Table 1.

Site-directed Mutagenesis of PanK3—Based on the PanK2 mutations identified in PKAN patients (10), site-directed mutagenesis in PanK3 was accomplished using the Stratagene QuikChange kit (Stratagene). Prior to protein expression, the sequences of all mutations were verified by sequencing, and each mutant protein was purified as described for the wild-type protein, except the concentration of dithiothreitol was 1.0 mM.

PanK Activity Assay—The effect of acetyl-CoA on the PanK1α and PanK3 Km values for ATP was investigated by using the assay previously described (2, 12). The standard reaction mixture contained 180 μM D-[1-14C]pantothenate (Amer sham Biosciences), 10 mM MgCl2, 0.1M Tris-HCl (pH 7.5), 200 ng of PanK1α or 100 ng of PanK3 and increasing concentrations of ATP and acetyl-CoA, as indicated. The mixture was incubated at 37 °C for 10 min and the radioactive product was quantitated by scintillation counting as described previously (11). Activities of PanK3 mutant proteins were compared with the wild-type using the standard pantothenate kinase assay (11). Briefly, the reaction mixture contained 90 μM D-[1-14C]pantothenate, 2.5 mM ATP, 10 mM MgCl2, 0.1M Tris-HCl (pH 7.5), and 100 ng of purified protein, and assayed as described above. Relative activities of the mutant proteins were checked three times and these data were combined.

UV-visible Spectra and Mass Spectrometry—UV-visible spectra of PanK3 wild-type and mutant proteins (1 mg/ml in 10 mM Tris-HCl (pH 8.0), 50 mM NaCl, 0.5 mM EDTA, 0.5 mM dithiothreitol, and 50% (v/v) glycerol) were recorded as described previously (27). The supernatant (100 μl) from a heat-denatured sample of PanK3 (1.8 mg/ml in 50 mM ammonium acetate, pH 7.0) was desalted on a C-8 MicroTip column (Harvard Apparatus), washed with water, and the protein-bound small molecules were eluted with methanol (100 μl), dried under N2 flow, and resuspended in methanol:water (1:1, v/v, 10 μl). Mass spectrometry analysis was performed using a Finnigan™ TSQ® Quantum™ (Thermo Electron Corporation, San Jose, CA) triple quadrupole mass spectrometer equipped with the nanospray ion source.

Stability Studies by Differential Static Light Scattering—The thermostabilities of PanK3 and its mutant proteins were studied using differential static light scattering (28, 29), monitoring protein stability by its aggregation properties. Protein samples at 0.2 mg/ml were heated from 25 to 80 °C at a rate of 1 °C/min
Structures of Human Pantothenate Kinases

TABLE 1
Crystallographic data and refinement statistics

| Crystal   | Hg for MAD | Native | PANK3, native |
|-----------|------------|--------|---------------|
| Space group | P6₁     | P6₁    | P6₁           |
| Unit cell parameters (Å) | a = b = 93.4  | a = b = 93.5  | a = b = 92.9  |
|                | c = 198.0  | c = 198.3  | c = 197.7     |
| Wavelength (Å) | 1.00510 | 1.00900  | 1.54178       |
| Resolution range (Å) | 99.7 (100) | 99.7 (100) | 99.5 (98.9)   |
| Completeness (%) | 23.04,183 | 23.60,276 | 1.66,635      |
| Total no. of reflections | 33,600 | 33,656  | 75,175        |
| No. of unique reflections | 7.6 (10.5) | 6.7 (12.5) | 6.7 (61)      |
| Rwork (%) | 50–1.9 | 22.2 (24) | 24.3 (31.8)   |
| Rfree (%) | 50–2.05 | 21.1 (27.8) | 25.4 (35.8)   |
| Completeness (%) | 50–2.5 | 50–1.9 | 50–2.05       |
| No. of atoms | 99.7 (100) | 99.7 (100) | 99.5 (98.9)   |
| Protein | 5,143 | 10,967  | 96,048        |
| Water | 58 | 116 | 92 |
| Root mean square deviations from ideality | 0.007 | 0.006 | 0.15 |
| Bond lengths (Å) | 1.5 | 1.2 |
| Ramachandran plot | 90.1 | 90.5 |
| Additional allowed (%) | 9.2 | 8.9 |
| Generously allowed (%) | 0.7 | 0.6 |
| Completeness (%) | 50–2.5 | 50–1.9 | 50–2.05       |
| No. of atoms | 99.7 (100) | 99.7 (100) | 99.5 (98.9)   |
| Protein | 5,143 | 10,967  | 96,048        |
| Water | 58 | 116 | 92 |
| Root mean square deviations from ideality | 0.007 | 0.006 | 0.15 |
| Bond lengths (Å) | 1.5 | 1.2 |
| Ramachandran plot | 90.1 | 90.5 |
| Additional allowed (%) | 9.2 | 8.9 |
| Generously allowed (%) | 0.7 | 0.6 |

* Values in parentheses are for the highest resolution shell.
* Rwork = Σ [ ||Fobs|| − ||Fc|| ]/Σ ||Fobs||, where ||Fobs|| and ||Fc|| are observed and calculated structure factor amplitudes.
* The Rfree value was calculated with a random 5% subset of all reflections excluded from refinement.

in clear bottom 384-well plates (Nunc) in 50 μL of buffer (100 mm Hepes, 150 mm NaCl, pH 7.5). Protein aggregation was measured by recording the scattered light using a CCD camera and taking images of the plate every 0.5 °C. The pixels intensities in a preselected region of each well were integrated using proprietary software to generate a value representative of the total amount of scattered light in that region. These intensities were plotted against temperature for each sample well and fitted to Equation 1 as described by Matulis et al. (30),

\[
I = I_f + \left( \frac{1}{1 + e^{-(\Delta H/R)/(1/T - 1/T_{agg})}} \right) \left( I_o - I_f \right)
\]

(Eq. 1)

where \(I_o\) and \(I_f\) are the intensities of light scattered by the protein before and after denaturation, \(\Delta H\) is the enthalpy of unfolding, \(R\) is the general gas constant, \(T_{agg}\) is the temperature of aggregation that represents the melting temperature, and \(\Delta C_p\) is the heat capacity difference between the unfolded and folded states of the protein. \(\Delta C_p\) of 4.93 kcal/mol was calculated based on the protein sequence and was used in all subsequence calculations (31). The free energy difference \(\Delta \Delta G\) between the wild-type and each mutant protein at the temperature of aggregation of the wild-type was calculated using Equation 2,

\[
\Delta \Delta G_{m(T_wt)} = \Delta H_m - [T_{agg}/T_m] \Delta H_m
\]

(Eq. 2)

where \(\Delta \Delta G_{m(T_wt)}\) is the denaturation free energy difference between the wild-type and mutant PanK3, \(T_{agg}\) and \(T_m\) are the aggregation temperatures of the wild-type and mutant proteins, respectively.

Stability Studies by Isothermal Aggregation (ITA)—Isothermal aggregation was measured at 37 °C. Protein samples at 0.2 mg/ml were incubated at 37 °C in clear bottom 384-well plates (Nunc) in 50 μL of buffer (100 mM Hepes, 150 mM NaCl, pH 7.5). Protein aggregation was monitored by detecting the scattered light by taking images every 30 s using a CCD camera for 2 h. Images were analyzed and the intensities were plotted versus time. All curves were normalized to compensate for small differences in initial intensities. The time it takes to have 50% of the protein in the aggregation state, here on, is termed ITA_{50}. Each experiment was repeated 5 times and the average has been reported here.

RESULTS

Overall Structure Description

The PanK1 and PanK3 proteins we crystallized represent the catalytic core domain, which is highly conserved in all human PanK isoforms. The identity between the catalytic core domains of the two isoforms is \(\sim 84\%\) (ClustalW 1.7). The overall monomer fold places the human PanKs in the actin fold protein kinase superfamily (32), and contains two domains (Fig. 1A). The A domain consists of a six-stranded \(\beta\)-sheet (strands Ab1–Ab6) flanked by five helices (Ah1–Ah3 on one side and Ah4 and Ah11 on the other side) and a glycine-rich loop, characteristic of many nucleotide-binding sites (33, 34), lies
Structures of Human Pantothenate Kinases

FIGURE 1. Overall view of the structure of human PanK3. A, stereo view of a ribbon representation of the PanK3 monomer structure. The secondary structure elements (h for helices and b for sheets) start with prefixes A and B to indicate A and B domains, respectively. The extended loop region and the predicted ATP binding site (arrow) are indicated. B, ribbon representation of the PanK3 dimer structure. Two molecules in the PanK3 dimer are in green and yellow, respectively. Acetyl-CoA is in ball-and-stick representation. Helix Bh9 of the B domain of the first monomer and helix Bh9' of the B' domain of the second monomer that form the dimerization interface are labeled. All the figures were generated by PyMol (pymol.sourceforge.net).

The major portion of the dimerization interface is composed of two long helices, Bh9 and Bh9' from each monomer, which interact with each other in parallel resulting in a V-shaped conformation (Fig. 1B and supplemental Fig. S2A). The extended loop of the second monomer contributing to the stabilization of the PanK dimer (supplemental Fig. S2B).

Dimerization Interface

The major portion of the dimerization interface is composed of two long helices, Bh9 and Bh9' from each monomer, which interact with each other in parallel resulting in a V-shaped conformation (Fig. 1B and supplemental Fig. S2A). Helices Bh9 and Bh9' widen at their NH2-terminal ends and form a four-helix bundle with two short helices Bh5 and Bh5' (supplemental Fig. S2A). The center of the four-helix bundle is lined by non-polar residues forming a compact hydrophobic core. The narrow COOH-terminal ends of helices Bh9 and Bh9' are held together by van der Waals interactions between Val304 and Met307 of each monomer, and are capped by Asn311 and Asn311', which face each other. The middle of helices Bh9 and Bh9' is attached by both direct and water-mediated hydrogen bonds between Asn300 and Asn300'. Additional van der Waals interactions exist between Bh5 of one monomer and the extended loop of the second monomer contributing to the stabilization of the PanK dimer (supplemental Fig. S2B).

The Acetyl-CoA Binding Site

In both PanK1 and PanK3 structures, the acetyl pantetheine diphosphate moieties of acetyl-CoA are bound parallel to the β strands of the B domain of one monomer and covered on the other side by the extended loop of the second monomer (Fig. 2A). The main interactions of acetyl-CoA with PanKs occur via the acetyldiphosphopantetheine moiety (Fig. 2B). The 3'-phosphoadenonsine moiety is poorly defined in the electron density map, suggesting that it is disordered and thus not included in the final PanK structures. Residues lining the binding site of the acetyldiphosphopantetheine are identical in both the PanK structures, with the exception of residues 266 and 340 involved in hydrophobic interactions (Fig. 2B). The acetyldiphosphopantetheine moiety forms direct or water-mediated hydrogen bonds and van der Waals interactions with the protein core. The polar residues involved in hydrogen bonds are from the β strands of one monomer (Bb9–Bb11) and the non-polar residues involved in van der Waals interactions are contributed by the extended loop and Bh10' of the second monomer (Fig. 2A). Acetyl-CoA was very tightly bound to PanK3 in the structure, consistent with the low IC50 value for acetyl-CoA inhibition of the enzyme (5).

Predicted Binding Sites for ATP and Pantothenate

The ATP binding site of PanK3 was modeled on the basis of its structural similarity to the SaPanK·AMP·PNP complex (Protein Data Bank code 2EWS) (17) and the location of the conserved glycine-rich loop (Fig. 3A). Superposition of the B domain of PanK3 and the corresponding domain of the SaPanK·AMP·PNP complex placed the ATP molecule between the two domains of the human PanKs. The two domains of PanK3 and also PanK1 were in an “open” conformation com-
Structures of Human Pantothenate Kinases

FIGURE 2. Acetyl-CoA interactions with human PanK3. A, stereo view of a F$_{o}$ — F$_{c}$ difference map of acetyl-CoA (the acetyl-pantetheine moiety, see “Results”) shown in black stick representation. The difference map was calculated before introducing the molecule acetyl-CoA and contoured at 2.5 a. Two PanK3 monomers forming the acetyl-CoA binding site are in green and yellow, respectively. A single prime superscript indicates the secondary structure elements from the second monomer. B, schematic representation of the acetyl-CoA interactions with PanK3. Blue dotted lines indicate hydrogen bonds, whereas red dotted semicircles mark hydrophobic interactions. The residues that form the hydrogen bonds to acetyl-CoA are boxed. S and M superscripts indicate the acetyl-CoA interactions with the side chain and main chain of residues, respectively. One water molecule (W) is displayed in the red sphere. The sequence differences between PanK3 and PanK1 are given in parentheses (e.g. Trp$_{266}$ and Tyr$_{330}$ of PanK3 are replaced by serine and phenylalanine in PanK1, respectively).

pared with the “closed ” conformation SaPanK-AMP-PNP (Fig. 3A). The binding of the acetyl-CoA kept PanK in the open conformation, characteristic of the ligand-free structures of actin-fold proteins (32). Acetyl-CoA was a competitive inhibitor of both PanK1 and PanK3 (Fig. 3, B and C), suggesting that the binding of acetyl-CoA interfered with the ATP binding by stabilizing the open conformation. Based on the superimposed structures of PanK3-acetyl-CoA and SaPanK-AMP-PNP, the $\alpha,\beta$-phosphates of acetyl-CoA and the $\beta,\gamma$-phosphates of ATP occupied the same space (Fig. 3A), explaining the competitive nature of acetyl-CoA and ATP for PanK. Also, we found that a high proportion of purified PanK3 was bound to acetyl-CoA. Heat denaturation of the protein released the bound compound into the supernatant, and analysis of this supernatant by mass spectrometry identified it as acetyl-CoA (Fig. 3D). This indicates that acetyl-CoA remained bound to PanK3 during purification.

The Pan binding site was predicted on the basis of the location of bound acetyl-CoA in PanK3. The same residues lining the binding pocket for the Pan moiety of acetyl-CoA (Fig. 2B) were considered for the binding of Pan. A key residue was Arg$^{207}$ that formed a salt bridge to the carboxamide group of the Pan moiety of acetyl-CoA (the carboxyl group of Pan) (Fig. 2B). This salt bridge interaction would anchor one end of Pan. A similar salt bridge interaction was observed in the Pseudomonas aeruginosa PanK structure (17). However, Glu$^{138}$, another key residue, predicted to act as a general base for the nucleophilic activation of the hydroxyl group of Pan in SaPanK (17), was not located adjacent to the 4’-hydroxy group of Pan (data not shown), suggesting that the enzyme exists in a catalytically inactive state. This finding is consistent with the idea that the acetyl-CoA-bound PanK3 structure adopts the open conformation (see above). Therefore, a switch to the closed conformation by the binding to ATP and/or Pan is required for the nucleophilic activation of the OH group of Pan by Glu$^{138}$.

Crystallographic Mapping of Missense Mutations Associated with PKAN Disease

The sequence similarity in the catalytic core domains of PanK3 and PanK2 allowed mapping of the PanK2 missense mutations (10) onto the PanK3 structure (Fig. 4A) and we correlated the locations of these mutations with their biochemical and disease phenotypes (Table 2). The mutations were grouped into three categories.

Active Site Mutations—Two PanK2 mutations (G219V and G521R) mapped to the ATP binding site and of these, G521R is most frequently observed in PKAN patients (10). Gly$^{321}$ (Gly$^{221}$ of PanK2) was located in the loop between Bh10 and Bh14 on one side of the ATP binding site, whereas Gly$^{138}$ (Gly$^{219}$ of PanK2) formed part of the glycine-rich loop (supplemental Fig. S3A). These glycine residues create space for the docking of the ATP phosphates, and the introduction of a side chain by mutation, especially a bulky one, was predicted to prevent proper ATP binding. A previous study of both mutations in PanK2 has shown that these mutant proteins were catalytically inactive, and the PanK2-G521R failed to properly fold (11).

Mutations at the Dimer Interface and Protein Interior—The PanK3 structure placed eight PanK2 missense mutations (L413P, D447N, S471N, I497T, N500I, I501T, A509V, and
N511D) at the dimer interface (Fig. 4A). Of these, Leu213 (Leu413 of PanK2) made van der Waals interactions with non-polar residues (supplemental Fig. S3B), and Asp247, Ser271, Asn300, and Asn311 (Asp447, Ser471, Asn500, and Asn511 of PanK2) formed direct or water-mediated hydrogen bonds to helix Bb9/H11032 of the second monomer (supplemental Figs. S2A and S3B, C). On the other hand, Ile297, Ile301, and Ala309 (Ile497, Ile501, and Ala509 of PanK2) were indirectly involved in PanK3 dimerization. Many water molecules were trapped at the dimer interface and formed a network through the interaction with amino acid residues along the dimer interface (supplemental Fig. S3B, C). This network, conserved in both PanK1 and PanK3 structures, is considered to tighten the dimerization of two monomers and the mutations at Asp247, Ser271, and Asn300 of PanK3 were predicted to perturb the ordered water molecules. Three PanK2 mutations, L282V, A398T, and L563P, were located in the protein interior (Fig. 4A) and the corresponding residues (Leu82, Ala198, and Leu363) in PanK3 were involved in the hydrophobic interactions with surrounding non-polar residues.

Mutations on the Surface of PanK—Fifteen PanK2 mutations were located on the surface of the PanK3 structure with their side chains exposed to solvent (Fig. 4B). Most notable was PanK2-(T528M), the second most common mutation in PKAN disease (10, 35). Thr528 of PanK2 was equivalent to Thr328 of PanK3 and Met553 of PanK1, which was located at the NH2 terminus of helix Bh10 (supplemental Fig. S3D). Based on our sequence alignment (data not shown) and both PanK structures, the side chain hydroxyl group of Thr528 in PanK2 was predicted to form a hydrogen bond to Arg532 and PanK2-(T528M) was predicted to break this interaction as shown in wild-type PanK1. The side chain of Lys332 in PanK3 formed water-mediated hydrogen bonds to residues in the loop between Bb7 and Bb8 (supplemental Fig. S3D), which could be achieved by the corresponding Arg532 in PanK2. Interestingly, this arginine mutation, PanK2-(R532W), was found in PKAN patients, supporting the idea that the Thr528–Arg532 interaction is important in positioning the Bb7–Bb8 loop for some physiological function of PanK2.

Mutations at Arg64, Arg78, Arg86, Ser151, and Asn155 of PanK3 (i.e. PanK2-(R264W), -(R278L or C), -(R286C/S351P), and -(N355S)) were predicted to disrupt salt bridges or hydrogen bonds with surrounding residues, also conserved between PanK3 and PanK2 (data not shown). Cys228 of PanK3 (Cys428 of PanK2) was located on the protein surface but its side chain pointed to the protein core (Fig. 4B). Finally, the PanK3 structure indicated that the side chains of five residues, Thr434, Arg499, Arg512, Thr527, and Asn540, corresponding to six PanK2 mutations, T234A, R249P, E322G or D, T327I, and N404I, were exposed to solvent with no other apparent interactions.

Biochemical and Thermodynamic Characterization of PanK3 Mutant Proteins

We cloned, expressed, and purified 27 PanK3 mutant proteins that correspond to the PanK2 missense mutations described above. The effects of these mutations on activity and stability of the protein were investigated using in vitro PanK...
activity and thermostability assays (Fig. 4C, Table 2). The mutations located at the ATP binding site severely affected the catalytic activity and stability of the protein. PanK3-(G19V) (PanK2-(G219V)) had negligible activity (~1%) and its $T_{\text{agg}}$ (36.7 °C) was lower than the wild-type (45.4 °C) with a significant destabilization energy of $-2.46 \text{ Kcal/mol}$ (Table 2). PanK3-(G321R)(PanK2-(G521R)) was insoluble and could not be studied with either technique.

Of the dimer interface mutant proteins, L213P (PanK2-(L413P)) and N300I (PanK2-(N500I)), eluted as monomers by gel filtration chromatography, showed low $T_{\text{agg}}$ values (36.9°C and 36.6 °C) with $\Delta G$ values of $-3.32$ and $-1.84 \text{ Kcal/mol}$, respectively, and they also exhibited compromised catalytic activities (Table 2). Three other PanK3 mutations along the dimer interface, D247N (PanK2-(D447N)), S271N (PanK2-(S471N)), and I297T (PanK2-(I497T)) did not affect dimer assembly but significantly altered their enzymatic activities and thermostabilities. The other dimer interface mutant proteins, I301T (PanK2-(I501T)), A309V (PanK2-(A509V)), and N311D (PanK2-(N511D)) eluted as dimers and showed moderate protein destabilizations, and no significant changes in their activities relative to the wild-type (Fig. 4C). A mutation in the interior of PanK3, L363P (PanK2-(L563P)), exhibited complete insolubility in protein purification trial, whereas the other interior mutant protein, L82V (PanK2-(L282V)) was soluble and displayed moderate destabilization with catalytic activity similar to the wild-type (Fig. 4C).

Most of the surface mutations exhibited moderate or no significant reduction in their thermostability ($T_{\text{agg}}$) or activity (Fig. 4C). Of these, PanK3-(R64W) (PanK2-(R264W)) was stable with a $T_{\text{agg}}$ value similar to the wild-type and low $\Delta G$ (0.16 Kcal/mol), but the activity was reduced to 36% of the wild-type. On the other hand, the $T_{\text{agg}}$ values of the three other surface mutant proteins, R86C (PanK2-(R286C)), S151P (PanK2-(S351P)), and K332W (PanK2-(R532W)), decreased to $-40^\circ \text{C}$ with $\Delta G$ between $-0.9$ and $-1.57 \text{ Kcal/mol}$; nevertheless, their activities were similar to that of the wild-type (Table 2). These findings suggest that the effects of these surface mutations are not always the same on both pro-
### DISCUSSION

PanK Regulation by Feedback Inhibition—Crystallographic identification of the acetyl-CoA binding sites of PanK1 and PanK3 provides an explanation for the weaker inhibitory effects of CoA. CoA and its thioesters (acetyl-CoA, malonyl-CoA, and palmitoyl-CoA) inhibit PanKs, but they show different potencies for mammalian PanK isoforms (2, 5). CoA inhibits mouse PanK3 with a significantly higher IC₅₀ than acetyl-CoA (5). CoA inhibits mouse PanK3 with a significantly higher IC₅₀ than acetyl-CoA (5). Unlike human PanKs, however, mouse PanK3 has a significantly higher IC₅₀ than acetyl-CoA (5).

### TABLE 2

| Table 2: Enzyme activities and thermostabilities of PanK3 mutant proteins |
|--------------------------|-------------|-------------|-------------|-------------|-------------|-----------|
| Mutations | Tag | ATP binding | Dimerization domain | Surface | Interior | 
| PanK3 | PanK2 | ΔΔG | PanK3 | PanK2 | PanK3 | PanK2 |
| ------------- | ------------- |-------------|-------------|-------------|-----------|
| ATP binding | G19V | G219V | 36.7 | −2.46 | 1.4 ± 1.2 | 0.4 |
| | G321R | G521R | ND | ND | ND | <0.2 |
| Dimerization domain | L213P | L413P | 36.9 | −3.32 | 2.0 ± 0.4 | ND |
| | D247N | D447N | 36.6 | −4.11 | 5.2 ± 0.5 | ND |
| | S271N | S471N | 36.1 | −4.36 | 2.6 ± 0.3 | 16 |
| | N300I | N500I | 36.6 | −1.84 | 3.5 ± 0.9 | 3 |
| | I297T | I497T | 37.7 | −3.0 | 14.5 ± 4.4 | ND |
| | I301T | I501T | 41.1 | −1.23 | 82.9 ± 2.8 | ND |
| | A309V | A509V | 42.2 | −0.77 | 62.4 ± 0.7 | 105 |
| | N311D | N511D | 40.5 | −1.71 | 164.9 ± 5.0 | ND |
| Surface | T34A | T234A | 46.2 | 0.12 | 153.5 ± 2.0 | 112 |
| | R49P | R249P | 45.8 | 0.07 | 122.9 ± 0.4 | ND |
| | R64W | R264W | 46 | 0.16 | 35.7 ± 0.7 | 58 |
| | R78L | R278L | 42.6 | −0.83 | 172.5 ± 4.3 | ND |
| | R78C | R278C | 43 | −0.74 | 184.9 ± 0.8 | ND |
| | R86C | R286C | 40.6 | −1.41 | 154.9 ± 1.7 | 176 |
| | E122G | E322G | 43.8 | −0.26 | 138 ± 6.6 | ND |
| | E122D | E322D | 43.2 | −0.59 | 149.6 ± 3.4 | ND |
| | T127I | T327I | 42.3 | −0.60 | 112 ± 3.1 | 91 |
| | S151P | S351P | 39.8 | −1.57 | 103.7 ± 8.1 | 78 |
| | N155S | N355S | 45 | −0.06 | 138.5 ± 7.0 | ND |
| | N204I | N404I | 42.4 | −0.65 | 159.4 ± 7.9 | 83 |
| | C228Y | C428Y | 44.7 | −0.14 | 147.4 ± 11 | ND |
| | T328M | T528M | 46.2 | 0.11 | 162.1 ± 7.9 | 146 |
| | K332W | K532W | 40.9 | −0.90 | 79.7 ± 1.7 | 95 |

*a* The number of early and late onset patients for PanK2 mutation is taken from Hayflick et al. (10).

*b* The value of activity for each PanK2 mutant protein is taken from Zhang et al. (11).

*c* Insoluble mutant protein.

*d* ND, not determined.

*e* Mutant not tested.

*f* Lys332 of PanK3 corresponds to Arg532 of PanK2.

PanK3 provides an explanation for the weaker inhibitory effects of CoA. CoA and its thioesters (acetyl-CoA, malonyl-CoA, and palmitoyl-CoA) inhibit PanKs, but they show different potencies for mammalian PanK isoforms (2, 5). CoA inhibits mouse PanK3 with a significantly higher IC₅₀ than acetyl-CoA (5). Similarly, CoA inhibits mouse PanK1α with the IC₅₀ of 80 μM, whereas acetyl-CoA has an IC₅₀ of 5 μM (5). The structures of catalytic core domains of two PanK isoforms suggest that the lower potency of CoA, compared with its thioesters, is attributed to the lack of the thioester carbonyl oxygen that forms a hydrogen bond with the main chain amide of a valine (supplemental Fig. S4). Additionally, the binding pocket for the acetyl moiety of acetyl-CoA is partially exposed to solvent, providing the opening for longer chain acyl moieties to interact with the same site.

The monomer folding and dimer assembly of human PanK catalytic core domains are similar to that of SaPanK (17). Unlike human PanKs, however, SaPanK is refractory to inhibition by CoA and its thioesters (18). Compared with the human PanK3 structure, SaPanK has two residues that line the acetyl-CoA pocket that would hinder the binding of feedback inhibitors (Ala₃₃⁷ of PanK3 versus Tyr²₄⁰ of SaPanK and Trp³⁴¹ of PanK3 versus Arg²₄⁸ of SaPanK) (supplemental Fig. S4). The
substitution of alanine to a bulky residue, tyrosine, may cause steric hindrance for the binding of CoA and the substitution of tryptophan to arginine introduces a positive charge, disrupting the hydrophobic binding pocket for the sulfhydryl group of CoA or the acyl groups of the thioesters. These two residues in PanK3 are also conserved in the other PanK isoforms (supplementary Fig. S4).

PanK2 Instability as a Determinant of PKAN—Structural analysis of the catalytic cores of PanK1 and PanK3 provides an opportunity to understand why many of the mutations that give rise to PKAN disease abolish the catalytic activity of PanK. Of the two mutations at the ATP-binding site, PanK3-(G19V) exhibits large reductions in both protein stability and enzymatic activity, and PanK3-(G321R) is insoluble (Fig. 4C). Low enzymatic activity, structural instability, and abnormal processing are observed with the corresponding human PanK2-(G219V) and PanK2-(G521R) mutations (4, 11). Many of the inactivating mutations are located in the dimerization interface and require between −0.77 and −4.36 Kcal/mol less free energy than the wild-type to unfold. Five of eight dimerization mutations give rise to catalytically inactive proteins in vitro, although three dimerization mutations have a less severe effect on activity (Fig. 4C). Similarly, one of the three mutations in the inner protein core, PanK3-(L363P), is insoluble (data not shown). The corresponding PanK2-(L563P) mutant protein is catalytically inactive (11). Our findings on the activities of the PanK3 mutant proteins reflect data obtained with the authentic PanK2 protein (4, 11), supporting the close relationship between the structures of these proteins. These findings suggest that proper monomer folding and dimerization are critical for the stability of PanK3 and enzymatic activity. More than 80% of patients who have the PanK2 active site or dimerization mutations are linked to the clinically early onset phenotype (10) (Table 2), indicating that the other PanK2 mutant proteins that correspond to PanK3 mutant proteins that alter protein stability may have a shorter half-life in vivo than their wild-type counterpart, and perhaps explains why these mutations are associated with early onset of the disease. However, the surface mutations have modest effects on PanK3 stability and similar in vivo experiments with the PanK2-(T528M) and PanK2-(T234A) did not reveal any differences between these two mutants and the wild-type with respect to activity, stability, and processing (4). Our biophysical studies on the analogous PanK3 isoform may suggest that there may still be more subtle effects on stability than noted in these experiments, or that alternatively, the surface mutations may relate to compromising the binding of PanK2 to as yet unidentified protein partners or regulatory ligands. On the other hand, PanK2 is translocated into the mitochondria (3, 4) and mitochondria produce most of the reactive oxygen species (46), suggesting the damage of mitochondrial PanK2 by reactive oxygen species is possible. Hydrogen peroxide is one of the major reactive oxygen species in mitochondria and modifies sulfur-containing amino acids (26, 38). Some PanK2 surface mutant proteins such as PanK2-(R278C), -(R286C), and -(T528M) may be more susceptible to hydrogen peroxide modification than wild-type and when modified, enhance protein destabilization in mitochondria.

In summary, we have presented the first structure of human PanK1 and PanK3 isoforms in complex with acetyl-CoA. The structures show a common architecture of the catalytic cores of human PanK isoforms and explain the allosteric regulation during CoA biosynthesis. Mapping of PanK2 mutations to the PanK3 structure and mutational studies on PanK3 isoform reveal that mutations at the ATP binding site, dimer interface and protein interior compromise their enzyme activities and stabilities and are mostly associated with the early onset of the PKAN phenotype, whereas surface mutations are catalytically active but show slight instabilities relative to the wild-type and are linked to the mixed onset of disease phenotype. These findings contribute to our understanding of the dysfunctional effects of polymorphic PanK2 mutations and help decipher the manner in which enzymatically active mutant proteins are linked to disease.

Acknowledgments—We thank members of Structural Genomics Consortium, notably Louis Wang, Limin Shen, and Peter Loppnau for expert technical assistance and Wolfram Tempel for assistance with x-ray data collection. We thank Cheryl H. Arrowsmith and Jinrong Min for helpful discussions. We also thank the IMCA-CAT beamline staffs of APS for help during data collection.
