

**STRUCTURAL AND FUNCTIONAL STUDY OF
THE TYPE III PANTOTHENATE KINASE FROM
*THERMOTOGA MARITIMA***

Approved by supervisory committee

Hong Zhang, Ph.D.

Margaret Phillips, Ph.D.

David Chuang, Ph.D.

Mischa Machius, Ph.D.

DEDICATION

To my wife JING WANG

and my parents

**STRUCTURAL AND FUNCTIONAL STUDY OF THE TYPE III
PANTOTHENATE KINASE FROM *THERMOTOGA MARITIMA***

by

KUN YANG

DISSERTATION

Presented to the Faculty of the Graduate School of Biomedical Sciences

The University of Texas Southwestern Medical Center at Dallas

In Partial Fulfillment of the Requirements

For the Degree of

DOCTOR OF PHILOSOPHY

The University of Texas Southwestern Medical Center at Dallas

Dallas, Texas

August, 2007

Copyright

by

KUN YANG, 2007

All Rights Reserved

ACKNOWLEDGEMENTS

First and foremost, I would like to thank my mentor, Dr. Hong Zhang, for her guidance, advice and encouragement. Her enthusiasm, knowledge and wisdom always inspire me. I am also very grateful to all the current and former members in Dr. Zhang's lab for their support, especially Yvonne Eyobo for her help with crystallization of the native PanK_{Tm} and Dr. Darek Martynowski for helping me with freezing crystals and crystallography.

This work could not be done without our excellent collaborators. Dr. Erick Strauss in Department of Chemistry, Stellenbosch University in South Africa, who first characterized the enzymatic properties of PanK-III, gave us the plasmids of PanK-III from several organisms, and carried out the mutagenesis studies.

My many thanks also go to my wonderful dissertation committee members: Dr. David Chuang, Dr. Mischa Machius, and committee chair Dr. Margaret Phillips. I would like to thank Dr. Margaret Phillips for valuable advices on pursuing the kinetics assay, Dr. Diana Tomchick, Dr. Mischa Machius and Frank Rotella for their help with synchrotron data collection and Dr. David Chuang and Dr. Scott Tso for their help with Isothermal Titration Calorimetry. I also want to thank my graduate program Biological Chemistry program and especially Shawana Criss for her help.

Finally I want to thank my wife Jing, our son Max, my parents and my parents-in-law for their support, encouragement and their unfailing love.

**STRUCTURAL AND FUNCTIONAL STUDY OF THE TYPE III
PANTOTHENATE KINASE FROM *THERMOTOGA*
*MARITIMA***

Publication No. _____

Kun Yang

The University of Texas Southwestern Medical Center at Dallas, 2007

Supervising Professor: Hong Zhang, Ph.D

ABSTRACT

Coenzyme A (CoA) is one of the most ubiquitous and essential cofactors in all living organisms. Pantothenate kinase (PanK) catalyzes the first step in the five-step universal pathway of CoA biosynthesis. Three types of PanK have been characterized so far. Prokaryotic PanK (PanK-I) and eukaryotic PanK (PanK-II) were identified previously. A third type of PanK (encoded by *coaX* gene) was identified by genetic complementation in 2005. PanK-III has a wider phylogenetic distribution than the long known PanK-I, and is nearly universally present in most of the major bacteria divisions, including many pathogenic bacteria. Different from the type I and type II PanKs, PanK-III is not feedback

inhibited by CoA, and can not use pantothenamide antibiotics as substrate. In addition, PanK-III has a high K_m for ATP (in the mM range) and requires a monovalent cation to have activity. The focus of my research is to unravel the underlying molecular basis for the unique enzymatic properties of PanK-III through crystallographic and other biochemical methods.

I have solved the first crystal structure of PanK-III from *Thermotoga maritima* (TmCoaX). As the structure reveals, PanK-III belong to the acetate and sugar kinase/heat shocks protein 70/actin (ASKHA) protein superfamily, same as PanK-II, whereas PanK-I belongs to P-loop kinase superfamily. Recently, I also solved the crystal structures of two binary complexes of PanK-III with substrate pantothenate and product phospho-pantothenate, respectively, as well as a ternary complex of PanK-III with pantothenate and ADP. Combined with isothermal titration calorimetry, we present a detailed structural and thermodynamic characterization of the interactions between PanK-III and its substrates ATP and pantothenate. Comparison of substrate binding and catalytic sites of PanK-III with that of eukaryotic PanK-II revealed drastic differences in the binding modes of both ATP and pantothenate, even though both PanK-II and PanK-III belong to the same ASKHA superfamily and may share a common catalytic mechanism. In conclusion, our studies not only are important for understanding the fundamental metabolic pathways in PanK-III-harboring pathogenic bacteria, but also provide a structural basis for designing specific inhibitors.

TABLE OF CONTENTS

ACKNOWLEDGEMENTS	v
ABSTRACT	vi
PRIOR PUBLICATIONS	xii
LIST OF FIGURES	xiii
LIST OF TABLES	xvi
LIST OF ABBREVIATIONS	xvii
CHAPTER I: General Introduction to Coenzyme A	
1.1 Macro View of Coenzyme A	1
1.2 Structure of Coenzyme A	4
1.3 Function of Coenzyme A	5
1.4 Biosynthesis of Coenzyme A	
1.4.1 Overview of the biochemical pathways for coenzyme A biosynthesis	11
1.4.2 Pantothenate: the essential precursor for coenzyme A	11
1.4.3 CoA biosynthesis from pantothenate	18
1.5 Regulation of CoA Levels	
1.5.1 Variation in cellular and tissue CoA content	22
1.5.2 Compartmentation of CoA	25
1.5.3 Regulation of CoA biosynthesis	25

1.5.4 Regulation by gene expression	28
1.5.5 Regulation by CoA degradation	29
CHAPTER II: Introduction to Pantothenate Kinases	
2.1 Introduction	31
2.2 Type I Pantothenate Kinase	31
2.3 Type II Pantothenate Kinase	34
2.4 Type III Pantothenate Kinase	37
2.5 Structural Fold Predictions for Type I II and III PanKs	41
2.6 Objectives of the Dissertation	44
CHAPTER III: Crystal Structure of PanK-III from <i>Thermotoga Maritima</i> and Phylogenetic Distribution of PanK-III	
Introduction	46
Materials and Methods	51
Results and Discussions	66
Description of PanK _{Tm} monomer structure	66
Structure of PanK _{Tm} dimer	72
The active site of PanK-III	74
Proposed ATP and pantothenate binding sites	79
Fold comparison of PanK-III with PanK-I and PanK-II	84
Type III PanK occurs in a wide range of bacterial species	88

CHAPTER IV: Thermodynamics Characterization of Substrate–Enzyme

Interactions in PanK-III

Introduction	91
Materials and Methods	91
Results and Discussions	93

CHAPTER V: Steady State Properties and Nucleotide Specificity of PanK-III

Introduction	100
Materials and Methods	100
Results and Discussions	101
Steady state kinetic parameters of PanK _{Hp} -III	101
Kinetics analysis and inhibition studies	102
Nucleotide specificity for PanK-III	105

CHAPTER VI: Structures of PanK-III Substrate / Product Complexes

Introduction	109
Materials and Methods	112
Results and Discussions	117
The overall structures of PanK _{Tm} -III complexes	117
Pantothenate binding site	120
ATP binding site	123

PanK _{Tm} -III product complex structure and catalytic	
mechanism of PanK-III	127
Comparison of PanK-III and PanK-II	129
CHAPTER VII: Conclusions and Future Directions	135
APPENDIX	141
REFERENCES	150
VITAE	166

PRIOR PUBLICATIONS

1. Singh SK, **Yang K**, Karthikeyan S, Huynh T, Zhang X, Phillips MA, Zhang H. (2004) “The thrH gene product of *Pseudomonas aeruginosa* is a dual activity enzyme with a novel phosphoserine:homoserine phosphotransferase activity.” *Journal of Biological Chemistry*. 279, 13166.
2. **Yang K**, Eyobo Y, Brand L, Martynowski D, Tomchick D, Strauss E, and Zhang H. (2006) “Crystal Structure of a Type III Pantothenate Kinase: Insight into the Mechanism of an Essential Coenzyme A Biosynthetic Enzyme Universally Distributed in Bacteria.” *Journal of Bacteriology*, 15, 5532.
3. Martynowski D, Eyobo Y, Li T, **Yang K**, Liu A, and Zhang H. (2006) “Crystal Structure of α -Amino- β -Carboxymuconate- ϵ -Semiaidehyde Decarboxylase (ACMSD): Insight into the Active Site and Catalytic Mechanism of a Novel Decarboxylation Reaction.” *Biochemistry*, 35, 10412.
4. **Yang K**, Erick Strauss, Margaret Phillips and Zhang H. “Structural basis for substrate binding and catalytic mechanisms of Type III Pantothenate kinase.” (in preparation)

LIST OF FIGURES

- Figure 1.1** The structure of coenzyme A and the formation of its thioesters
- Figure 1.2** Two general modes of reactivity of Acetyl-CoA
- Figure 1.3** The citric acid cycle: the major catabolic pathway for acetyl-CoA in aerobic organisms
- Figure 1.4** Comparison of coenzyme A and acyl carrier protein (ACP)
- Figure 1.5** The pantothenate biosynthetic pathway in *E. coli*
- Figure 1.6** CoA biosynthetic pathway from pantothenate
- Figure 1.7** Expression of the CoASy mRNA in human tissues and cancer cell lines
- Figure 3.1** Crystal of Native PanK_{Hp}
- Figure 3.2** X-ray Diffraction Pattern of PanK_{Hp} Crystal
- Figure 3.3** Crystals of PanK_{Pa}
- Figure 3.4** Crystal of PanK_{Tm}
- Figure 3.5** X-ray Diffraction Pattern of PanK_{Tm} Crystal
- Figure 3.6** Electron density map after density modification
- Figure 3.7** Structure of PanK_{Tm} monomer
- Figure 3.8** Ribbon diagram of PanK_{Tm} dimer
- Figure 3.9** Fold comparison of PanK-III with representative members of the ASKHA superfamily

Figure 3.10 Multiple sequence alignment of representative sequences of PanK-III (group I) and actin/hsp70/sugar kinase superfamily with known structures (group II)

Figure 3.11 Model of MgATP and pantothenate binding in PanK_{Tm} active site

Figure 3.12 Crystal structural comparison of PanK_{Tm}-III and PanK_{Ec}-I

Figure 3.13 Crystal structural comparison of PanK_{Tm}-III and PanK_{Sa}-II

Figure 4.1 Calorimetric titration of PanK_{Tm} with substrates pantothenate in the absence or presence of AMPPNP

Figure 4.2 Calorimetric titration of PanK_{Hp} with substrates pantothenate in the absence or presence of AMPPNP

Figure 4.3 Calorimetric titration of PanK_{Tm} with AMPPNP

Figure 4.4 Calorimetric titration of PanK_{Hp} with substrates AMPPNP in the absence or presence of pantothenate

Figure 5.1 Steady state kinetic parameters of PanK_{Hp}

Figure 5.2 Initial rate plot of PanK_{Hp} versus pantothenate concentration

Figure 5.3 Kinetics of inhibition of PanK_{Hp} by AMPPNP

Figure 5.4 Nucleotide specificity of PanK-III

Figure 6.1 Binary complex crystal structure of PanK_{Tm}•pantothenate

Figure 6.2 Ternary complex crystal structure of PanK_{Tm}•pantothenate•ADP

Figure 6.3 Binary complex crystal structure of PanK_{Tm}•phosphor-pantothenate

Figure 6.4 Electron density of the bound substrate and product

Figure 6.5 Stereoview of pantothenate and ATP binding site of PanK_{Tm}-III

Figure 6.6 Superposition of substrate binding site of PanK_{Tm}-III (Cyan) and PanK_{Pa}-III (yellow)

Figure 6.7 Interactions with product phospho-pantothenate (P-Pan) in PanK_{Tm}-III•P-Pan product complex

Figure 6.8 Comparison of the nucleotide binding in PanK-III and PanK-II

Figure 6.9 Stereoview of the superposition of the pantothenate binding sites in PanK_{Tm}-III (light and dark blue) and human PanK-II (orange)

LIST OF TABLES

Table 2.1 Classification of kinase activity by fold group

Table 3.1 Summary of PanK-III expression, crystallization and diffraction from
five different organisms

Table 3.2 Data collection, phasing, and refinement statistics for native crystal of
PanK_{Tm}

Table 3.3 Proteins structurally similar to PanK_{Tm} and structural alignment
statistics from DALI

Table 3.4 Effect of mutation of the active site aspartate residues on the activity of
PanK_{Hp}

Table 3.5 Phylogenetic distributions of different types of PanK in bacteria

Table 4.1 Thermodynamic parameters for interactions between substrate and
PanK-III

Table 6.1 Data collection and refinement statistics for complex crystals

LIST OF ABBREVIATIONS

CoA	– coenzyme A
Pan	– pantothenate
PanK	– pantothenate kinase
PanK-I	– type I PanK
PanK-II	– type II PanK
PanK-III	– type III PanK
ACP	– acyl carrier protein
TCA	– tricarboxylic acid
THF	– tetrahydrofuran
ATP	– adenosine 5'– triphosphate
PKAN	– pantothenate kinase associated neurodegeneration
Baf	– Bvg accessory factor
RNase	– ribonuclease
ASKHA	– acetate and sugar kinase / heat-shock protein 70 (hsp70) / actin
Hsc	– heat-shock cognate protein
ADO	– adenosine motif
PHO1	– phosphate motif 1
PHO2	– phosphate motif 2
GK	– glycerol kinase
SK	– sugar kinase

Ts – temperature sensitive

ITC – isothermal titration calorimetry

IPTG – isopropyl-1-thio- β -D-galactopyranoside

Tm – *Thermotoga maritima*

Ba – *Bacillus anthracis*

Ec – *Escherichia coli*

Sa – *Staphylococci aureus*

Hp – *Helicobacter pylori*

Pa – *Pseudomonas aeruginosa*

CHAPTER I

General Introduction to Coenzyme A

1.1 Macro View of Coenzyme A

The metabolism of organisms involves a giant array of chemical reactions, but most of them fall under a few basic types of group transfer reactions (Wimmer and Rose, 1978). Such common chemistry allows the cells to utilize a limited set of intermediates that are usually small organic non-protein molecules to carry chemical groups between different reactions. These group-transfer intermediates are named coenzymes, such as ATP, the biochemical carrier of phosphate groups, and coenzyme A, the coenzyme that carries acyl groups (acetyl, fatty acyl and others) (Jackowski and Rock, 1981; Leonardi et al., 2005b). Each class of group-transfer reaction is carried out by a certain coenzyme.

Among all the coenzymes, coenzyme A (CoA, where A stands for *acetylation*) is one of the most ubiquitous and essential cofactors in all living organisms (Jackowski, 1996). It plays critical roles in the synthesis and oxidization of fatty acids, and the oxidation of pyruvate in the citric acid cycle. Since CoA is chemically a thiol (Figure 1.1), it can react with carboxylic acids to form thioesters, thus functioning as a carrier for acyl group. A molecule of CoA carrying an acetyl group is also referred to as acetyl-CoA. Along with its

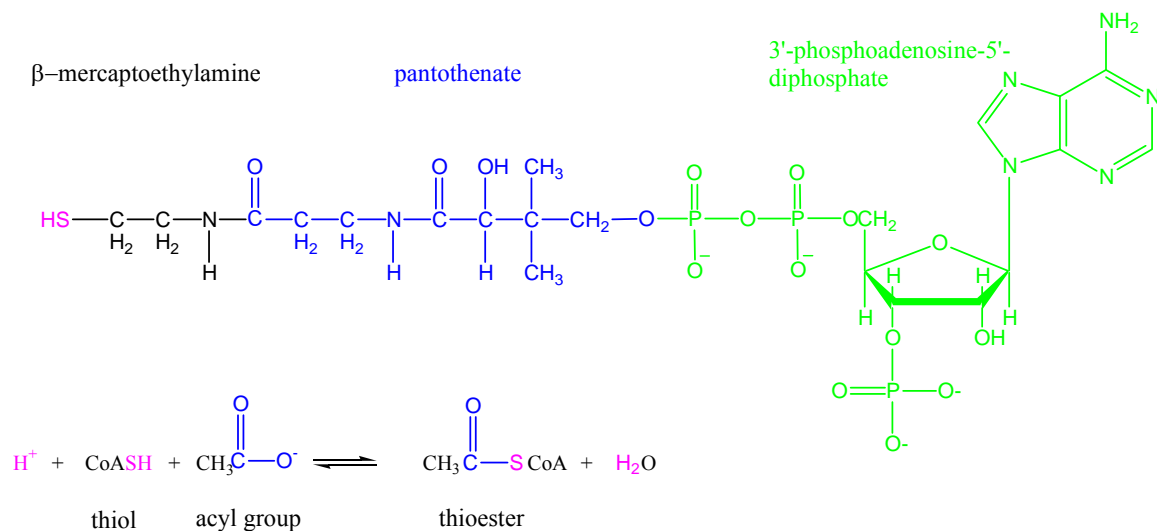


Figure 1.1 The structure of coenzyme A and the formation of its thioesters

CoA consists of 3'-phosphoadenosine-5'-diphosphate (in green) coupled to pantothenate (in blue), which in turn forms an amide bond with β -mercaptoethylamine. The sulfhydryl group of CoA (in red, a thiol group) is the key functional group of the molecule. The formation of a thioester from CoA and acyl group is also shown.

thioesters, CoA is used as a substrate for approximately 9% of all enzyme activities, where it participates in a variety of acyl transfer reactions (Begley et al., 2001). It has been estimated that CoA is involved in over 100 different reactions in intermediary metabolism (Begley et al., 2001).

Five enzymatic steps are involved in the biosynthetic pathway of CoA from pantothenate or vitamin B₅, and all genes in this pathway are essential for cell survival and growth (Jackowski, 1996, Jackowski and Rock, 1981). The biosynthesis of CoA is studied extensively due to its recognition as a novel target for antibacterial drug discovery and the association of human neurodegenerative disorder with mutations in pantothenate kinase (Zhou et al., 2001; Leonardi et al., 2007).

Because of CoA's ubiquitous nature and its critical role as a cofactor in cell metabolism, its level must be stringently regulated (Jackowski and Rock, 1981). In addition, the production of one molecule of CoA uses three ATP equivalents. Therefore, the CoA biosynthetic pathway is an energetically expensive pathway, and it is necessary to have the pathway itself be regulated as not to waste cellular energy.

In the following part of the chapter, the structure, function, biosynthesis and regulation of coenzyme A will be discussed in detail.

1.2 Structure of Coenzyme A

Coenzyme A was discovered in 1945 by the German-born biochemist Dr. Fritz Albert Lipmann, who was the first to show that a coenzyme was required to facilitate the biological acetylation reactions (Lipmann et al., 1950; Baddiley et al., 1953). In recognition of his pioneering work in elucidating the role of this very important coenzyme, Dr. Lipmann was awarded the Nobel Prize in physiology and medicine in 1953.

The structure of CoA was first reported in 1953 (Baddiley et al., 1953). As shown in Figure 1.1, CoA consists of 3'-phosphoadenosine-5'-diphosphate (in green) in a phosphoryl ester linkage with pantothenate (vitamin B₅, in blue), which in turn forms an amide bond with β -mercaptoethylamine. The adenine moiety of CoA serves as the recognition site for enzyme to bind CoA, which increases the affinity and specificity of CoA when it binds to the enzyme in reaction (Garrett, 1999; Mishra and Drueckhammer, 2000). The sulfhydryl group of CoA (shown in red, also called a thiol group) is the key functional group of the molecule (Wharton, 1981). Its reaction with acyl groups to form activated thioesters like acetyl-CoA is illustrated in Figure 1.1. This reaction is thermodynamically unfavorable in the forward direction shown ($\Delta G^{\circ} = +7.5$ kcal/mole). The generation of activated acyl groups in the form of acyl-CoA molecules is coupled to energy releasing processes such as oxidative

decarboxylation or hydrolysis of high-energy phosphate bonds as in the activation of fatty acids for oxidation. Because the thioester bond has a large negative standard energy of hydrolysis (-7.5 kcal/mole), this thioester bond is a high energy bond and the acyl group attached to CoA is a highly activated group.

1.3 Functions of Coenzyme A

Coenzyme A is the common carrier for activated acyl groups in prokaryotic and eukaryotic cells (Jackowski, 1996). It is required for a multitude of reactions for both biosynthetic and degradative pathways, forming derivatives that are key intermediates in energy metabolism. In general, CoA has two major functions: it activates acyl groups for transfer by nucleophilic acyl substitution, and it activates the α -hydrogen of the acyl group for deprotonation. These two functions are illustrated in Figure 1.2.

Acetyl-CoA is the most common CoA thioester and acts as a central “hub” in metabolism. It is the intermediate in the breakdown of carbohydrates, fat, and amino acids, and a precursor in the synthesis of fat, cholesterol, and ketone bodies. Chemically acetyl-CoA is a thioester between CoA and acetic acid (an acyl group carrier). It is produced during the second step of aerobic cellular respiration, pyruvate decarboxylation, which occurs in the matrix of the mitochondria (Murray et al, 2006). Acetyl-CoA then enters citric acid cycle (Figure 1.3). Several enzymes are responsible for the formation of acetyl-CoA including

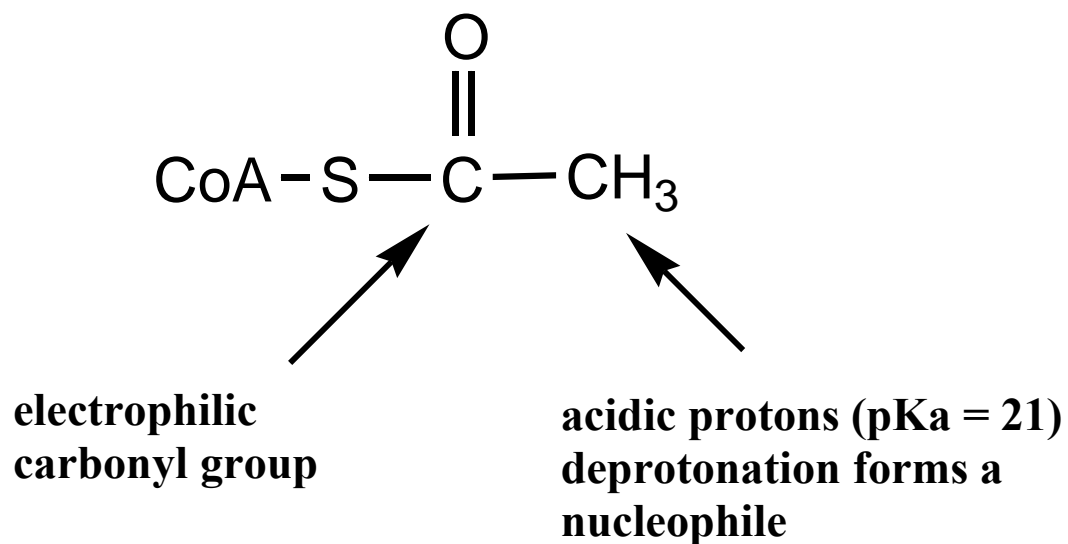


Figure 1.2 Two general modes of reactivity of Acetyl-CoA

1. The thioester carbonyl can act as an electrophile to react with a nucleophile co-substrate.
2. Upon deprotonation the thioester α -carbon can react as a nucleophile.

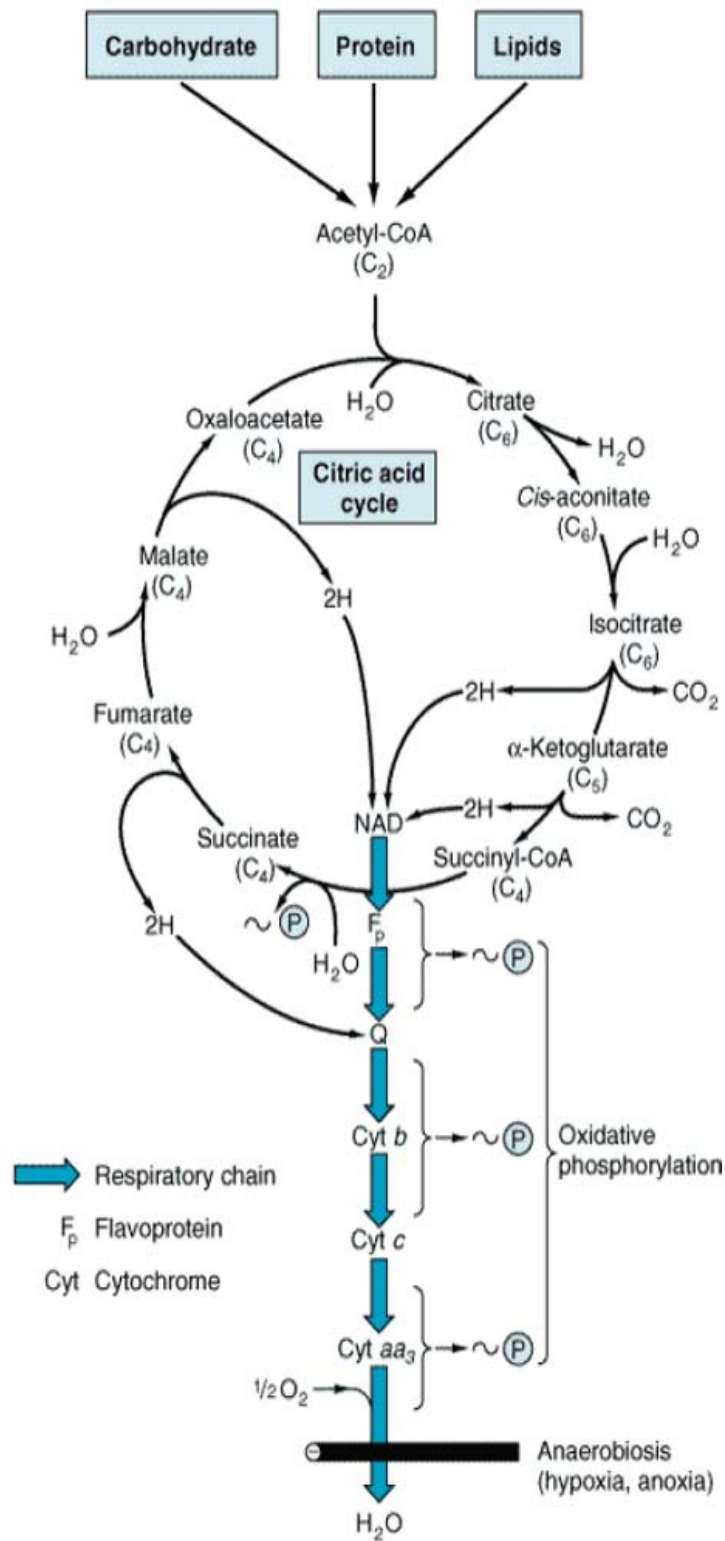


Figure is adapted from (Murray et al., 2006).

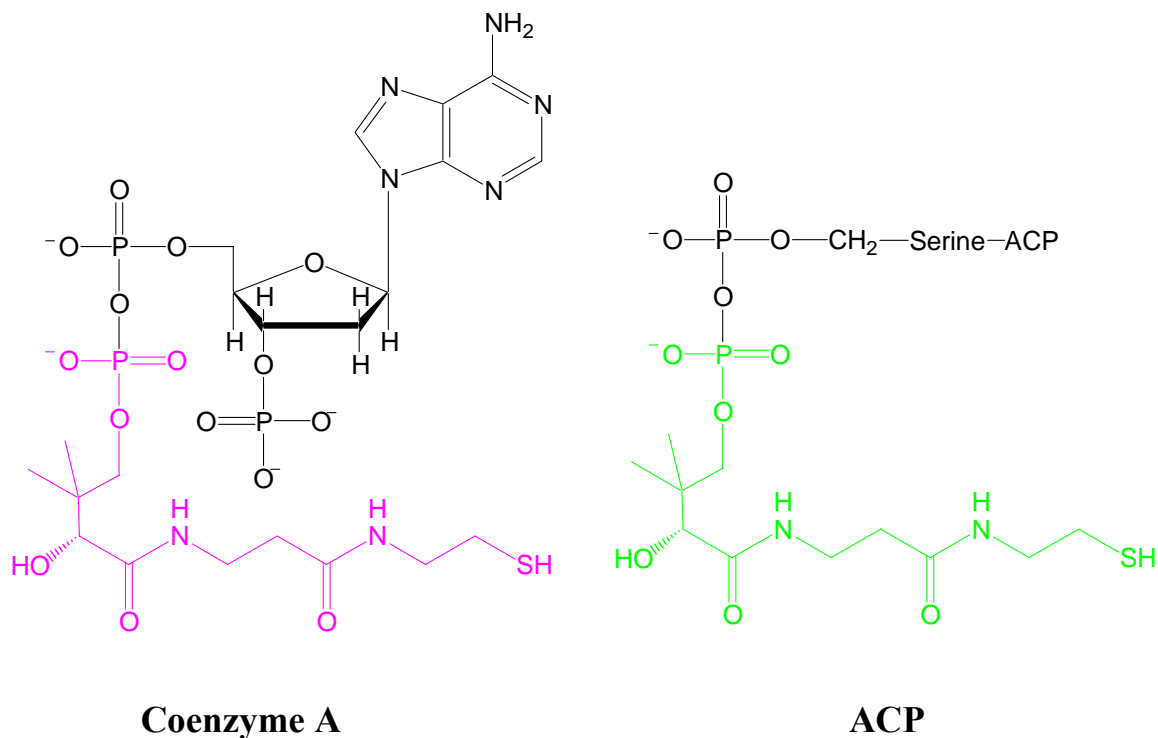
Figure 1.3 The citric acid cycle: the major catabolic pathway for acetyl-CoA in aerobic organisms

Acetyl-CoA, the product of carbohydrate, protein, and lipid catabolism, is taken into the cycle and oxidized to CO_2 with the release of reducing equivalents (2H). Subsequent oxidation of 2H in the respiratory chain leads to phosphorylation of ADP to ATP.

acetyl-CoA synthetase, phosphotransacetylase, ATP citrate lyase and thiolase (Mishra and Drueckhammer, 2000).

The enzymes that utilize acetyl-CoA can be divided into two major classes. They are the Claisen enzymes and acetyltransferases (Mishra and Drueckhammer, 2000). Claisen enzymes catalyze reactions involving deprotonation of the α -carbon (Figure 1.2). They utilize acetyl-CoA as a nucleophilic substrate via deprotonation of the methyl group. Acetyltransferases catalyze the nucleophilic acyl substitution reactions at the carbonyl carbon (Figure 1.2) (Mishra and Drueckhammer, 2000). As their name suggests, acetyltransferases catalyze the transfer of the acetyl group from acetyl-CoA to a nucleophile acceptor, most commonly an alcohol or amine. Acetyltransferases have broad biological significance. Bacterial acetylation of antibiotics inactivates the drugs, which explains the antibiotic resistance in many bacteria. Acetylation of histones catalyzed by histone-N-acetyltransferase is a critical control element in gene transcription. As acetylcholine is a major neurotransmitter, acetylation also plays a key role in the transmission of nerve impulses.

CoA also serves as a precursor for acyl carrier protein (ACP), which is a larger version of CoA and also uses the phosphopantetheine group as a functional group (Figure 1.4) (Rawlings and Cronan, 1992). CoA and ACP are the two



Phosphopantetheine group of CoA

Phosphopantetheine prosthetic group of ACP

Figure 1.4 Comparison of coenzyme A and acyl carrier protein (ACP)

The phosphopantetheine group of CoA and ACP are highlighted. Fatty acids are conjugated to both CoA and ACP through the sulfhydryl of the phosphopantetheine group.

predominant acyl group carriers in cells: ACP is used in fatty acid biosynthesis, whereas CoA is used in β -oxidation of fatty acids.

1.4 Biosynthesis of Coenzyme A

1.4.1 Overview of the biochemical pathways for coenzyme A biosynthesis

The biosynthesis of coenzyme A can be divided into two parts in bacteria (Jackowski, 1996). First, pantothenate is synthesized *de novo*, and next the universal biosynthesis of CoA from pantothenate occurs. The second part here, biosynthetic pathway from pantothenate to CoA, is essential in both prokaryotes and eukaryotes. All the genes coding for the enzymes that catalyze the reactions in the biosynthetic pathway are known, although not in all species.

1.4.2 Pantothenate: the essential precursor for coenzyme A

Pantothenate (or pantothenic acid), also known as vitamin B₅, is one of the B complex of vitamins (Tahiliani and Beinlich, 1991). In the 1930s, pantothenate was identified when investigators were looking for a substance necessary for yeast to grow. Researchers found that diets lacking this substance caused certain disorders in animals, including a retarded growth rate, anemia, degenerated nerve tissue, decreased production of antibodies, ulcers, and malformed offspring. Thereafter, pantothenate was found to play a fundamental

role in all organisms. Because this vitamin is found virtually everywhere in biology, it was designated “pantothenate”, derived from the Greek word “*pantother*” meaning “from everywhere” (Tahiliani and Beinlich, 1991).

Animals and some microbes lack the capacity to synthesize pantothenate. They are totally dependent on the uptake of pantothenate from their diets. On the contrary, most bacteria, plants and fungi are capable of synthesizing pantothenate *de novo* (Tahiliani and Beinlich, 1991; Begley et al., 2001). It has been reported that *Escherichia coli* (*E. coli*), for example, produce and secrete 15 times more pantothenate than that is required for intracellular CoA biosynthesis (Jackowski and Alix, 1990). The excess pantothenate is released into the medium and is available to organisms that harbor this microorganism. Because of pantothenate's ubiquitous nature, very few cases of pantothenate deficiency in human have been reported (Jackowski and Rock, 1981; Leonardi et al., 2005b).

1.4.2.1 The Biosynthetic Pathway of Pantothenate

The pantothenate biosynthetic pathway is best characterized in *Escherichia coli* (Figure 1.5). It involves four steps catalyzed by enzymes encoded by *panB*, *panC*, *panD*, and *panE* genes (Merkel and Nichols, 1996). Most bacteria synthesize pantothenate from aspartate, α -ketoisovalerate and ATP. Since animals do not have this pantothenate biosynthesis pathway, it offers targets for developing drugs against microbial pathogens.

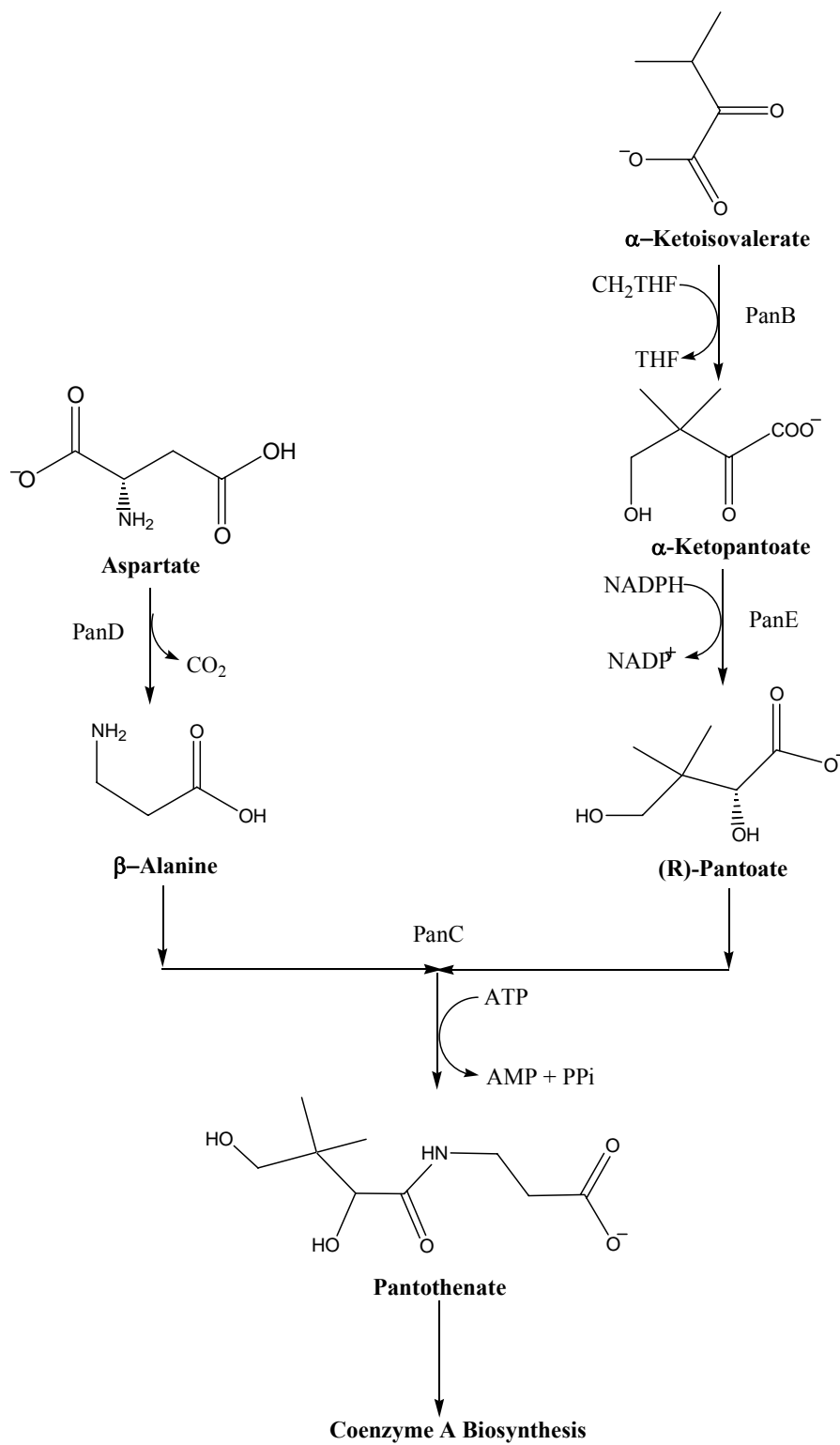


Figure 1.5 Pantothenate biosynthetic pathway in *E. coli*

Four steps are involved in pantothenate biosynthetic pathway, and they are catalyzed by enzymes encoded by *panB*, *panC*, *panD* and *panE* genes. (THF: tetrahydrofuran)

The first step in the biosynthesis of pantothenate is the transfer of a hydroxymethyl group from $N^{5,10}$ -methylenetetrahydrofolate to α -ketoisovalerate by α -ketopantoate hydroxymethyltransferase, the product of the *panB* gene (Jones et al., 1993). Next, (R)-Pantoate is synthesized from α -ketopantoate by α -ketopantoate reductase, which is the product of the *panE* gene. This step proceeds stereospecifically, with an inversion of the configuration at the C-3 carbon of α -ketoisovalerate (Jackowski, 1996). On the other branch, aspartate decarboxylase (product of *panD* gene) converts aspartate to β -alanine (Cronan, 1980; Williamson and Brown, 1979). It has been shown that *panD* mutants can grow with the supplementation of β -alanine (Cronan, 1980; Williamson and Corkey, 1979).

Pantothenate synthetase, the product of *panC* gene, is the enzyme responsible for the ATP-dependent condensation of pantoate with β -alanine (Cronan et al., 1982; Maas, 1952). It was found to be a dimer from the crystal structure (von Delft et al., 2001), although earlier studies indicated that it could be a tetramer in solution (Miyatake et al., 1978). Since the activity of pantothenate synthetase is not tightly regulated, *E. coli* secretes most of the synthesized pantothenate into the medium (Maas and Davis, 1950). The overproduction of pantothenate by bacteria highlights a role of intestinal flora in providing this vitamin to the mammalian host (Jackowski and Rock, 1981).

Like microbes, some higher organisms such as plants can also synthesize pantothenate *de novo* (Rathinasabapathi and Raman, 2005). Genes encoding PanB and PanC have been identified from *Arabidopsis thaliana* genome, and the cDNAs of these genes were able to complement the pantothenate auxotrophic phenotype of *E. coli* mutants (Ottenhof et al., 2004). Subcellular localization studies show that PanB is targeted to mitochondria whereas PanC is found in cytosol (Ottenhof et al., 2004), which implies that pantothenate synthesis in plants occurs in different subcellular compartments and transporters for pathway intermediates are required (Rathinasabapathi and Raman, 2005).

1.4.2.2 Pantothenate transports

Bacteria are able to move pantothenate across the membrane bidirectionally. The best-characterized transport system is in *E. coli*, where pantothenate uptake is mediated by pantothenate permease (also named as the PanF protein, encoded by the *panF* gene) (Nakamura and Tamura, 1973; Vallari and Rock, 1985). Pantothenate permease utilizes a sodium-cotransport mechanism to concentrate pantothenate from the medium (Jackowski and Alix, 1990). PanF is predicted to contain 12 transmembrane hydrophobic domains connected by short hydrophilic chains, which is a topological motif characteristic of other cation-dependent permeases of the major facilitator superfamily of

proteins (Jackowski and Alix, 1990). The transport system is highly specific for pantothenate, with a K_t of 0.4 μM and a maximum velocity of 1.6 pmol/min/ 10^8 cells (Nakamura and Tamura, 1973).

In bacteria lacking *de novo* pantothenate biosynthesis, such as *Streptococcus pneumoniae*, *Lactobacillus lactis* and *Hemophilus influenzae*, pantothenate permease transport system is indispensable (Gerdes et al., 2002). Pantothenate permease is only responsible for uptaking pantothenate into the cells, whereas another yet uncharacterized transport system is responsible of secretion of pantothenate from the cells.

In higher eukaryotes, it was first thought that the transport of pantothenate in the rat small intestine occurred by simple diffusion (Shibata et al., 1983), but now there is strong evidence for a specific transport mechanism (Fenstermacher and Rose, 1986; Prasad et al., 1997). The process is unidirectional, sodium-dependent and active, with a pantothenate K_t of 17 μM measured across the brush-border membrane of rat jejunum and chick intestinal cells (Fenstermacher and Rose, 1986). As first described in human placental epithelial cells, pantothenate, biotin and lipoate all use the same transporter (Grassl, 1992; Prasad et al., 1997). This multivitamin transporter was cloned by several groups from different sources, including rat placenta, a human choriocarcinoma cell line and rabbit intestine (Prasad et al., 1999; Prasad et al., 1998; Wang et al., 1999). The mammalian

pantothenate transporter belongs to the sodium-coupled glucose transporter family and the proteins in this family share a common core of 13 transmembrane helices (Hirabayashi et al., 2004).

1.4.3 CoA biosynthesis from pantothenate

The biosynthesis of CoA from pantothenate is a universal and essential pathway in prokaryotes and eukaryotes (Daugherty et al., 2002). CoA is synthesized in five steps, and they are catalyzed by enzymes encoded by *coaA*, *B*, *C*, *D*, and *E*, respectively (Figure 1.6). The first step is the phosphorylation of pantothenate (**1**, numbers correspond to the ones shown in figure 1.6) to 4'-phosphopantothenate (**2**) by pantothenate kinase. Next, 4'-phosphopantothenate (**2**) is condensed with cysteine and decarboxylated to form 4'-phosphopantetheine (**4**). The AMP moiety of ATP is then added to form dephospho-CoA (**5**), which is subsequently phosphorylated to produce CoA (**6**). Metabolite labeling experiments have only detected pantothenate (**1**) and 4'-phosphopantetheine (**4**) as intermediates of the reactions in the cells, suggesting that the enzymes utilizing these two compounds (pantothenate kinase, CoaA, and 4'-phosphopantetheine adenylyltransferase, CoaD) catalyzes the rate-limiting steps in the biosynthesis of CoA (Jackowski and Rock, 1981; Zhyvoloup et al., 2002).

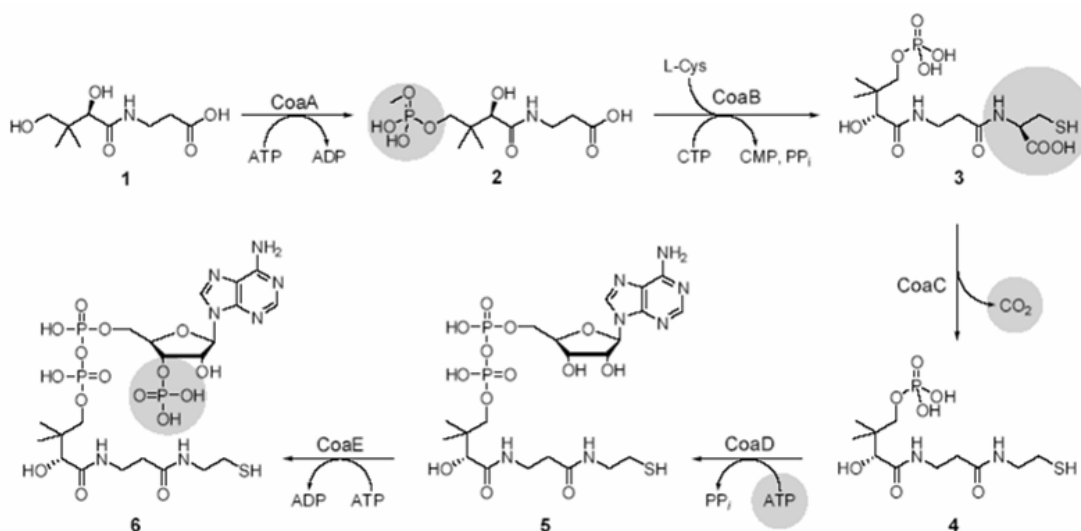


Figure 1.6 CoA biosynthetic pathway from pantothenate

Pantothenate (vitamin B₅) is first phosphorylated to 4'-phosphopantothenate by pantothenate kinase (CoaA), and then condensed with cysteine and decarboxylated to form 4'-phosphopantetheine. These two reactions are catalyzed by the 4'-phosphopantetheinoylcysteine synthase (CoaB) and 4'-phosphopantetheinoylcysteine decarboxylase (CoaC) domains of a bifunctional enzyme in prokaryotes and by two distinct proteins in eukaryotes (Strauss et al., 2001; Gerdes et al., 2002; Daugherty et al., 2002). 4'-Phosphopantetheine is subsequently converted to dephospho-CoA by phosphopantetheine adenylyltransferase (CoaD) and phosphorylated by dephospho-CoA kinase (CoaE) at the 3'-OH of the ribose to form CoA. The CoaD and CoaE activities are associated with two separate enzymes in prokaryotes and plants but fused in a bifunctional enzyme, also termed the CoA synthase, in mammals (Geerloff et al., 1999; Mishra et al., 2001; Gerdes et al., 2002).

1.4.3.1. Phosphorylation of Pantothenate

The first committed step in CoA biosynthesis is ATP-dependent phosphorylation of pantothenate (**1**) to form 4'-phosphopantothenate (**2**), which is catalyzed by pantothenate kinase (also known as PanK or CoaA) (Rock et al., 2002; Song and Jackowski, 1992). PanK is encoded by the gene *coaA*, and is an essential regulatory enzyme in CoA biosynthesis. This dissertation will focus on the structural analysis of one type of PanK. Part II in this chapter focuses on PanKs and discusses the enzyme in greater detail.

1.4.3.2. Formation of 4'-Phosphopantetheine

The second step in this pathway is the condensation of 4'-phosphopantothenate (**2**) with cysteine to produce 4'-phosphopantothenoylcysteine (**3**). This reaction is catalyzed by an enzyme known as 4'-phosphopantothenoylcysteine synthetase, (PPCS/CoaB, *coaB* gene product). Next, decarboxylation of 4'-phosphopantothenoylcysteine (**3**) by 4'-phosphopantothenoylcysteine decarboxylase (PPCDC/CoaC, *coaC* gene product) yields 4'-phosphopantetheine (**4**). 4'-phosphopantetheine also functions as an acyl-group carrier, which activates carboxylic acids for biological Claisen reactions and the formation of peptides and esters (Strauss et al., 2001).

It was shown that the NH₂-terminal domain of the Dfp protein (a flavin mononucleotide (FMN)-containing enzyme) from bacteria had CoaC activity

(Kupke et al., 2000). Later, Strauss *et al.* (Strauss et al., 2001) showed that the Dfp protein is a bifunctional enzyme and catalyzes not only the decarboxylation of 4'-phosphopantetheinoylcysteine (**3**) but also its synthesis from 4'-phosphopantothenate (**2**) and cysteine using cytidine 5'-triphosphate (CTP) as the activating nucleotide. Therefore, Strauss *et al.* (Strauss et al., 2001) renamed the *dfp* gene to *coaBC*. The eukaryotic counterparts of these two enzymes are monofunctional and show very little sequence similarity to the bacterial enzymes (Daugherty et al., 2002; Genschel, 2004). No accumulation of 4'-phosphopantetheinoylcysteine (**3**) or 4'-phosphopantothenate (**2**) is detected *in vivo* (Jackowski and Rock, 1981), reflecting a rapid conversion of compound **2** to **3** and **3** to **4** (Figure 1.6).

1.4.3.3 Conversion of 4'-phosphopantetheine to Coenzyme A

Two enzymatic steps are involved in the conversion of 4'-phosphopantetheine (**4**) to CoA (**6**). 4-phosphopantetheine adenylyltransferase (CoaD or PPAT, encoded by the *coaD* gene) catalyzes the adenylation of 4'-phosphopantetheine (**4**) to yield dephospho-coenzyme A (**5**). In this step CoaD transfers the AMP moiety from ATP to 4'-phosphopantetheine (**4**) (Geerlof et al., 1999). At last, phosphorylation of the 3'-hydroxyl of dephospho-coenzyme A (**5**) by dephospho-coenzyme A kinase (CoaE, product of the *coaE* gene) completes the biosynthesis of coenzyme A (**6**).

In mammals CoaD and CoaE are co-purified and exist as a bifunctional protein called CoA synthase (CoASy) (Zhyvoloup et al., 2002; Daugherty et al., 2002). Interestingly, some organisms like *D. melanogaster* and *C. elegans* have homologues of mammalian CoASy, whereas in lower organisms, such as *S. cerevisiae* and bacteria, the CoaD and CoaE activities reside on different proteins (Zhyvoloup et al., 2002). These differences may suggest different modes of regulation for CoA biosynthesis among different organisms.

1.5 Regulation of CoA Levels

As discussed previously, CoA and its thioesters function as carriers of acetyl and acyl groups in all organisms and are essential for numerous biosynthetic, energy-yielding, and degradative metabolic pathways. Therefore, it is necessary to have CoA level tightly controlled. The regulation of CoA levels can be achieved by compartmentalization of CoA, regulation of CoA biosynthesis, regulation of gene expression and regulation by degradation (Leonardi et al., 2005b).

1.5.1 Variation in cellular and tissue CoA content

Metabolic labeling experiments showed that the cellular CoA content fluctuates depending on the carbon source for growth in *E. coli* (Vallari and

Jackowski, 1988). The CoA is highest in bacteria growing on 400 μ M glucose (Vallari et al., 1987), and the CoA is much lower in bacteria growing on a mixture of amino acids, indicating that CoA is important for the conversion of glucose to amino acids via citric acid cycle (Vallari et al., 1987). It has also been shown that the intracellular concentration of CoA can be manipulated by varying the amount of pantothenate or 3-alanine in the growth medium of mutant strains that can not synthesize pantothenate or one of its precursors (Jackowski and Rock, 1986).

The tissue level of CoA is regulated by various extracellular stimuli, including hormones, nutrients, and cellular metabolites (Vallari et al., 1987; Jackowski et al., 1986; Vallari and Jackowski, 1988). Insulin, glucose, fatty acids, pyruvate, and ketone bodies inhibit CoA biosynthesis, whereas glucocorticoids and glucagon, and drugs like clofibrate, increase tissue concentration of CoA (Leonardi et al., 2005b; Zhyvoloup et al., 2002). It has been observed that the homeostasis of CoA is altered in several disease states, including diabetes, starvation, alcoholism, hypertension, and certain types of tumors (Daugherty et al., 2002; Reibel et al., 1983; Reibel et al., 1981; Zhyvoloup et al., 2002). Daugherty et al. (2002) reported the variation for mRNA level of CoASy (CoA synthase, a point of regulation in CoA biosynthesis, and a bifunctional enzyme that carries the function of both CoaD and CoaE) in normal and cancer cell lines (Figure 1.7). All cancerous tissues examined had higher expression of the CoASy mRNA

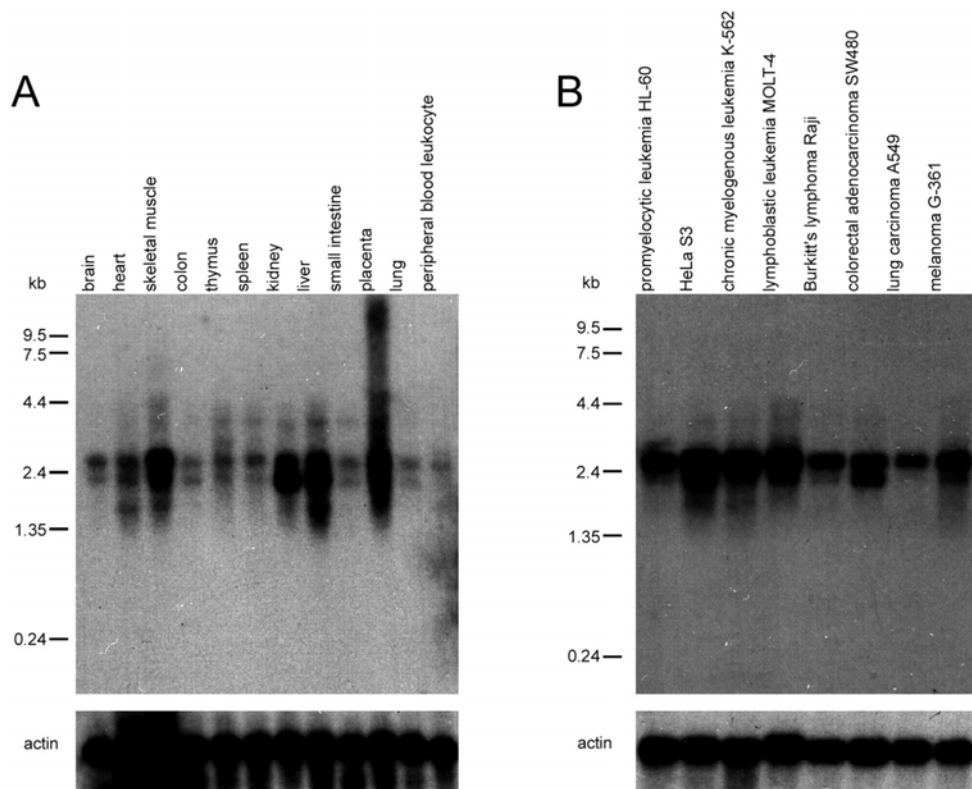


Figure 1.7 Expression of the CoASy mRNA in human tissues and cancer cell lines

mRNAs from normal human tissues (A) and human cancer cell lines (B) were analyzed by Northern blot hybridization with a radiolabeled probe corresponding to the central domain (PPAT) of human CoASy. Blots were exposed to an X-ray film at -70°C . The blots were subsequently washed and rehybridized with radiolabeled β -actin to normalize for mRNA loading levels. RNA marker positions are shown. (Figure was adapted from Daugherty et al., 2002.)

than any of the normal tissues, with the exception of skeletal muscle, liver, and kidney (Daugherty et al., 2002).

1.5.2 Compartmentation of CoA

Eukaryotic cells contain sequestered pools of CoA (Leonardi et al., 2005b). The availability and concentration of CoA in these sequestered pools controls the rate of the reactions CoA involved in. For example, mitochondria CoA in mammals is used in the citric acid cycle and fatty acid β -oxidation, and the concentration of CoA and its thioesters regulates the rate of these reactions (Wang et al., 1991).

1.5.3 Regulation of CoA biosynthesis

There are two sites of regulation in CoA biosynthesis pathway. Pantothenate kinase (PanK, CoaA) is proposed to be the master regulator (Rock et al., 2000), and CoA is a feedback inhibitor of PanK activity by competitively binding to the ATP-binding site on the enzyme (Yun et al., 2000). CoaD is the secondary site for the pathway regulation, although the mechanism for this regulation remains to be elucidated.

1.5.3.1 Feedback regulation by Pantothenate Kinase

Pantothenate kinase is the key regulatory point in the control of CoA levels in the cells (Leonardi et al., 2005b; Rock et al., 2000). Since 15 times more pantothenate is produced than is phosphorylated in *E. coli* growing on glucose (Jackowski and Alix, 1990), PanK is the primary rate-limiting step in CoA biosynthesis. Most of the pantothenate kinase activity is controlled by CoA, the end-products of the pathway, and less affected by CoA thioesters, including acetyl-CoA (Song and Jackowski, 1992). The lower sensitivity to acetyl-CoA (the major CoA species when cells grow on glucose) allows the CoA pool to expand as CoA biosynthesis pathway proceeds, and accommodate the metabolic demands of cells dividing rapidly using glucose as a carbon source (Leonardi et al., 2005b).

Structural analysis revealed the mechanism of feedback inhibition of PanK by CoA. The crystal structure of *E. coli* PanK in complex with either ATP or CoA has been published (Yun et al., 2000). Based on the structures, site-directed mutagenesis studies were carried out to generate mutants that were resistant to feedback inhibition of CoA by decreased binding efficiencies to CoA (Rock et al., 2003). Mutant CoaA[R106A] retains 50% activity but is refractory to inhibition by CoA. Arg 106 was proposed to be critical for CoA binding because it forms a salt bridge with the phosphate attached to the 3'-hydroxyl of the CoA ribose (Yun et al., 2000). Cells expressing CoaA[R106A] mutant protein produce double amount of CoA, compared to that synthesized by the wild-type CoaA protein

expressed in the *coaA15*(Ts) strain which lacks PanK activity at elevated temperature (Rock et al., 2003). These data confirms that the feedback inhibition of PanK by product CoA *in vivo* is to limit the amount of CoA produced (Leonardi et al., 2005b). In addition, since CoA binding to PanK is competitive with ATP binding at the active site, the biosynthetic activity can also be coordinated with the energy level of the cell. A reduction in the ATP level would allow more feedback inhibitors to bind PanK, thus decelerates the rate of CoA biosynthesis (Leonardi et al., 2005b). Altogether, the CoA levels are controlled by the predominant CoA species and ATP levels in the cell (Vallari et al., 1987).

The eukaryotic PanKs are also feedback inhibited by CoA thioesters (Calder et al., 1999). Acetyl-CoA and palmitoyl-CoA inhibit the *Aspergillus nidulans* PanK strongly and selectively, in a competitive manner with ATP (Calder et al., 1999). The human PanK2 is very sensitive to long-chain acyl-CoA, acetyl-CoA and malonyl-CoA, with an IC₅₀ of 1 µM, while nonesterified CoA is reported to be much less effective, with an IC₅₀ of 60-75 µM (Kotzbauer et al., 2005).

Some novel type of PanK can not be inhibited by CoA and will be described in Chapter II.

1.5.3.2 Secondary regulation by CoaD

The release of 4'-phosphopantetheine from the bacteria to the outside medium suggests that 4'-phosphopantetheine adenylyltransferase (CoaD) is the secondary regulation point (Jackowski and Rock, 1984; Vallari and Jackowski, 1988). Regulation by this enzyme becomes more important when the regulation at the PanK site is disrupted or when the PanK protein is overexpressed (Rock et al., 2003; Song and Jackowski, 1992). The amount of intracellular and extracellular 4'-phosphopantetheine increases in both cases, indicating the restriction of the rate of CoA biosynthesis at the step catalyzed by CoaD. It was proposed that, like PanK, CoaD is feedback regulated by CoA as well. When purified from *E. coli*, CoA remains bound to CoaD enzyme (Geerlof et al., 1999), and a crystal structure of CoaD in complex with CoA shows that CoA binds CoaD at the 4'-phosphopantetheine site (Izard, 2003).

The mammalian CoA synthase that includes both CoaD and CoaE activities appears to be feedback inhibited by CoA and CoA thioesters (Rock et al., 2000). In mammalian cells the 4'-phosphopantetheine pool is almost as high as the pantothenate pool, and when PanK1 is overexpressed, the 4'-phosphopantetheine pool increases 3 fold (Leonardi et al., 2005b). More direct examination on the CoA synthase inhibition by different CoA species has not been done yet.

1.5.4 Regulation by gene expression

CoA biosynthesis is regulated not only by PanK enzymatic activity but also by *coaA* gene expression. In *E. coli* the CoaA protein is found at low abundance compare to the average bacterial proteins, because the *coaA* promoter has poor homology with the consensus *E. coli* promoter sequences and the *coaA* coding sequence uses low usage codons (Song and Jackowski, 1992). No evidence for *coaA* transcriptional control in bacterial has been reported, and bacteria like *E. coli* control CoA levels primarily by biochemical regulation (Song and Jackowski, 1992).

Mammalian cells and tissues modulate PanK expression to change CoA levels in the long-term response to diet and disease, such as starvation, alcoholism, diabetes and cancer (Wittwer et al., 1990). However, the mechanisms for these adaptive responses are still unknown.

1.5.5 Regulation by CoA degradation

CoA degradation is another way to modulate CoA levels. CoA can be dephosphorylated by a lysosomal phosphatase to produce dephospho-CoA, or hydrolyzed by cleavage of the phosphodiester bond to produce 4'-phosphopantetheine and 3',5'-adenosin mononucleotide (Bremer et al., 1972). In addition to the direct CoA degradation, the 4'-phosphopantetheine moiety of CoA can be transferred to carrier proteins, such as the acyl carrier protein (ACP) in

bacteria or the fatty acid synthase (FAS) in eukaryotes (Jackowski and Rock, 1984; Lambalot and Walsh, 1995).

Chapter II

Introduction to Pantothenate Kinase

2.1 Introduction

Chapter I highlights the central role of coenzyme A in metabolism, its biosynthesis and regulation. Since CoA is an indispensable cofactor in all living organism, enzymes involved in its biosynthesis become important research subjects. We are particularly interested in one enzyme in this pathway, pantothenate kinase (ATP:D-pantothenate 4'-phosphotransferase, also known as PanK or CoaA). PanK catalyzes the first committed step in CoA biosynthesis from pantothenate: the ATP-dependent conversion of pantothenate to 4'-phosphopantothenate (Rock et al., 2002; Song and Jackowski, 1992). Three distinct types of PanK have been characterized so far. This chapter focuses on different types of pantothenate kinase, giving an overview of how they differ from one another in terms of enzymetic characteristics and structure.

2.2 Type I Pantothenate Kinase

Pantothenate kinase (PanK or CoaA) was first identified in *Salmonella typhimurium* (Dunn and Snell, 1979) and *E. coli* (Vallari and Rock, 1987). *E. coli*

PanK has been exclusively characterized and is considered the prototypical prokaryotic PanK that is also referred to as type I PanK or PanK-I.

The *E. coli* PanK gene was cloned by functional complementation of the temperature sensitive mutant, in which the *coaA15(Ts)* mutation caused a temperature-sensitive growth phenotype and temperature-dependent inactivation of pantothenate kinase activity (Vallari and Rock, 1987). The *E. coli coaA* transcript has two translation initiation sites and the *E. coli* PanK protein was first isolated as a mixture of two peptides, with the shorter peptide lacking the first eight N-terminal residues (Song and Jackowski, 1992). Both peptides are active and the enzyme primarily exists as a homodimer (Song and Jackowski, 1994). There is one nucleotide binding site for each monomer, and ATP binding to the homodimer is highly cooperative (Song and Jackowski, 1994). The K_m for pantothenate and ATP were measured to be 36 μ M and 136 μ M, respectively. Kinetics studies indicated that the phosphorylation proceeds via an ordered sequential mechanism, with ATP binding first followed by pantothenate (Song and Jackowski, 1994).

PanK-I activity is inhibited, *in vivo* and *in vitro*, by CoA and less effectively by its thioesters (Vallari et al., 1987). Early studies have shown that the CoA inhibition is competitive with ATP and that both ligands bind to kinetically indistinguishable site on PanK (Vallari et al., 1987; Yun et al., 2000). The kinetics data is confirmed by the site-directed mutagenesis study. A

conserved lysine within the Walker A ATP binding motif (GX₄GKS) (this motif will be explained in section 2.5) was mutated to methionine to yield the PanK[K101M] mutant. The mutant kinase is catalytically inactive, and does not bind either ATP or CoA (Song and Jackowski, 1994).

The crystal structures of the *E. coli* PanK in complex with CoA or AMPPNP, a nonhydrolyzable ATP analog, has been reported (Yun et al., 2000). Their structures provide detailed insight into the binding of ATP and CoA to PanK. Superimposition of the PanK•CoA and PanK•AMPPNP structures show that the ADP moieties of the two ligands bind PanK at very different locations and orientations. Nonetheless, the α - and β -phosphates of CoA occupy the same location as the β - and γ -phosphates of AMPPNP near Lys¹⁰¹, which explains the competition between CoA and ATP for PanK binding. The thiol group of CoA fits tightly in a pocket defined by the aromatic residues Phe²⁴⁴, Phe²⁵², Phe²⁵⁹ and Tyr²⁶² (Yun et al., 2000). Therefore, the presence of an acyl chain in CoA thioesters like acetyl-CoA may cause steric hindrance to make them less potent than unacylated CoA in inhibiting PanK activity.

In addition, PanK•ADP•pantothenate ternary structure (Ivey et al., 2004) reveals the mechanism for the binding of another substrate pantothenate. The residues for pantothenate binding are highly conserved in bacterial PanK-I (Calder et al., 1999; Ivey et al., 2004). Superposition of the

PanK•ADP•pantothenate structure with PanK•CoA structure reveals that the substrates and the allosteric regulator are located within the same binding groove (Ivey et al., 2004). However, PanK has distinct confirmation when it binds to different substrates (Yun et al., 2000; Ivey et al., 2004).

Although *E. coli* PanK is considered the model bacterial pantothenate kinase, its homolog is not universally present in bacteria (Osterman and Overbeek, 2003). For example, the PanK from *Staphylococcus aureus* and the putative PanK from *Bacillus anthracis* are moderately related to the eukaryotic PanK proteins and unrelated to the *E. coli* PanK. The eubacteria *Pseudomonas aeruginosa* and *Helicobacter pylori* have a distinct type of PanK that is different from the classical prokaryotic and eukaryotic PanK.

2.3 Type II Pantothenate kinase

The second type of PanK (PanK-II) is found mainly in eukaryotes, including yeast, fungi, plants, and mammals (Calder et al., 1999; Rock et al., 2000). Interestingly, PanKs from a few gram-positive bacteria, such as *Staphylococcus aureus* (PanK_{Sa}) and several bacilli, also belong to this group based on sequence homology, although the bacterial enzymes exhibit certain catalytic characteristics different from their eukaryotic counterparts (Choudhry et al., 2003; Leonardi et al., 2005a).

The first eukaryotic PanK gene was identified in fungus *Aspergillus nidulans*, and this protein bears no sequence similarity to the *E. coli* enzyme (Calder et al., 1999). The *A. nidulans panK* gene is located on chromosome 3 and is interrupted by three small introns. This gene is expressed constitutively. The enzyme has a K_m of 60 μ M for pantothenate and 145 μ M for ATP (Calder et al., 1999). In contrast to *E. coli* PanK, which is feedback inhibited by CoA and to a less extent by its thioesters, *A. nidulans* PanK activity is selectively and potently inhibited by acetyl-CoA. The acetyl-CoA inhibition is also competitive with respect to ATP. Therefore, the eukaryotic (type II) PanK has a distinct primary structure and unique regulatory properties compared to the prokaryotic (type I) PanK (Calder et al., 1999).

The first mammalian PanK was discovered in mouse and was termed PanK1 (Rock et al., 2000). The predicted sequence of the murine PanK1 catalytic domain has significant homology with the *A. nidulans* PanK and is not related to *E. coli* PanK. The *mPanK1* gene consists of seven introns and eight exons and is located on chromosome 19 (19C2-3). Two isoforms exist, termed mPanK1 α and mPanK1 β , which share the same catalytic domain (exons 2 through 7) and different amino termini (exon 1). The difference in amino termini results in different regulatory properties on mPanK1. The exon 1 alpha encodes a regulatory domain at the amino terminus of PanK1 α that confers feedback

inhibition by CoA and acyl-CoA, and more potently inhibited by acetyl-CoA and malonyl-CoA (Rock et al., 2002), whereas the PanK1 β activity *in vitro* is not inhibited by CoA and weakly inhibited by acetyl-CoA (Rock et al., 2000). The differential regulation of two mPanK1 isoforms is physiologically significant, because the varied expression of PanK1 α and PanK1 β would alter the amount of CoA produced in cells as a function of the ratio of free CoA to acetyl-CoA, an indication of the metabolic status of the tissue (Rock et al., 2002).

The human *PANK1*, *PANK2*, *PANK3* and *PANK4* genes were first described in conjunction with the mapping of Hallervorden-Spatz syndrome (HSS) to the *PANK2* gene (Zhou et al., 2001). HSS is an autosomal recessive neurodegenerative disorder that is associated with iron accumulation in the brain. After identifying the linkage of mutations in *PANK2* gene with this disorder, HSS is also named as pantothenate kinase associated neurodegeneration (PKAN). However, *PANK2* mutations may not be associated with the onset of these degenerative conditions and the neurodegeneration with brain iron accumulation is genetically heterogeneous (Matarin et al., 2006).

Even though *S. aureus* is a bacterium, its PanK protein has a distinct primary sequence that does not resemble *E. coli* PanK but shows readily detectable homology to the mammalian PanK (Leonardi et al., 2005a). PanK_{Sa} exist as homodimer. The K_m value for pantothenate and ATP are 23 and 34 μ M,

respectively. The low K_m for ATP of PanK_{Sa} makes this kinase more reactive in the presence of lower intracellular ATP concentrations compared with *E. coli* enzyme whose K_m for ATP is 136 μ M (Leonardi et al., 2005a). Different from all the other known prokaryotic and eukaryotic PanKs, the PanK_{Sa} is refractory to feedback inhibition by CoA or its thioesters. The lack of regulation results in the accumulation of millimolar concentration of CoA in the cells. *S. aureus* produce CoA in proportion to the precursor in the medium, and there is no evidence for regulation at PanK_{Sa} or other downstream steps (Leonardi et al., 2005a; Leonardi et al., 2005b). This surprising phenomenon could be explained by the unique physiology of this organism that depends on CoA and a NADPH-dependent CoA reductase rather than glutathione to maintain the intracellular redox balance (delCardayre et al., 1998; Luba et al., 1999).

To understand the catalytic mechanism, Hong et al (Hong et al., 2006) solved the first crystal structure of PanK-II from *Staphylococcus aureus* in complex with AMPPNP. The structure of the human PanK2 is also available now (Hong et al., 2007; pdb code: 2i7n and 2i7p). The structural analysis of PanK-II will be discussed later in comparison with the structure of type III PanKs.

2.4 Type III Pantothenate Kinase

As discussed before, PanK is an essential gene product in all living organisms. However, neither PanK-I nor PanK-II can be found in a number of bacteria, including *Pseudomonas aeruginosa* and *Helicobacter pylori* (Gerdes et al., 2002). Since these organisms still possess the other genes (*coaBC*, *coaD*, and *coaE*) for CoA biosynthesis, it is likely that they contain a novel PanK that had not been characterized. Moreover, this new analogue has no or little sequence homology to the well-known type I and type II PanKs.

A putative third isoform of PanK was reported in a patent application (Yocum, 2002). A new gene sequence was identified from the *Bacillus subtilis* genome, and, when cloned in *trans*, can suppress the effects of an *E. coli coaA* temperature-sensitive mutant. Interestingly, *B. subtilis* already possess a gene homologous to the prokaryotic PanK-I, and interruption of this known *coaA* gene gave a normal-growing phenotype. It was speculated that the simultaneous deletion of both genes is lethal to *B. subtilis*. To distinguish it from the known *coaA* gene, the new gene was dubbed *coaX*.

Recently, Dr. Strauss group cloned, overexpressed and characterized CoaX from *Bacillus subtilis* and its homologue from *H. pylori* and show that they catalyze the ATP-dependent phosphorylation of pantothenate (Brand and Strauss, 2005). These enzymes do not share any sequence similarity with bacterial PanK-I, and is only very distantly related to PanK-II (Cheek et al., 2005). Through

homology searches based on the *coaX* gene sequence, the authors also identified CoaX orthologue in several bacterial genomes, including various pathogenic bacteria like *Bordetella pertussis* (the causative agent for whooping cough) and the category A biodefence pathogen, *Francisella tularensis*.

The *coaX* homologue in *B. pertussis* was originally annotated as related to Baf, a *Bvg* accessory factor, and was found to be an essential gene in this organism (DeShazer et al., 1995; Wood and Friedman, 2000). These studies attributed this protein to functioning in pertussis toxin production via interaction with the two-component transcriptional regulator BvgAS. However, Baf does not have any significant sequence homology to any of the known bacterial transcriptional regulators. In addition, later studies showed that these results were dependent on the components in the growth medium, which causes a doubt on the true nature of the BvgAS/Baf relationship (Bock et al., 2001). Since Baf protein shows 28% identity and 49% similarity to CoaX from *B. subtilis*, and it is essential for survival in *B. pertussis*, these evidences support its potential role in CoA biosynthesis rather than in a non-vital cellular process.

In comparison to the PanK-I and PanK-II, PanK-III (CoaX) exhibits different enzymatic characteristics (Brand and Strauss, 2005). First, PanK-III enzymes show a very low specificity for ATP, which is due to the surprisingly high K_m (nearly 10 mM in the case of *H. pylori* PanK-III). Second, unlike the

well-known bacterial PanK-I and eukaryotic PanK, the activity of PanK-III is not inhibited by either CoA or acetyl-CoA. Third, PanK-IIIs do not accept the pantothenate antimetabolite *N*-substitute pantothenamide as a substrate, and are not inhibited by it. Studies with type I PanK_{Ec} and type II PanK_{Sa} show that *N*-pantothenamide like *N*-pentyl- and *N*-heptylpantothenamide pantothenamide acts as antimicrobial agents through their function as CoA antimetabolites (Choudhry et al., 2003; Leonardi et al., 2005a; Strauss and Begley, 2002). These compounds act as substrates of the CoA biosynthetic enzymes in both *E. coli* and *S. aureus*, and they are converted to inactive CoA analogues, resulting in the inhibition of CoA-dependent cellular processes. The striking contrast of PanK-III with the other well-known PanKs declares it a distinct form of PanK that accounts for the only known activity for this kinase in many pathogens. The following chapters of this dissertation focus on the structural and biochemical analysis of PanK-III and its comparison with the type I and type II analogues. Our structural data provide solid explanation for the unique enzymatic characteristics listed above.

Finally, the discovery of the third type of PanK is a significant contribution to the understanding of how identical enzyme activity arises through convergent evolution. In addition, since genomic analysis suggests that this homologue is the only known pantothenate kinase in many pathogenic bacteria, the knowledge we gain from studying this type of PanK may have application in the development of specific antibacterial agents.

2.5 Structural Fold of Type I, II, and III PanKs

In order to investigate the relation between kinase structural fold and functional specifications, Cheek *et al* have done a comprehensive survey of all available kinase sequences (>17,000) and classified them into 30 distinct families based on homology (Cheek et al., 2002). Later a new survey was conducted in 2005 of 59,402 kinase sequences (Cheek et al., 2005). The kinase sequences were classified into a final 25 families of homologous proteins. Among them 22 families (about 98.8% of all sequences) contain known three-dimensional structures and fall into 10 different fold groups. For the remaining kinase families without a solved structure, fold predictions are made. The authors highlighted two novel kinase structural folds. Altogether, 12 fold groups were described (Table 2.1) (Cheek et al., 2005).

PanK-I is placed into the P-loop kinase family of the Rossmann-like fold group, which is the largest family in this group and also includes adenylate, thymidylate and shikimate kinases (Cheek et al., 2005; Yun et al., 2000). P-loop family contains one three-layered ($\alpha/\beta/\alpha$) domain, and the central parallel β -sheet is five-stranded with strand order 23145. The nucleotide binding region is distinguished by the presence of two conserved motifs, Walker A (GXXXXGKT/S) and Walker B (ZZZZD, where Z is any hydrophobic residue)

Table 2.1 Classification of kinase activity by fold group

Fold Groups	Examples of Kinase Activities
Group1: Protein S/T-Y kinase/ atypical protein kinase/ lipid kinase	Choline Kinase
Group 2: Rossmann-like	Pantothenate kinase Type I (P-loop kinase family)
Group3 : Ferredoxin-like fold kinases	Creatine kinase
Group4: Ribonuclease H-like	Pantothenate kinase Type II Pantothenate kinase Type III Hexokinase
Group5: TIM b/a-barrel kinases	Pyruvate kinase
Group6: GHMP kinase	Galactokinase
Group7: AIR synthetase like	Thiamine-phosphate kinase
Group8: Riboflavin kinase	Riboflavin kinase
Group9: Dihydroxyacetone kinase	Glycerone kinase
Group10: Putative glycerate kinase	Glyceratekinase
Group11: Polyphosphate kinase	Polyphosphate kinase
Group12: Integral membrane kinase	Dolicholkinase

(Saraste et al., 1990). The phosphate binding loop (P-loop), which is located at the end of the first β -strand and includes the first half turn of the following α -helix, is formed by Walker A motif and is found in a variety of different proteins that bind nucleotides, including type I pantothenate kinase (Yun et al., 2000). The PanK_{Ec} structure shows that AMPPNP is bound in a groove formed in part by the residues from the P-loop.

Type II and III PanKs were predicted to adopt a different fold pattern from PanK-I, due to the lack of sequence identity between these proteins and PanK-I in conjunction with the distinct predicted secondary structure pattern. Standard sequence similarity search methods failed to obtain any reasonable structure assignment. However, using state-of-the-art fold prediction methods, Cheek *et al.* (Cheek et al., 2005) have predicted that both these PanK types adopt an ribonuclease (RNase) H-like fold. This fold is composed of three layers ($\alpha/\beta/\alpha$), including a 5-stranded mixed β -sheet with strand order 32145, and strand 2 is antiparallel to the rest of the sheet. Importantly, the RNase H-like fold contains the acetate and sugar kinase/heat shock protein 70 (hsp70)/actin (ASKHA) superfamily (Bork et al., 1992; Hurley, 1996). The structures in this work and those by others have confirmed that both type II and type III PanKs belong to the ASKHA superfamily (Hong et al., 2006; Yang et al., 2006). A more detailed discussion will be present in the following chapters.

2.6 Objectives of the Dissertation

This dissertation is focused on type III pantothenate kinase which accounts for the only PanK activity in many bacteria, including some pathogens like *Helicobacter pylori* and *Pseudomonas aeruginosa*. When I first started this project, little was known about this new yet important type of PanK. Here, we aim to fully characterize this enzyme, and our strategy is two-folded: the first will address the biochemical characterization of PanK-III, and the second will focus on a detailed structural analysis of the enzyme. Altogether, these results will provide us with the necessary data to perform a comprehensive comparative characterization of the PanK analogues and to identify pathogen-specific inhibitors.

Aim 1. Understand the general fold of type III PanK (Chapter 3). We solved the first crystal structure from *Thermotoga maritima* (PanK_{Tm}) at 2.0 Å resolution. As the structure reveals, PanK-III indeed belong to the acetate and sugar kinase/heat shocks protein 70/actin (ASKHA) protein superfamily, similar to that of PanK-II. Comparative structural analysis of the PanK_{Tm} active site configuration suggested several aspartate residues as critical for PanK-III catalysis, which have been confirmed by the mutagenesis studies. Furthermore, the analysis also provides an explanation for the lack of CoA feedback inhibition

by the enzyme. Work described in this chapter has been published in Yang et al., 2006.

Aim 2. Understand the enzymatic properties of PanK-III. To fulfill this aim, we studied the thermodynamics characterization of substrate-enzyme interactions of PanK-III by isothermal titration calorimetry (Chapter 4), and the steady state properties and nucleotide specificity of PanK-III (Chapter 5).

Aim 3. Characterize the active site of PanK-III. We solved the crystal structures of PanK-III from *Thermotoga maritima* (PanK_{Tm}-III) complexed with substrate pantothenate and product phosphopantothenate, respectively, as well as a ternary complex of PanK_{Tm}-III with pantothenate and ADP. These structures revealed the detailed interactions between both substrates and the enzyme, shed new light into the catalysis and the unique kinetic properties of PanK-III, and should facilitate the structural based approach for developing specific inhibitors targeting PanK-III. Work described in chapters 4, 5 and 6 will be submitted.

Finally, our studies not only are important for understanding the fundamental metabolic pathways in PanK-III-harboring pathogenic bacteria and substrate specificity for PanK-III, but also provide a structural basis for inhibitor design (Chapter 7).

CHAPTER III

Crystal Structure of PanK-III from *Thermotoga Maritima* and Phylogenetic Distribution of PanK-III

Introduction

Pantothenate kinase (PanK; EC 2.7.1.33) catalyzes the ATP-dependent phosphorylation of pantothenate (vitamin B₅) to give 4'-phosphopantothenate. This reaction represents the first and committed step in the universal biosynthetic pathway of coenzyme A (CoA) (Begley et al., 2001; Leonardi et al., 2005b). Because CoA is a ubiquitous and essential cofactor in all organisms, genes coding for the five enzymes that make up this pathway—including PanK—are essential for their survival and growth (Begley et al., 2001).

Three distinct types of PanK, as differentiated by primary sequence analysis and kinetic properties, have been characterized so far. Type I PanKs (PanK-I) are found exclusively in eubacterial species and are exemplified by the *Escherichia coli* enzyme encoded by the *coaA* gene (Song and Jackowski, 1992, 1994). The second type of PanK (PanK-II) is found mainly in eukaryotes, including yeast and various fungi, plants, and mammals (Calder et al., 1999; Rock et al., 2000; Rock et al., 2002). Interestingly, PanKs from a few gram-positive bacteria, such as *Staphylococcus aureus* (PanK_{Sa}) and several bacilli, are also included in this group based on sequence homology, although the bacterial

enzymes exhibit certain catalytic characteristics different from their eukaryotic counterparts (Choudhry et al., 2003; Leonardi et al., 2005a). Recently, a third type of PanK (PanK-III) was identified which represents the only known pantothenate kinase activity in many pathogenic bacteria, including *Helicobacter pylori*, *Pseudomonas aeruginosa*, and *Bordetella pertussis*, as well as the category A biodefense pathogen *Francisella tularensis* (Brand and Strauss, 2005). Moreover, some bacteria, such as *Bacillus subtilis* and *Mycobacterium tuberculosis*, have genes that code for both PanK-I and PanK-III. To distinguish these from one another, the gene that codes for PanK-III was dubbed *coaX*, in contrast to *coaA* genes that produce PanK-I.

Substantial biochemical data have been accumulated for both PanK-I and PanK-II. These data show that although evolutionarily unrelated, both types of PanK are feedback inhibited by the end product of the pathway, CoA, as well as its thioesters, although the extent of inhibition depends on the system and the specific inhibitor (Calder et al., 1999; Rock et al., 2000; Rock et al., 2002; Song and Jackowski, 1994; Vallari and Rock, 1987; Yun et al., 2000). This feedback inhibition of PanK activity by CoA and its derivatives represents a key regulatory mechanism that controls intracellular CoA levels in response to a cell's metabolic status (Leonardi et al., 2005b). One exception to this observation is the PanK enzyme from *S. aureus*, which is not inhibited either by CoA or its thioesters, most probably due to this organism's unique redox biology that depends on high

concentrations of CoA and an NADPH-dependent CoA disulfide reductase to maintain its intracellular redox balance (delCardayre et al., 1998; Leonardi et al., 2005a; Luba et al., 1999).

In contrast to the large body of data gathered on type I and II PanKs, relatively little is known about the mechanism and regulation of PanK-III. Current knowledge does, however, clearly indicate a unique position for these enzymes among PanKs: while their k_{cat} and K_m values for pantothenate are comparable to those of PanK-I and -II, they exhibit an unusually high (in the mM range) K_m for ATP, a 30- to 100-fold increase over the other types (Brand and Strauss, 2005). Furthermore, unlike other PanKs the type III enzymes are not inhibited by CoA or any of its thioesters, a characteristic that might be singularly significant to organisms that predominantly harbor this type of PanK.

Numerous studies have identified the potential of CoA biosynthetic enzymes as targets for drug development. Most recently, a comprehensive in vivo analysis of *Salmonella enterica* has highlighted the five enzymes of CoA biosynthesis to be among those previously known but as yet unexploited antimicrobial targets of important human pathogens (Becker et al., 2006). This analysis is based on the essential requirement of these enzymes for survival and/or virulence and on the lack of homology between bacterial PanK-I enzymes and their mammalian PanK-II counterparts. Development of inhibitors targeting these enzymes is being actively pursued (Choudhry et al., 2003; Gerdes et al., 2002;

Strauss and Begley, 2002; Virga et al., 2006; Zhang et al., 2004; Zhao et al., 2003). Among these, the N-substituted alkylpantothenamides have shown the greatest promise as growth inhibitors of both *E. coli* and *S. aureus* (Choudhry et al., 2003; Leonardi et al., 2005a; Strauss and Begley, 2002; Zhang et al., 2004). These compounds act as alternative substrates of PanK and two other CoA biosynthetic enzymes, allowing their conversion to CoA analogs that subsequently inhibit CoA- and acetyl-CoA-utilizing enzymes and inactivate proteins with CoA-derived prosthetic groups, such as the acyl carrier protein (ACP) (Leonardi et al., 2005a; Strauss and Begley, 2002; Zhang et al., 2004). Importantly, PanK-III enzymes are not affected by the *N*-alkylpantothenamide family of inhibitors; neither do they accept these compounds as alternative substrates (Brand and Strauss, 2005). A structural characterization of PanK-III active site configuration will, therefore, greatly facilitate the development of inhibitors targeting this type of PanK.

Such structure-based drug development strategies targeting PanK-I are already possible due to the availability of the three-dimensional structure of *E. coli* PanK-I protein (PanK_{Ec}) (Ivey et al., 2004; Yun et al., 2000) and have allowed a structure-activity relationship (SAR) analysis to be performed on the pantothenamide-type inhibitors of this enzyme (Virga et al., 2006). While the sequence and structure of PanK-I indicate that it belongs to the "P-loop kinase" superfamily (Cheek et al., 2005; Cheek et al., 2002), no structures of any PanK-II

or PanK-III were previously known. Using state-of-the-art fold prediction methods, we have predicted that both these PanK types adopt an RNase H-like fold (Brand and Strauss, 2005; Cheek et al., 2005) and are distantly related to the acetate and sugar kinase/heat shock protein 70 (hsp70)/actin (ASKHA) superfamily (Bork et al., 1992; Hurley, 1996). Because ASKHA and P-loop superfamilies belong to two different protein folds, presumably PanK-II and PanK-III will have completely different active site architectures from that of PanK-I.

To verify our fold predictions and to address the general lack of knowledge of type III PanKs, we have determined the crystal structure of the PanK-III enzyme from *Thermotoga maritima* (PanK_{Tm}) at a 2.0 Å resolution. The structure confirms that PanK-III belongs to the ASKHA superfamily, which allowed us to identify its active site and to model the interactions between the substrates and the active site residues by comparison with other members of this superfamily. Based on this model, mutagenesis and kinetic analysis of highly conserved aspartate residues were carried out to investigate their roles in catalysis. Finally, we provide a comprehensive survey of the phylogenetic distribution of all three types of PanK and show that PanK-III has a much wider distribution in the bacterial kingdom than originally anticipated. Taken together, these results add significantly to our current knowledge of this key metabolic enzyme.

Materials and Methods

Construction of Expression Plasmid

The *Thermotoga maritima* *coaX* gene (accession no. NC_000853, region: 905791...906531) was amplified from *T. maritima* genomic DNA (ATCC 43589D) by the PCR using Pfu DNA polymerase and the following Primers: 5'-GGTGGATCCATGTACCTCCTCGTGGAC-3' (forward primer), introducing a *Bam*HI site (underline) at the start of the gene, and 5'-GAACTCGAGTCAATCTCCGAAGCAG-3' (reverse primer), introducing an *Xho*I site (underline) at the end of the gene. The resulting PCR product was digested with BamHI and XhoI and cloned into the BamHI/XhoI-digested pProEX-HTa expression vector (Invitrogen, Carlsbad, CA) containing a *trc* promoter, N-terminal 6xHis-tag, and a tobacco etch virus (TEV) protease cleavage site. The *Bacillus subtilis* *coaX* gene was subcloned into pProEx-HTa vector (Invitrogen) from pET28a-*Bs*CoaX that was described before (Brand and Strauss, 2005). The *Mycobacterium tuberculosis* *coaX* was amplified from the *M. tuberculosis* genomic DNA by PCR and cloned into pProEx-HTa expression vector. The pET28a-*Hp*CoaX and pET28a-*Pa*CoaX plasmids are gifts from our collaborator Dr. Erick Strauss (Department of Chemistry, Stellenbosch University, Matieland, South Africa).

Protein Expression and Purification

The pProEx-*Tm*CoaX plasmid was transformed into *E. coli* strain BL21 (DE3) for protein expression. Cells were grown in liquid Luria-Bertani (LB) medium containing 100 µg/ml ampicillin at 37°C until OD₆₀₀ reached 0.6 and induced with 0.8mM isopropyl-1-thio-β-D-galactopyranoside (IPTG). Growth was continued overnight at 20°C. The cells were harvested by centrifugation and were frozen at –80°C. The thawed cells were resuspended in the lysis buffer (20 mM imidazole, 0.1 M NaCl, 20 mM HEPES, pH 8.0, 0.03% Brij-35, 5mM β-mercaptoethanol, 2mM PMSF, and protease inhibitors cocktail (Sigma-Aldrich, St. Louis, MO) and were passed twice through a high-pressure homogenizer (AVESTIN Inc., Ottawa, Canada). The clarified cell extract was then loaded on to a nickel-nitrilotriacetic acid-agarose column (Qiagen, Valencia, CA), and eluted with a gradient of imidazole (0-250 mM). The N-terminal 6xHis tag was removed by treatment with TEV protease produced in house using a vector kindly provided by Dave Waugh (NCI, Frederick, MD. (Kapust and Waugh, 1999), followed by purification of the cleaved protein with ion exchange chromatography on a Resource Q column (GE Healthcare Life Sciences, Piscataway, NJ). The selenomethionine-substituted PanK_{Tm} protein was expressed in minimum medium supplemented with selenomethionine and other nutrients following standard protocols (Doublié, 1997) and was purified as described above. The purification of *Hp*CoaX and *Pa*CoaX are the same as above.

Crystallization and Data Collection

Numerous effects have been made to crystallize type III PanK from 5 organisms: *Helicobacter pylori*, *Pseudomonas aeruginosa*, *Thermotoga maritime*, *Bacillus subtilis* and *Mycobacterium tuberculosis* (Table 3.1). The PanK-III from the first three organisms could be expressed, but not the other two. Next, crystallization of type III PanK protein from *Helicobacter pylori* was tried first, and the resolution was 3.5 Å (Figure 3.1 and 3.2). To improve the diffraction of the crystal, first, different cryoprotectants (glycerol, MPD, PEG-400 and sugar) were tried, but there was no significant improvement. Second, annealing was pursued, but the resolution was not increased. Third, according to our experience, crystal of native protein without any tag can give better resolution than that of a fusion protein. Therefore, we decided to remove the his-tag before crystallization. Unfortunately, no crystal grew after his-tag being removed. Next, crystallization of type III PanK from other organisms was pursued, including *Pseudomonas aeruginosa* (Figure 3.3). However, the resolution of PanK_{Pa} with his tag was only 8 Å. After removing the his tag, the resolution of the crystal was improved to 2.9 Å. Finally, a crystal of *Thermotoga maritime* PanK-III were obtained that gave diffraction data of sufficient quality (at 2 Å) to allow the structural analysis (Figure 3.4 and 3.5).

Crystals of PanK_{Tm} were grown at 4°C using sitting drop vapor diffusion method. Drops containing 1.5 µl of PanK_{Tm} protein (concentration 18 mg/ml in 20 mM HEPES, pH 8.0, and 200 mM NaCl) mixed with an equal volume of the

Table 3.1 Summary of PanK-III expression, crystallization and diffraction from five different organisms

Organisms	Expression	Crystallization	Crystal Resolution (Å)
<i>Helicobacter pylori</i>	+	+	3.5
<i>Pseudomonas aeruginosa</i>	+	+	2.9
<i>Thermotoga maritima</i>	+	+	2.0
<i>Bacillus subtilis</i>	-	N/A	N/A
<i>Mycobacterium tuberculosis</i>	-	N/A	N/A

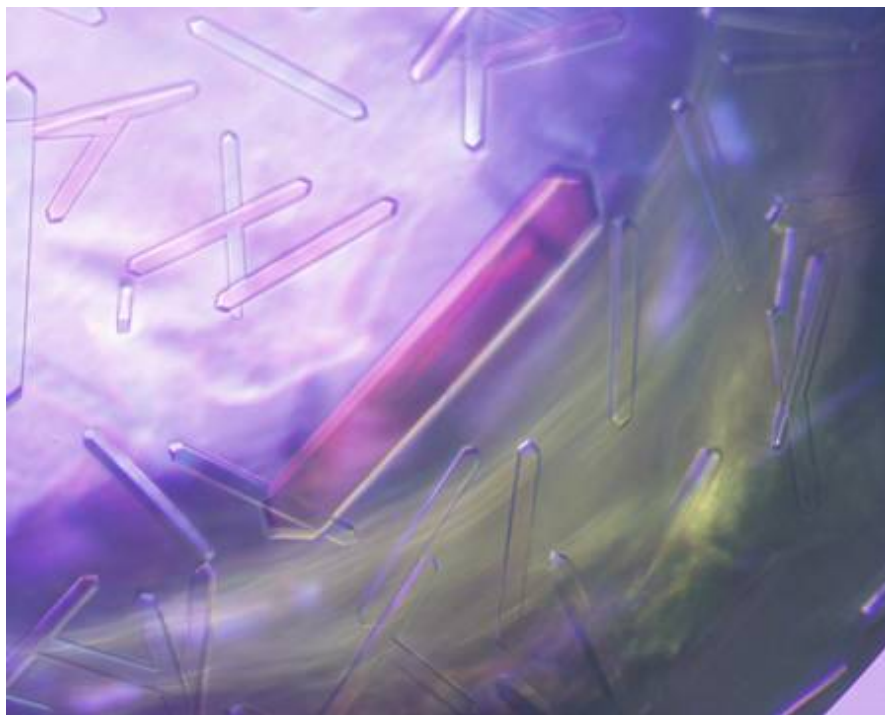


Figure 3.1 Crystal of Native PanK_{HP}

The crystals were grown using hanging drop vapor diffusion method at 20 °C. The optimized condition contains 0.8 M sodium citrate, 0.1 M CHES (pH=10), 0.2 M MgCl₂. The protein concentration was 20 mg/ml.

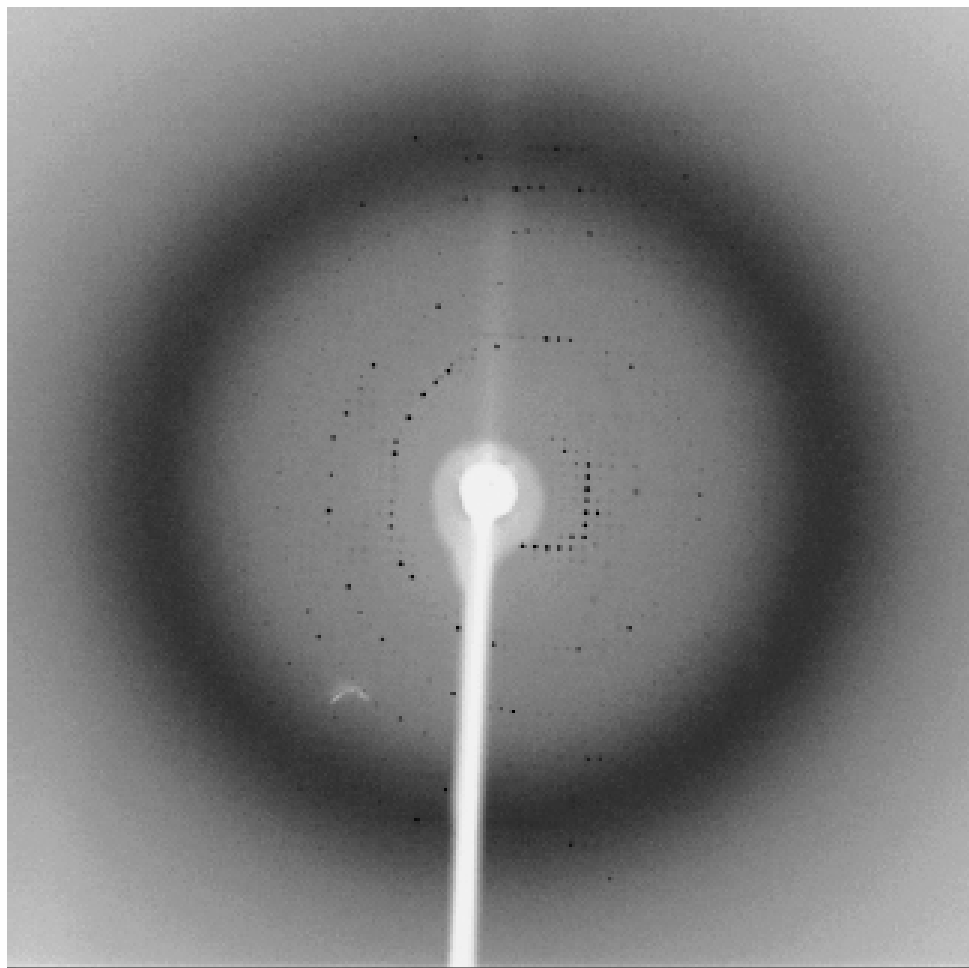


Figure 3.2 X-ray Diffraction Pattern of a PanK_{Hp} Crystal

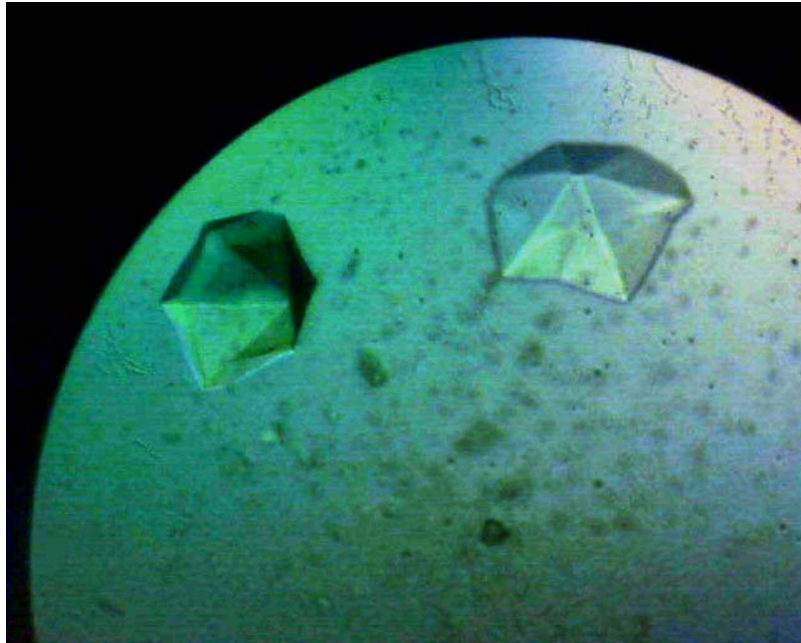
Diffraction data was collected in-house on an RAXIS-IV⁺⁺ image plate detector equipped with a Rigaku FR-E SuperBright X-ray generator and VarimaxTM HF mirrors.

Resolution: 3.5 Å.

Unit cell:

$a=b=108.309$ Å, $c=105.358$ Å. $\alpha=\beta=\gamma=90^\circ$.

A



B



Figure 3.3 Crystals of PanK_{Pa}

A. Crystals of the native PanK_{Pa} with His₆ tag

The crystals were grown using hanging drop vapor diffusion method at 20°C. The optimized condition contains 1.2M K/Na tartrate, 0.1 M Imidazole pH 7.3, 0.2M NaCl. The protein concentration was 20 mg/ml.

B. Crystals of the native PanK_{Pa} without His₆ tag

The crystals were grown using hanging drop vapor diffusion method at 20°C. The optimized condition contains 30% PEG400, 0.1M Tris pH 8.2, and 0.2M MgCl₂. The protein concentration was 20 mg/ml.

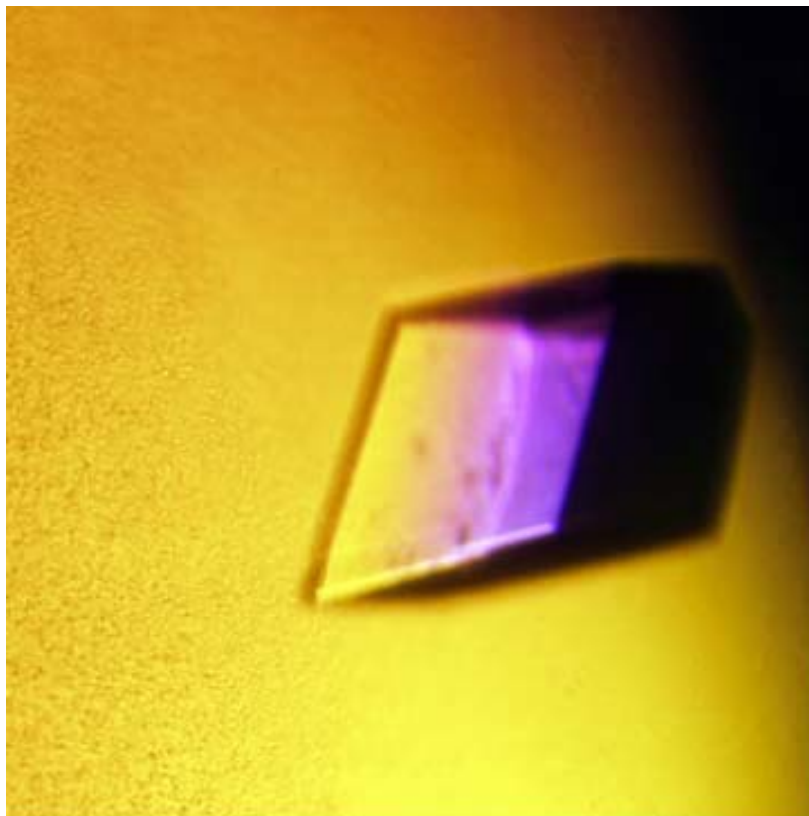


Figure 3.4 Crystal of PanK_{Tm}

The crystals were grown using sitting drop vapor diffusion method at 4°C. The optimized condition contains 15% PEG 3350. The protein concentration was 18 mg/ml.

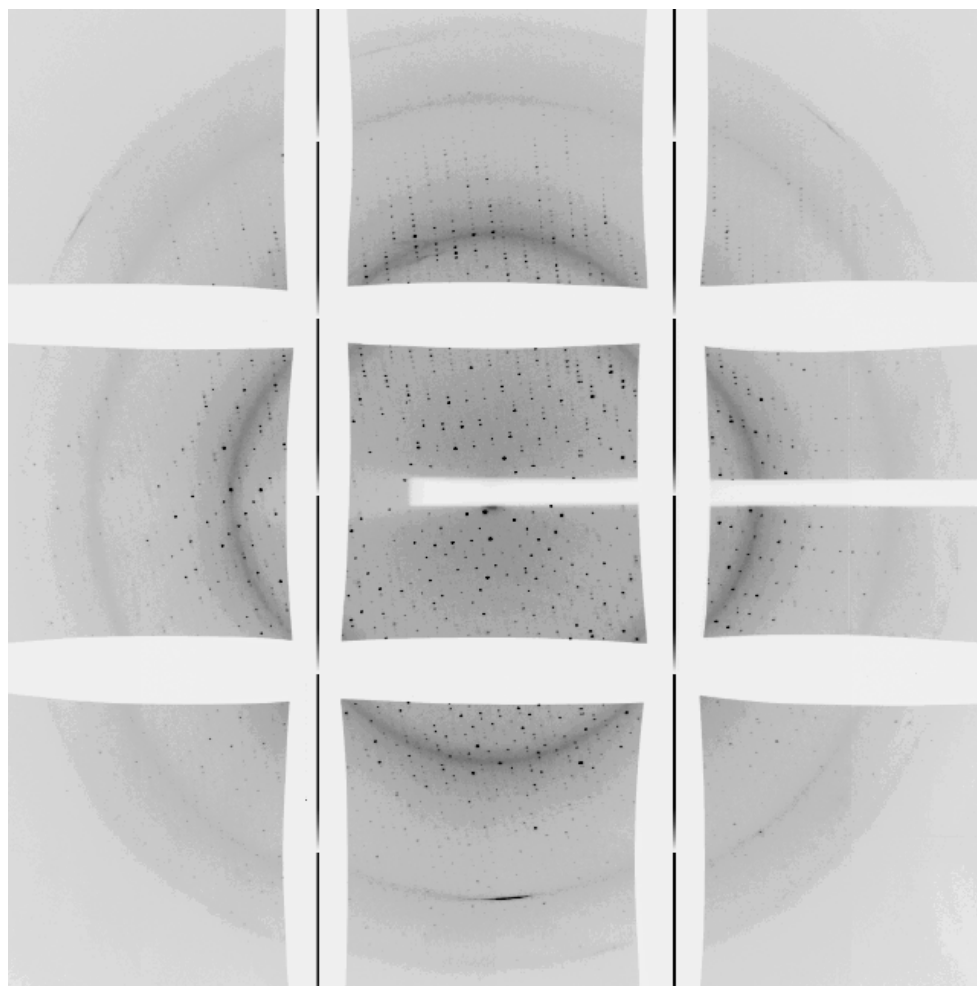


Figure 3.5 X-ray Diffraction Pattern of a PanK_{Tm} Crystal

Diffraction data was collected at beamline 19-BM at the Advance Photon Source, Argonne National Laboratory (Argonne, IL).

Resolution: 2.0 Å.

Unit cell:

$a = 75.11 \text{ Å}$, $b = 137.79 \text{ Å}$, $c = 75.22 \text{ Å}$, and $\beta = 109.22^\circ$.

reservoir solution containing 15% PEG-3350 were equilibrated against the reservoir over a period of several days. PanK_{Tm} crystals of sizes around 0.2-0.4 mm appeared typically within 1 week. Prior to data collection, crystals were transferred sequentially to the cryoprotectant solutions containing 20 mM HEPES, pH 8.0, 100 mM NaCl, and additional PEG-3350 at a final concentration of 15%, 25%, and 35% before flash-freezing in liquid propane. Diffraction data were collected at the beamline 19-BM at the Advanced Photon Source (Argonne National Laboratory, Argonne, IL). The diffraction data were indexed, integrated, and scaled using the HKL2000 program package (Otwinowski and Minor, 1997). The crystals belong to the primitive monoclinic space group P2₁ with cell dimensions $a = 75.11 \text{ \AA}$, $b = 137.79 \text{ \AA}$, $c = 75.22 \text{ \AA}$, and $\beta = 109.22^\circ$.

Structure Determination and Refinement

The initial phases were obtained by the multiwavelength anomalous dispersion (MAD) phasing method from a crystal of the selenomethionyl variant of PanK_{Tm}. An X-ray fluorescence scan of a selenomethionyl PanK_{Tm} crystal was conducted near the absorption edge (K-edge) of selenium. The MAD data were collected at two wavelengths that correspond to the peak and inflection points of the K-edge of selenium. Twenty-eight selenium sites were located by the program SHELXD (Schneider and Sheldrick, 2002). Refinement of the heavy atom parameters and phase calculation were performed using the program MLPHARE

in the CCP4 package. Initially a fivefold noncrystallographic symmetry was identified by program RESOLVE (Terwilliger, 2000, 2003) from the 28 Se sites. Density modification including a fivefold molecular averaging was subsequently carried out using RESOLVE. The resulting map was of excellent quality (Figure 3.6) and revealed that there are actually a total of six PanK_{Tm} monomers in the asymmetric unit. The majority of the model was automatically built by RESOLVE and was completed manually using the O program (Jones et al., 1991). The crystallographic refinement was carried out with Refmac5 (Murshudov et al., 1997) of the CCP4 package. The current model contains six PanK_{Tm} monomers, each from residue 1 to residue 245 (of total 246 residues) plus 3 additional residues, Met⁻²-Asp⁻¹-Pro⁰, which are introduced upstream of the first methionine during cloning. There are thus a total of 1,488 protein residues and 1,366 water molecules in the current model. The crystal data and refinement statistics are listed in Table 3.2.

Modeling of the substrates in the PanK_{Tm} active site

The structural superposition of PanK_{Tm} with the following ASKHA proteins was performed manually using the O program (Jones et al., 1991): 2-hydroxyglutaryl-CoA dehydratase component A (Protein Data Bank [PDB] code 1hux) (Locher et al., 2001), acetate kinase (1g99) (Buss et al., 2001), the C-terminal half of human hexokinase (1d9k) (Aleshin et al., 2000), hsp70 (1ba1)

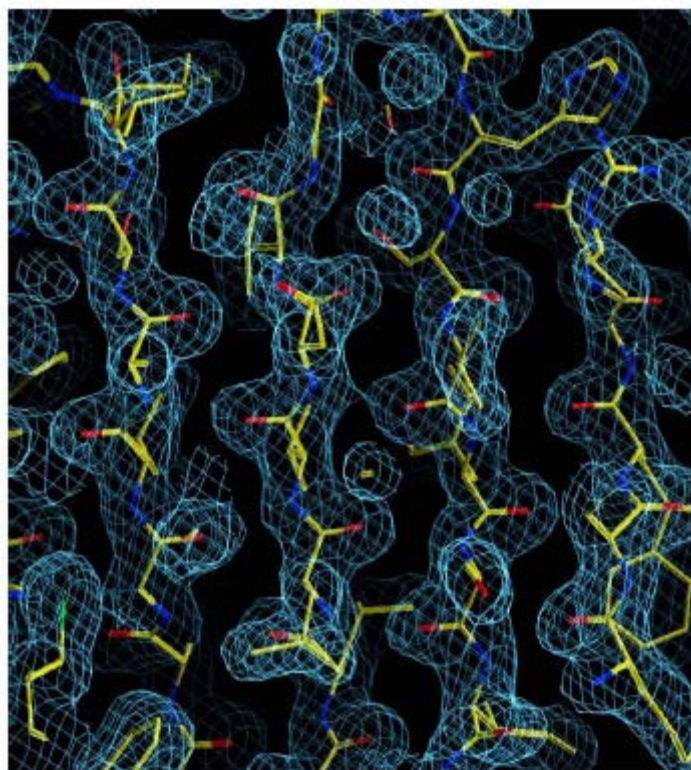


Figure 3.6 Electron density map after density modification of PanK_{Tm}.

Table 3.2 Data collection, phasing, and refinement statistics for native crystal of PanK_{TM}

Data Statistics	SeMet peak ^a	SeMet inflection
Wavelength (Å)	0.97872	0.97886
Resolution (Å)	50-2.00 Å	50-2.00 Å
Total observations	377,376	367,288
Unique reflections	95,595	95,673
Completeness (outer shell)	97.7% (93.3%)	96.3% (88.7%)
R_{sym} (outer shell) ^b	0.077(0.226)	0.097 (0.881)
I/σ (outer shell)	28.78 (5.59)	21.46 (2.10)
Figure of merit	0.78	
Refinement		
Resolution range (Å)	30-2.00 Å	
R_{work} ^c	17.7(%)	
R_{free} ^d	24.8(%)	
Protein atoms (Avg. B factor)	11574(22.78)	
Solvent atoms (Avg. B factor)	1345(39.41)	
R.m.s.d. bond length	0.012 Å	
R.m.s.d. bond angle	1.338°	
Ramachandran Plot		
Most favored region	92.8(%)	
Additional allowed region	7.1(%)	
Disallowed region	0.1(%)	

^a Bijvoet pairs were treated as equivalent reflections during data processing.

$$^b R_{\text{sym}} = \sum_{hkl} \{ (\sum_j |I_j| - \langle I \rangle) \sum_j |I_j| \}$$

^c $R_{\text{work}} = \sum_{hkl} |F_o - F_c| / \sum_{hkl} |F_o|$, where F_o and F_c are the observed and calculated structure factors, respectively.

^d Five percent of the reflections were used in the calculation of R_{free} .

(Wilbanks and McKay, 1998), and glycerol kinase (1glc) (Hurley et al., 1993). Each of the two domains of these template structures was superimposed separately onto the corresponding domain of PanK_{Tm} guided by the multiple sequence alignment. In general, when the second or the C-terminal domains are superimposed, the bound ATP or ADP in the templates falls into the cleft of the PanK_{Tm} active site without any serious steric clash. An ATP molecule was then placed in the PanK_{Tm} active site at the consensus position for ATP that is commonly shared among all members of the superfamily. The pantothenate molecule taken from the complex structure of *E. coli* PanK (PanK_{Ec}) with ADP and pantothenate (PDB code 1sq5) (Ivey et al., 2004) was placed in the general location of the second substrate of the ASHKA superfamily. The following assumptions were made during the modeling. Assumption 1 was that the Asp6 side chain directly coordinates an Mg²⁺ ion which would interact with the β - and γ -phosphates of ATP. Assumption 2 was that Asp105 acts as the catalytic base and is within hydrogen bond distance from the 4-hydroxyl group of pantothenate. Assumption 3 was that the two C3 methyl groups of pantothenate would fit into a small hydrophobic pocket formed by residues Ala129, Ile145, and Leu163' from the second monomer of the dimer. This manually docked model was then energy minimized using the CNS program (Brunger et al., 1998).

Protein structure accession number

Coordinates of PanK_{Tm} have been deposited in the Research Collaboratory for Structural Bioinformatics (RCSB) Protein Data Bank under accession code 2GTD.

Results and Discussions

Description of PanK_{Tm} monomer structure

The crystal structure of PanK_{Tm} was solved using the MAD phasing method and refined to a 2.0 Å resolution. Six PanK_{Tm} monomers are found in the crystallographic asymmetric unit of the PanK_{Tm} crystal. The enzyme's gel filtration profile indicates that PanK_{Tm} exists in solution as a dimer, which is likely the functional unit of PanK_{Tm}. The hexameric appearance of PanK_{Tm} in the crystal could be a consequence of the given crystallization conditions.

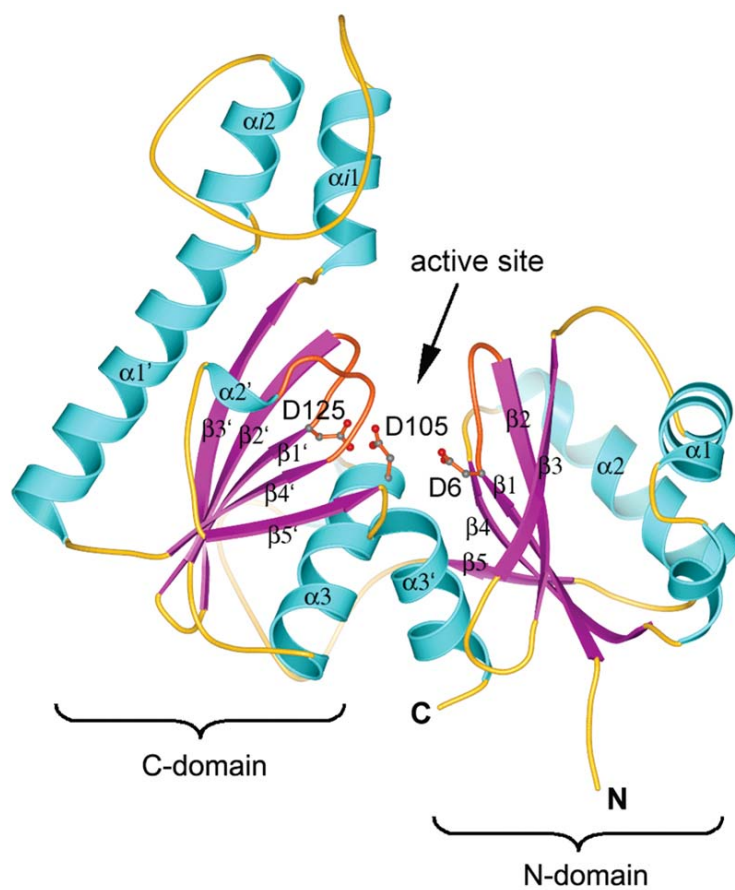
The structure of the PanK_{Tm} monomer confirmed our previous prediction that it adopts the "RNase H-like fold" as classified in the SCOP database (Andreeva et al., 2004; Murzin et al., 1995). More specifically it belongs to the "actin-like ATPase domain" superfamily of proteins, often referred to as the ASKHA (acetate and sugar kinase/hsp70/actin) superfamily (Bork et al., 1992; Hurley, 1996). Proteins in this superfamily include the ATPase domain of actin

(Holmes et al., 1990; Kabsch et al., 1990) and hsp70 (Flaherty et al., 1990), acetate kinase (Buss et al., 2001), several sugar kinases, such as hexokinase (Rosano et al., 1999; Steitz et al., 1976), glycerol kinase (Hurley et al., 1993), ATP- and ADP-dependent glucokinases (Ito et al., 2003; Lunin et al., 2004), as well as 2-hydroxyglutaryl-CoA dehydratase component A (2-HG-CoA dehydratase Comp A) (Locher et al., 2001). Similar to all other members of the ASKHA superfamily, the PanK_{Tm} monomer contains two domains that have the same fold and are considered to be a result of gene duplication (Figure 3.7A). The core of each domain consists of a five-stranded mixed β -sheet with strand order 3-2-1-4-5, where strand 2 is antiparallel to the rest of the sheet. The topology of the core of this fold is $\beta_3\beta_2\beta_1\alpha_1\beta_4\alpha_2\beta_5\alpha_3$, with the first helix α_1 following strand β_3 .

Two of the helices (α_1 and α_2) are located on one side of the β -sheet, while the third helix (α_3) is on the other side. Notably, the last helix α_3 interacts intimately with the β -sheet of the C-terminal domain and should be considered as part of the C-terminal domain, while the corresponding helix near the C terminus (α_3') actually contributes to the core of the N-terminal domain (Figure 3.7A). This structural arrangement is characteristic of all members of the ASKHA superfamily.

The nucleotide binding and divalent metal coordination are achieved by several motifs conserved within the ASKHA superfamily. These motifs are the

A



B

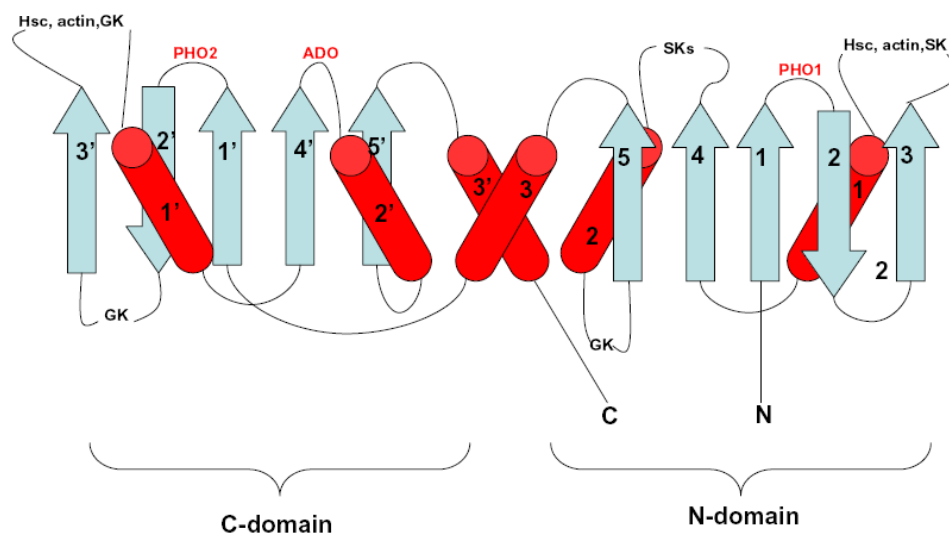


Figure 3.7 Structure of PanK_{Tm} monomer

A. Ribbon diagram of PanK_{Tm} monomer. The corresponding secondary structure elements in the two duplicate domains (N- and C-terminal domains) are labeled $\beta 1$ to $\beta 5$ and $\beta 1'$ to $\beta 5'$ for β -strands and $\alpha 1$ to $\alpha 3$ and $\alpha 1'$ for $\alpha 3'$ for α -helices, respectively. The two helices $\alpha i1$ and $\alpha i2$ between $\beta 3'$ and $\alpha 1'$ are considered to be insertions to the core of the fold. Three highly conserved aspartate residues, Asp6, Asp105, and Asp125, are shown in the ball-and-stick representation. The loop regions corresponding to the three conserved motifs PHOSPHATE 1 (between strands $\beta 1$ and $\beta 2$), PHOSPHATE 2 (between strands $\beta 1'$ and $\beta 2'$), and ADENOSINE (between strands $\beta 4'$ and $\alpha 2'$) are colored orange.

B. Schematic of the topology of the conserved core subdomains in the ASKHA (acetate and sugar kinase/heat-shock protein 70/actin) superfamily. The C-domain and N-domain represent an internal duplication of the RNase H-like fold. The red helix 3 and 3' cross over to form the base of an interdomain cleft. Three conserved motif are labeled in red (PHO1: phosphate motif 1; PHO2: phosphate motif 2; ADO: adenosine). GK: glycerol kinase; SK: sugar kinases; Hsc: heat-shock cognate protein. All four proteins have unique extensions at either their N- or C-termini, or both. The unique biological functions are mediated by subdomains inserted at four different topologic positions.

ADENOSINE motif that interacts with the ribosyl and the α -phosphoryl group of ATP, the PHOSPHATE 1 motif that interacts with Mg^{2+} through coordinated water molecules, and the PHOSPHATE 2 motif that interacts with the β - and γ -phosphoryl groups of ATP.

All the proteins in the ASKHA superfamily contain a conserved core domain that contains two α/β subdomains (topology: $\beta\beta\alpha\beta\alpha\beta\alpha$) related to each other by approximated dyad symmetry (Hurley, 1996). The structure of each member in the superfamily shows a distinct pattern of insertions at several different loci within the conserved fold. The two core subdomains together with the insertions make up the two domains of each structure (Figure 3.7B). ATP binds in a deep cleft between the two domains, and the metal ion that is catalytically essential binds to the ATP phosphates directly and to the enzyme via water-mediated interactions (Hurley, 1996).

The search for similar structures in the protein data bank using program DALI (Holm and Sander, 1995) returns as its top hit 2-HG-CoA dehydratase Comp A (PDB code 1hux) with a Z-score of 14.5 and overall root mean square deviation (rmsd) of 202 superimposed C_{α} s of 3.7 Å (Table 3.3). The structural similarities of PanK_{Tm} to other members of the ASKHA superfamily are also high as shown in Table 3.3. Clearly, PanK-III proteins share the same fold as the ASKHA proteins and are likely to have evolved from the same ancestral protein as

Table 3.3 Proteins structurally similar to PanK_{Tm} and structural alignment statistics from DALI^a

^a For details, see reference (Holm and Sander, 1995).

Protein (pdb code – chain ID)	Z-score	Rmsd (Å)	Aligned C_α's	Sequence identity (%)	Total number of residues in protein
HG-CoA DH CompA (1hux-A)	14.5	3.7	202	16	259
FtsA (1e4f-T)	14.0	3.9	216	10	378
Glycerol kinase (1glc-G)	13.2	3.7	222	13	489
Acetate kinase (1g99-A)	13.1	3.6	220	11	398
Hexokinase (1qha-A)	11.8	4.2	226	11	903

the rest of the superfamily. Notably, the sequence identities between PanK-III and the rest of the family are low (17% and below) and the rmsd's between the superimposed C_{α} atoms are quite large, ranging from 3.6 Å to 5.0 Å. To some extent, this large structural deviation is due to the difference in relative orientations between the two domains of the structures, but it also reflects the large evolutionary distance between PanK-III and the rest of the ASKHA superfamily.

In PanK_{Tm}, the only deviation from the minimum core of the RNase H-like fold is a small insertion between strands β_3' and α_1' of the C-terminal domain, which consists of a pair of antiparallel helices and a long connecting loop (Figure 3.7). Insertion at this particular site has been universally observed in all members of the ASKHA superfamily characterized so far. While many members in the superfamily are extensively decorated by long insertions at various sites of the RNase H-like fold core, PanK_{Tm} and 2HG-CoA dehydratase Comp A (1hux) appear to contain minimally inserted elements aside from the core.

Structure of PanK_{Tm} dimer

PanK_{Tm} forms a tight dimer in the crystal as well as in solution, as demonstrated by the size exclusion chromatography profile (data not shown). The helical insertion (α_{i1} and α_{i2}) between β_3' and α_1' along with helix α_1' form an extensive dimer interface with the corresponding region of the second monomer

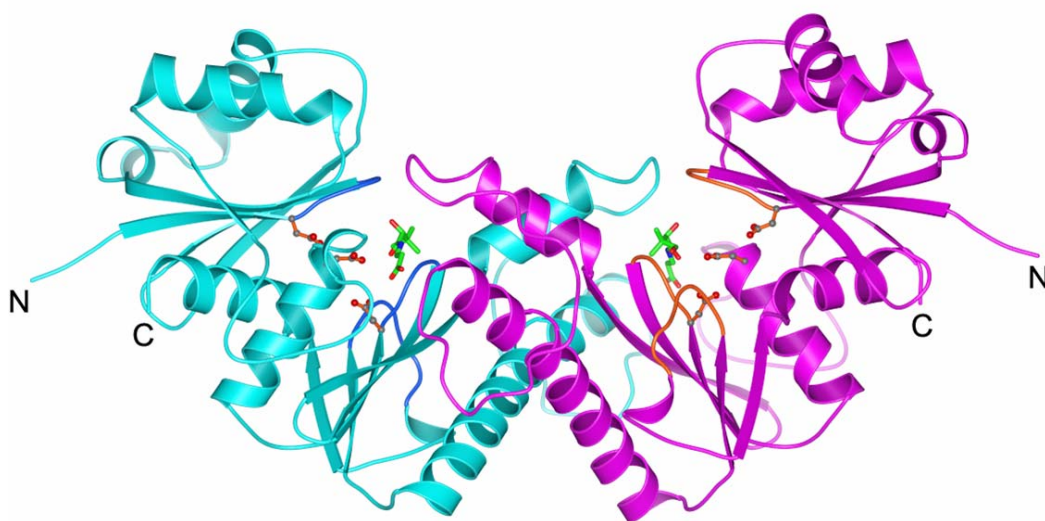


Figure 3.8 Ribbon diagram of PanK_{Tm} dimer

The two monomers are colored cyan and magenta, respectively. The active site of each monomer is marked by the ball-and-stick representation of the conserved aspartate residues. Modeled pantothenate (in green; see "Materials and Methods" for details) is also shown to indicate its location near the dimer interface.

(Figure 3.8). This dimer interface buries about a 2,034 Å² surface area and is largely hydrophobic in nature. The conformations of each monomer in the dimer are very similar, with an average rmsd in C_α positions of 0.2 to 0.3 Å even though no noncrystallographic symmetric restraints are imposed during refinement.

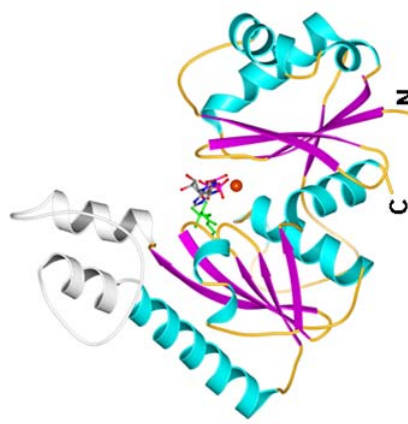
Several other members of the ASKHA superfamily also form dimers through insertions located at the same site (between β₃' and α₁') and through a helix corresponding to the α₁' helix of the second domain, such as those observed in 2-HG-CoA dehydratase Comp A (Locher et al., 2001) and acetate kinase (Buss et al., 2001). However, the details of this interface differ, and the relative orientations of the two monomers are also different between PanK-III and the other members of the superfamily. As will be discussed later, the proposed pantothenate binding site in PanK_{Tm} contains residues from both monomers of the dimer, strongly suggesting that the dimer is the functional unit for PanK_{Tm} and probably for all other PanK-III enzymes as well.

The active site of PanK-III

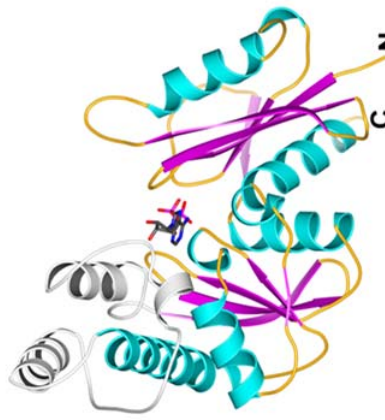
Several structures in the ASKHA superfamily have been solved in the presence of the bound substrates, those being either the ATP (or ADP) nucleotide or the phosphoryl acceptor substrate. In glycerol kinase and hexokinase, the ternary complexes with ADP and the phosphoryl acceptor substrate (or product in glycerol kinase) were also obtained (Aleshin et al., 2000; Hurley et al., 1993). In

all these structures, the nucleotide binds at the same general location in a cleft formed between the two domains (Figure 3.9). The structure-based multiple sequence alignment of representative PanK-III enzymes and a diverse set of the structures in the ASKHA superfamily (Figure 3.10) shows that several conserved motifs that interact with the bound substrate, in particular ATP, are also conserved in PanK-III (Bork et al., 1992). These include the so-called PHOSPHATE 1 motif that encompasses the loop connecting strands β_1 and β_2 of the N-terminal domain and contains an invariant aspartate residue (Asp6 in PanK_{Tm}). This Asp residue has been shown to coordinate the Mg²⁺ ion that interacts with the β - and γ -phosphates of ATP (Hurley et al., 1993; van den Ent and Lowe, 2000). The second highly conserved aspartate residue (Asp105 of PanK_{Tm}) is located at the beginning of helix α_3 . This Asp residue is close to the phosphoryl acceptor group and has been proposed to act as a catalytic base (Hurley, 1996; Hurley et al., 1993). The PHOSPHATE 2 motif, also present in PanK-III, is located in the C-terminal domain between β_1' and β_2' and is somewhat structurally symmetrical to the PHOSPHATE 1 motif in the N-terminal domain. The third conserved motif, ADENOSINE, located in a loop after strand β_4' in the C-terminal domain, forms part of the pocket that binds the adenosine moiety of the nucleotide.

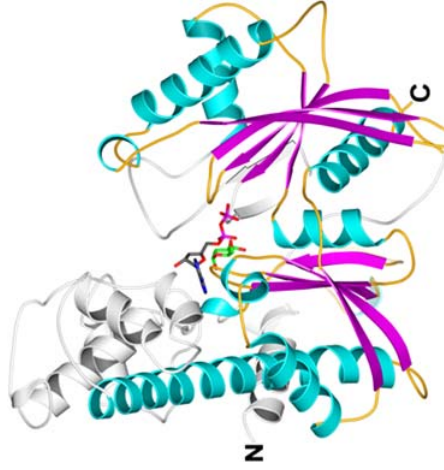
The conservation of these sequence and structural motifs in the PanK-III proteins indicates that these motifs are likely to play similar roles in substrate binding and catalysis as in other members of the superfamily. Since the complex



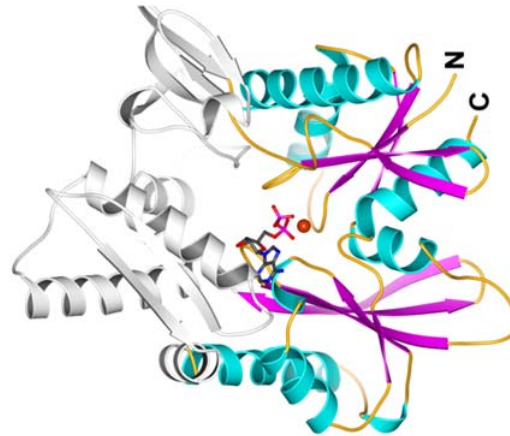
Pantothenate kinase Type III



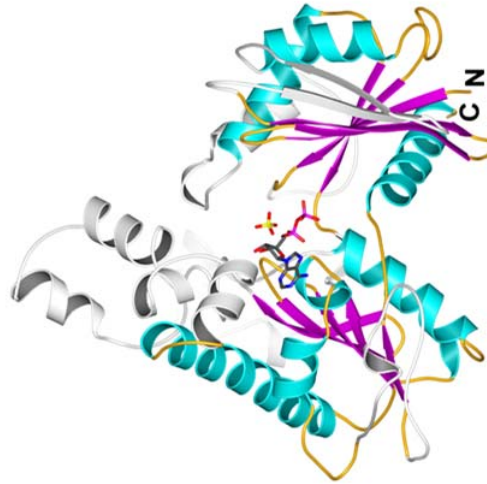
HG-CoA dehydratase Comp A (1hux)



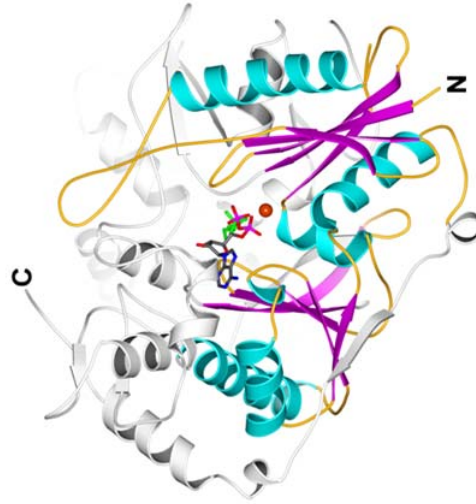
Hexokinase C-term domain (1dgk)



Hsp70 ATPase domain (1ba1)



Acetate kinase (1g99)



Glycerol kinase (1glc)

Figure 3.9 Fold comparison of PanK-III with representative members of the ASKHA superfamily

The corresponding secondary structure elements in the core of each structure are colored accordingly, with β -strands in magenta and α -helices in cyan. The regions that are considered insertions to the RNase H-like fold core are gray. Bound substrates in each structure are shown in a stick representation. The modeled substrates of PanK_{Tm} are also shown.

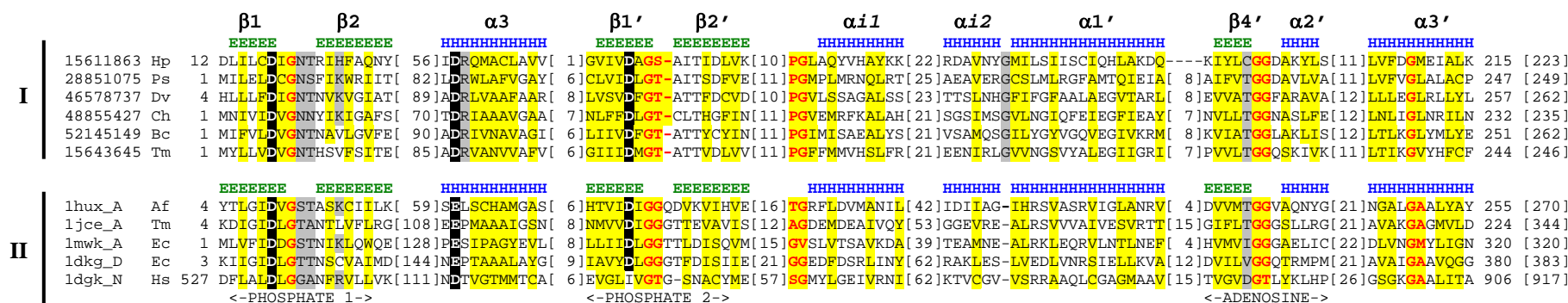


Figure 3.10 Multiple sequence alignment of representative sequences of PanK-III (group I) and actin/hsp70/sugar kinase superfamily with known structures (group II)

Sequences are labeled according to the gi number or PDB code and species name. The first and last residue numbers are indicated before and after each sequence, with the lengths of insertions specified in square brackets and the total sequence lengths of proteins following in parentheses. Residue conservation is denoted by the following scheme: uncharged, highlighted in yellow; charged/polar, in gray; small, in red; identical, bold and highlighted in black. The PHOSPHATE 1, PHOSPHATE 2, and ADENOSINE motifs are indicated at the bottom of the alignment. The secondary structure elements (E, b-strand; H, a-helix) for PanKTm (gi 15611833) and (PDB 1hux_A) are marked above each sequence block, respectively. Abbreviations of species names are as follows: Hp, *Helicobacter pylori*; Ps, *Pseudomonas syringae*, Dv, *Desulfovibrio vulgaris*; Ch, *Cytophaga hutchinsonii*; Bc, *Bacillus cereus*; En, *Emericella nidulans*; Mm, *Mus musculus*; Ce, *Caenorhabditis elegans*; Hs, *Homo sapiens*; Sa, *Staphylococcus aureus*; Af, *Acidaminococcus fermentans*; Tm, *Thermotoga maritima*; and Ec, *Escherichia coli*.

structure of PanK-III with substrate is at present not yet available, superposition of PanK_{Tm} with several ASKHA proteins enabled us to model both the ATP and pantothenate substrates in the PanK_{Tm} active site (Figure 3.11). This model provided a general placement of the substrates based on which further mutagenesis and kinetic analysis may be carried out to investigate the precise roles of the active site residues.

Proposed ATP and pantothenate binding sites

In the current model of the PanK_{Tm}-substrate complex (Figure 3.11), the ATP molecule interacts directly with the PHOSPHATE 1, PHOSPHATE 2, and ADENOSINE motifs. Asp6 of the PHOSPHATE 1 motif is in position to coordinate the divalent metal ion that interacts with the ATP phosphates, while Asp125 in the PHOSPHATE 2 motif may coordinate Mg²⁺ indirectly through a water molecule. The ADENOSINE motif in PanK_{Tm} adopts a conformation very similar to that in other members of the family and is predicted to play a similar role in binding the adenosine moiety of ATP.

The modeling of pantothenate binding in PanK_{Tm} revealed that the second monomer of the dimer likely contributes to the pantothenate binding site of the first monomer and vice versa (Figure 3.11). The loop following helix α_{i1} in the C-terminal domain insertion of the second monomer (hereafter termed "Pan cap," for "pantothenate binding site cap") is in close contact with the PHOSPHATE 1 loop

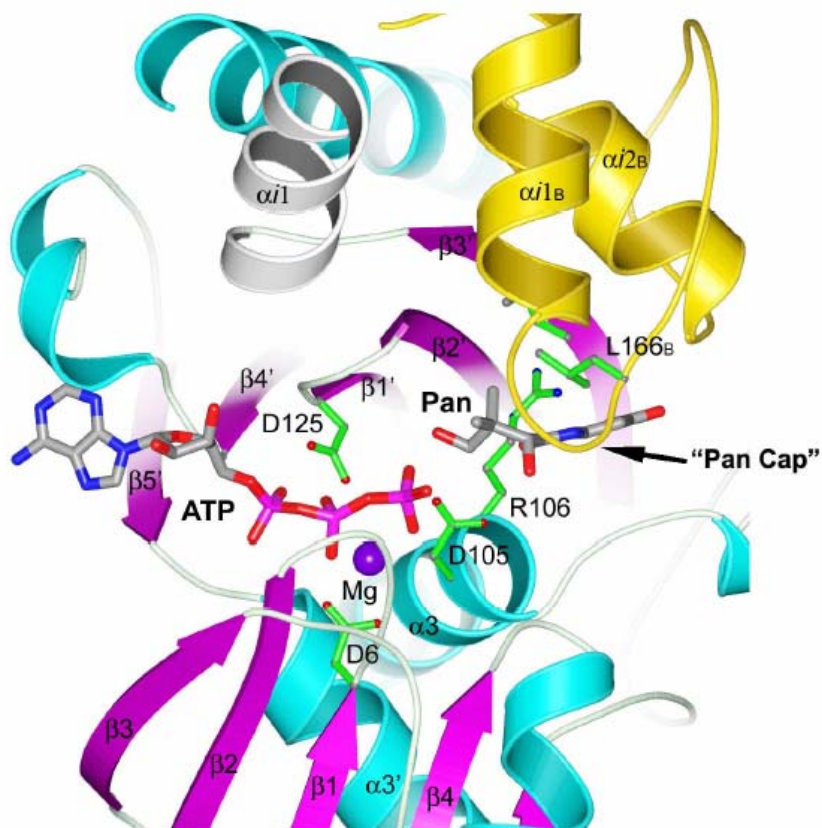


Figure 3.11 Model of MgATP and pantothenate binding in PanK_{Tm} active site

The color scheme is the same as that in Figure 3.9 for the first monomer of the dimer, while the region corresponding to the second monomer is in yellow. The substrates ATP and pantothenate are shown as thick bonds, while the side chains of several active site residues are shown in the ball-and-stick representation. The Mg²⁺ ion is shown as a purple ball.

of the first monomer. A hydrogen bond is formed between the side chains of Asn9 of the first monomer to the backbone amide group of residue Ala161' of the second monomer. Additionally, Leu163' of the second monomer, together with Ala129 and Ile145 of the first monomer, forms a small hydrophobic site that would accommodate the two methyl groups of the pantothenate substrate. As a result, this "Pan cap" loop forms a cap that would close in on the bound pantothenate and sequester the substrate from the surrounding solvent. The three hydrophobic residues (Ala129, Ile145, and Leu163) are highly conserved among all PanK-III proteins, suggesting their importance in either dimerization or the formation of the pantothenate binding site. The shared active site between two monomers of the PanK-III homodimer appears to be a novel feature unique among members of ASKHA superfamily.

The phosphoryl transfer reactions catalyzed by ASKHA enzymes are believed to proceed via nucleophilic attack by the phosphoryl acceptor group on the γ -phosphoryl moiety of ATP, followed by the direct transfer of the terminal phosphate to the acceptor molecule (Blattler and Knowles, 1979; Hurley, 1996). Residues corresponding to Asp105 of PanK_{Tm} have been proposed to act as a catalytic base activating the hydroxyl group of the phosphoryl acceptor for the nucleophilic attack. In the PanK_{Tm}-substrate complex model, Asp105 is in good position for playing such a role (Figure 3.11).

To investigate the proposed roles for the highly conserved Asp residues, mutagenesis studies on PanK_{Hp} have been carried out in our collaborator Dr. Strauss lab. PanK_{Hp} is closely related to PanK_{Tm} with which it shares about 32% sequence identity (Figure 3.10).

Based on sequence and structural analysis, it is almost certain that the conserved residues in these two proteins have the same function. PanK_{Hp} residues Asp17, Asp87, and Asp102 (corresponding to Asp6, Asp105, and Asp125, respectively, in PanK_{Tm}) were each mutated to either Asn or Glu. With the exception of the Asp17Glu mutant, all the mutant proteins expressed well and were purified by immobilized metal ion affinity chromatography. The enzymatic activities of the mutants were subsequently measured and compared to that of the native enzyme (Table 3.4). These results show that even a conservative substitution of any of the three proposed active site aspartate residues reduced enzyme activity drastically to less than 6% of that of the wild type. Two of these residues, Asp17 and Asp102, are proposed to be the metal ligands, while Asp87 is proposed to be the catalytic base. The mutagenesis data underscore the critical roles these residues play in the PanK-III-catalyzed reaction and are consistent with the mutant data for hexokinases and other members of the ASKHA superfamily (Arora et al., 1991; Wilbanks et al., 1994; Wilbanks and McKay, 1998).

Table 3.4 Effect of mutation of the active site aspartate residues on the activity of PanK_{Hp}^a

PanK_{Hp} protein	Relative rate of activity (%)
Native	100 ± 2.1
Asp17Asn	4.7 ± 1.4
Asp87Asn	2.8 ± 1.1
Asp87Glu	5.0 ± 1.7
Asp102Asn	2.7 ± 2.7
Asp102Glu	2.7 ± 1.3

^a The native enzyme and mutants were assayed under identical conditions in reaction mixtures containing 15 mM ATP plus 500 μM pantothenate in 100 mM HEPES, pH 7.6, in three separate experiments for each protein. The initial rates of reaction were determined and are reported relative to the rate of the native enzyme, which was set at 100. Reported errors are the standard deviation of the three experiments (Yang et al., 2006).

Fold comparison of PanK-III with PanK-I and PanK-II

Although PanK-III and PanK-I share the same functions, they adopt different protein folds, belonging to ASKHA and P-loop kinase superfamily, respectively. Comparison of the active sites of PanK_{Tm}-III and PanK_{Ec}-I (Figure 3.12) shows the dramatic difference between the two structures, which may provide a structural explanation for their distinct enzymatic characteristics. Both PanK-III and PanK-II share the same actin-like monomer fold, but they associate to form two very different dimeric architectures (Figure 3.13). These studies provide a structural framework for understanding the substrate specificity and their distinct catalytic mechanisms. A detailed discussion of the comparison between PanK-III and PanK-II will be described later in Chapter VI.

Comparison of the crystal structure of PanK_{Tm} and its active site configuration with those of PanK_{Ec} may provide a structural explanation for its lack of feedback inhibition by CoA and its inability to phosphorylate the *N*-alkylpantothenamide antimetabolites. As illustrated in a series of structures of PanK_{Ec} complexed with CoA, MgATP, as well as with both ADP and pantothenate (Ivey et al., 2004; Yun et al., 2000), CoA binds to PanK_{Ec} tightly in a site that partially overlaps with the ATP site, with 5'-phosphate of CoA overlapping with the γ -phosphate of ATP. Surprisingly, the 3'-phosphoadenosine moiety of CoA binds at a completely different site from that of the adenosine group of ATP (Yun et al., 2000). In contrast, the pantetheine tail of CoA largely

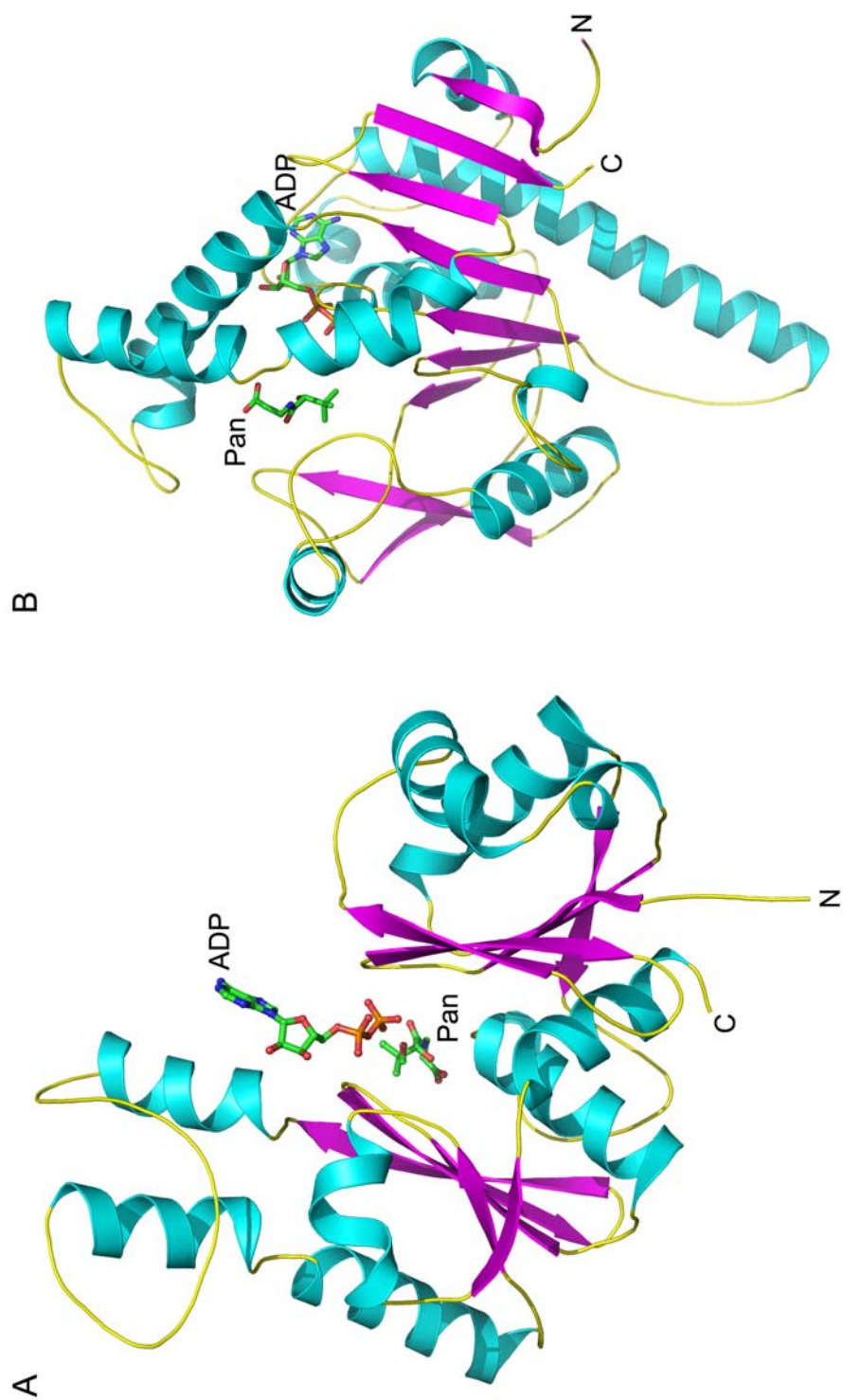


Figure 3.12 Crystal structural comparison of PanK_{Tm}-III and PanK_{Ec}-I

The pantothenate and ADP bound to PanK_{Tm}-III (A) and PanK_{Ec}-I (B) are shown as a stick representation with green carbons.

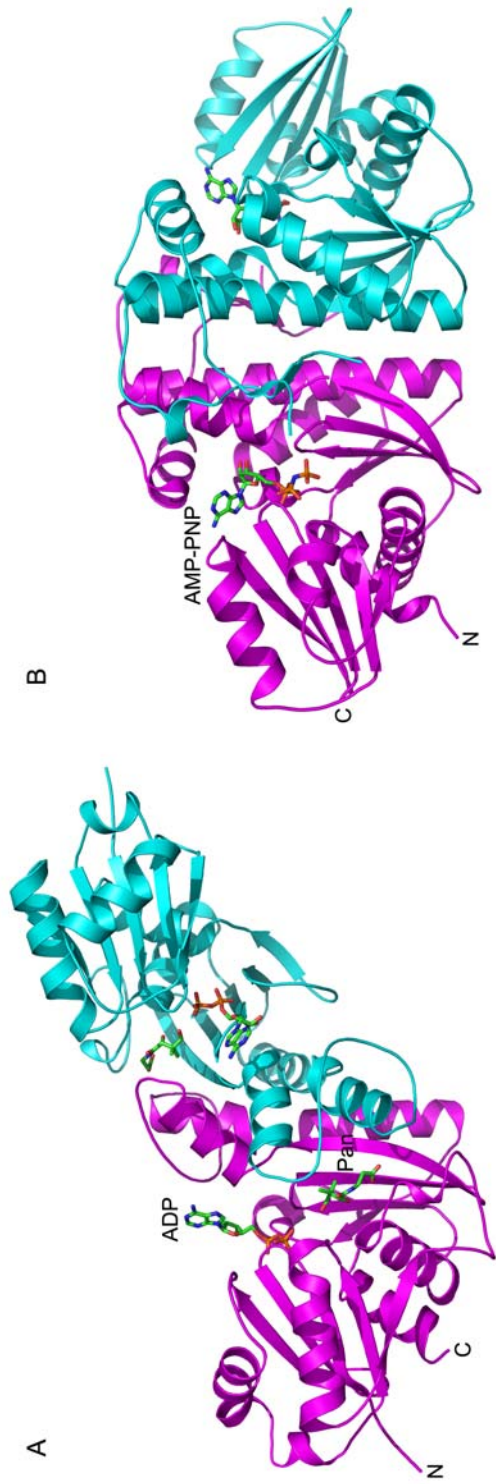


Figure 3.13 Crystal structural comparison of PanK_{Tm}-III and PanK_{Sa}-II

The first monomer (magenta) of PanK_{Tm}-III (A) and the first monomer (magenta) of PanK_{Sa}-II (B) are shown in identical orientations to highlight their structural similarities and their different packing interactions with their dimeric partners (cyan). The pantothenate and ADP bound to PanK_{Tm}-III and the AMP-PNP bound to PanK_{Sa}-II are shown as a stick representation with green carbons.

overlaps with the pantothenate binding site. The additional thiol group of CoA is accommodated in a hydrophobic pocket lined with mostly aromatic residues (Yun et al., 2000). It is hypothesized that this hydrophobic pocket would also be able to accommodate the longer hydrophobic tail of the *N*-substituted alkylpantothenamides (Ivey et al., 2004), which can thus serve as alternative substrates of the enzyme and are converted to CoA antimetabolites (Strauss and Begley, 2002). Inspection of a potential pantothenate binding site in PanK_{Tm} reveals no such hydrophobic pocket that could accommodate the longer tail of either CoA or *N*-alkylpantothenamides, which may explain, at least partially, why these molecules are not inhibitors or substrates of PanK-III and why PanK-III is not feedback inhibited by CoA or its thioesters.

At present, it is difficult to speculate why PanK-III has such a high K_m for ATP. It should be noted that PanK-III is not the only enzyme in the ASKHA superfamily that possesses such a high K_m towards its substrate. The K_m s of *Methanosarcina thermophila* acetate kinase for its substrate are also quite high: 2.8 mM for ATP and 22 mM for acetate (Aceti and Ferry, 1988). The consequences and implications of such steady-state kinetic properties on the metabolic fluxes in the organism remain to be illustrated. Clearly, further kinetic and structural studies are required to fully understand the underlying mechanisms of the PanK-III-catalyzed reaction.

Type III PanK occurs in a wide range of bacterial species

It has been noted previously that PanK-III appears to be more common in the bacterial world than the "classical" PanK-I (Overbeek et al., 2005; Ye et al., 2005). To examine this assertion, we conducted a comprehensive survey of the phylogenetic distribution of type I, II and III PanKs in over 300 complete or nearly complete genomes from the *Archaea*, *Eukarya*, and 13 major groups of *Bacteria* using the expert annotated SEED genomic integration platform (<http://theseed.uchicago.edu/FIG/index.cgi>). The same database has previously been used to establish the existence of the five-step universal CoA biosynthetic pathway in the majority of these organisms (Overbeek et al., 2005; Ye et al., 2005). The results of this survey clearly show that PanK-III exists in 12 of the 13 major bacterial groups, the exception being the *Chlamydiae*, for which no candidate PanK has yet been identified (Table 3.5; for more details, see Appendix A). PanK-I, on the other hand, is present in only four groups of *Bacteria*: the *Actinobacteria*, *Chloroflexi* (green non-sulfur bacteria), *Firmicutes* (gram-positive bacteria), and *Proteobacteria* (purple bacteria and relatives). This surprisingly widespread distribution of PanK-III further underscores the importance of understanding the mechanism of this important enzyme and the regulation of CoA biosynthesis in organisms harboring PanK-III. Interestingly, a number of bacteria have more than one type of PanK: in mycobacteria and several bacilli both type I and type III PanKs are present, while *Bacillus anthracis* and *B. cereus* contain

both PanK-II and PanK-III. The physiological significance of this functional redundancy is currently unclear. Moreover, genes coding for the PanK activity have not been identified in the *Archaea* kingdom, suggesting the existence of another, as yet uncharacterized, type of PanK (Overbeek et al., 2005).

In summary, we have demonstrated that type III PanK encoded by *coaX* has a more widespread phylogenetic distribution than the long-known PanK-I and is nearly universally present in most of the major bacterial groups. The crystal structure of PanK_{Tm} revealed that type III PanK belongs to the ASKHA superfamily and adopts an entirely different fold from that of type I PanK. Mutagenesis and comparative structure analysis of PanK-III uncovered features of the enzyme and provided a structural explanation for the lack of product feedback inhibition of PanK-III. Since the currently established inhibitors of type I and type II PanKs are ineffective against PanK-III, the biochemical and structural elucidation of PanK-III not only is important for understanding the fundamental metabolic pathways in many PanK-III-harboring organisms but also provides a structural basis for the computer-aided design of specific inhibitors targeting PanK-III which may lead to novel antibacterial therapeutics.

Table 3.5 Phylogenetic distributions of different types of PanK in bacteria^a

Bacterial Groups		Type I	Type II	Type III
Actinobacteria		√		√
Aquificae				√
Bacteroidetes/Chlorobi				√
Chloroflexi		√		√
Chlamydiae ^b				
Cyanobacteria				√
Deinococcus/Thermus				√
Firmicutes		√	√	√
Fusobacteria				√
Planctomycetes				√
Proteobacteria	α	√		√
	β			√
	δ			√
	ϵ			√
	γ	√		√
Spirochaetes				√
Thermotogae				√

^a The presence of each type of PanK is indicated. For a more detailed distribution of different types of PanK in individual species, see the supplemental material.

^b No candidate for any type of PanK can as yet be identified in the *Chlamydiae* (Raman, 2004).

CHAPTER IV

Thermodynamics Characterization of Substrate–Enzyme

Interactions in PanK-III

Introduction

Characterization of the thermodynamics of substrate-enzyme interactions is important in improving our understanding of enzymatic mechanism. Isothermal titration calorimetry (ITC) is one of the methods for undertaking such studies. The power of ITC lies in its unique ability to measure binding reactions by the detection of the heat change during the binding interaction, and it measures the association/dissociation constant (K_a/K_d), stoichiometry (n), free energy change (ΔG), enthalpy change (ΔH) and entropy change (ΔS) of binding. Since heat changes occur during many physicochemical processes, ITC has a broad application, including in enzyme kinetics.

In order to learn the nature of the enzyme-substrate interaction and obtain an unbiased binding affinity of each substrate (Pan or ATP) toward the enzyme, we carried out ITC experiments with PanK-IIIs from *Thermotoga maritima* and *Helicobacter pylori*.

Materials and Methods

Isothermal Titration Calorimetry (ITC)

PanK-III from *T. maritime* and *H. pylori* were cloned, expressed, and purified as described before (Brand and Strauss, 2005; Yang et al., 2006). To prepare for the ITC assay, proteins were exhaustively exchanged into a buffer containing 100 mM NaCl, 20 mM Tris, pH 8.0. Protein concentrations were determined by measuring the absorbance at 280 nm and were calculated according to Beer's Law ($A = \epsilon cl$, while A is absorbance, ϵ is the molar extinction coefficient, c is the protein concentration, and l is the path-length). The extinction coefficients were obtained from ProtParam tool available at the ExPASy Proteomics Server (www.ExPASy.org). The extinction coefficient at 280nm used are $27055 \text{ M}^{-1} \text{ cm}^{-1}$, $13910 \text{ M}^{-1} \text{ cm}^{-1}$ and $18950 \text{ M}^{-1} \text{ cm}^{-1}$ for PanK_{Tm}-III, PanK_{Hp}-III and PanK_{Pa}-III proteins, respectively. The protein concentration was further confirmed using Bradford assay (Bradford, 1976) with bovine serum albumin (Sigma) as a standard. The isothermal calorimetric titration was carried out at 20°C in a VP-ITC titration microcalorimeter (MicroCal, Northampton, MA). Ligand (pantothenate or ATP analog AMPPNP) was titrated into a sample cell (1.8 ml) containing PanK-III protein in 31 serial injections of 10 µl each. The heat change from protein-ligand interaction was monitored by the VP-ITC instrument until the target protein was saturated with the ligands. Data were processed and fitted using Microcal-ORIGIN software (OriginLab, Northampton, MA).

Results and Discussions

ITC results show that pantothenate binds to both PanK-IIIs with high affinity (Figure 4.1 and 4.2), and the $K_{d, \text{Pan}}$ range from 2.7 to 6.4 μM (Table 4.1). Interestingly, the titration profile revealed that two pantothenate binding sites in PanK_{Tm}-III dimer are not equal under the assay condition, and the data is best fitted with a sequential two-site model with the first site having a K_d of 2.7 μM , and the second 6.4 μM (Figure 4.1). This observation suggests a slightly negative cooperativity of the two pantothenate binding site in PanK_{Tm}-III dimer. The negative enthalpy change ΔH for the binding of pantothenate to PanK_{Tm}-III is -8.4 kcal/M (Table 4.1), indicating that the reaction is exothermic (releasing heat), and there is an overall increase in bonding. The entropy change (ΔS) upon pantothenate binding to PanK_{Tm}-III is $\sim -3 \text{ cal/mole/deg}$, indicating a decrease in disorder, which is often associated with an increase in bonding and presence of hydrophobic interactions between ligand and protein. All these are consistent with the extensive enzyme-pantothenate interactions including both specific hydrogen bonds and hydrophobic interactions as observed in the PanK-III•Pan complex structures (see Hong *et al.*, 2006 and chapter VI). Surprisingly, the binding of ATP or ATP analog AMPPNP is endothermic ($\Delta H = 2.9 \text{ kcal/M}$), and is accompanied by a more favorable entropy component ($\Delta S = 21.3 \text{ cal/mole/deg}$) (Table 4.1, Figure 4.3). Overall the favorable entropy contribution is sufficient to

off-set the unfavorable enthalpic changes, and the free energy change ΔG for ATP binding is negative (-3.2 kcal/mole).

The substrate binding affinities for PanK_{Hp}-III are very similar to those of PanK_{Tm}-III with $K_{d, Pan}$ of 5.6 μ M and the $K_{d, ATP}$ of 2.3 mM (Table 4.1). Notably, no cooperativity is observed between the two substrate binding sites of PanK_{Hp}-III dimer and the two sites appear to be independent of each other (Figure 4.2).

In order to determine whether pantothenate binding would affect the binding of ATP and vice versa, we also determined the dissociation constants for Pan or AMPPNP in the presence of the second substrate (Figure 4.1B, 4.2B and 4.4B). The results show that the substrate binding constants in the absence and the presence of the second substrate are essentially identical, suggesting that the Pan and ATP binding may be independent of each other. Although these data may imply a random mechanism for the enzyme, due to the extreme difference in the affinities of the two substrates, *i.e.*, pantothenate binds to the enzyme about ~400 times tighter than ATP, an ordered sequential mechanism can not be ruled out. In fact the structural data of PanK-III has suggested an ordered mechanism with Pan binding first to the enzyme (Hong et al., 2006). This is in contrast to the ordered mechanism of PanK_{Ec}-I, in which ATP is the leading substrate (Yun et al., 2000).

Table 4.1 Thermodynamic parameters for interactions between substrate and PanK-III

		Pan				ATP			
		$K_d(\mu\text{M})$	$\Delta H(\text{kcal/M})$	$-T\Delta S(\text{kcal/M})$	$\Delta G(\text{kcal/M})$	$K_d(\text{mM})$	$\Delta H(\text{kcal/M})$	$-T\Delta S(\text{kcal/M})$	$\Delta G(\text{kcal/M})$
TmPanK-III	site 1	2.7±0.16	-8.4±0.07	0.94	-7.5	3.0±0.50	2.9±0.36	-6.2	-3.3
	Site 2	6.4±0.24	-15.2±0.11	8.20	-7.0				
HpPanK-III		5.6±0.34	-6.4±0.06	-0.59	-7.0	2.3±0.15	2.0±0.10	-5.6	-3.4

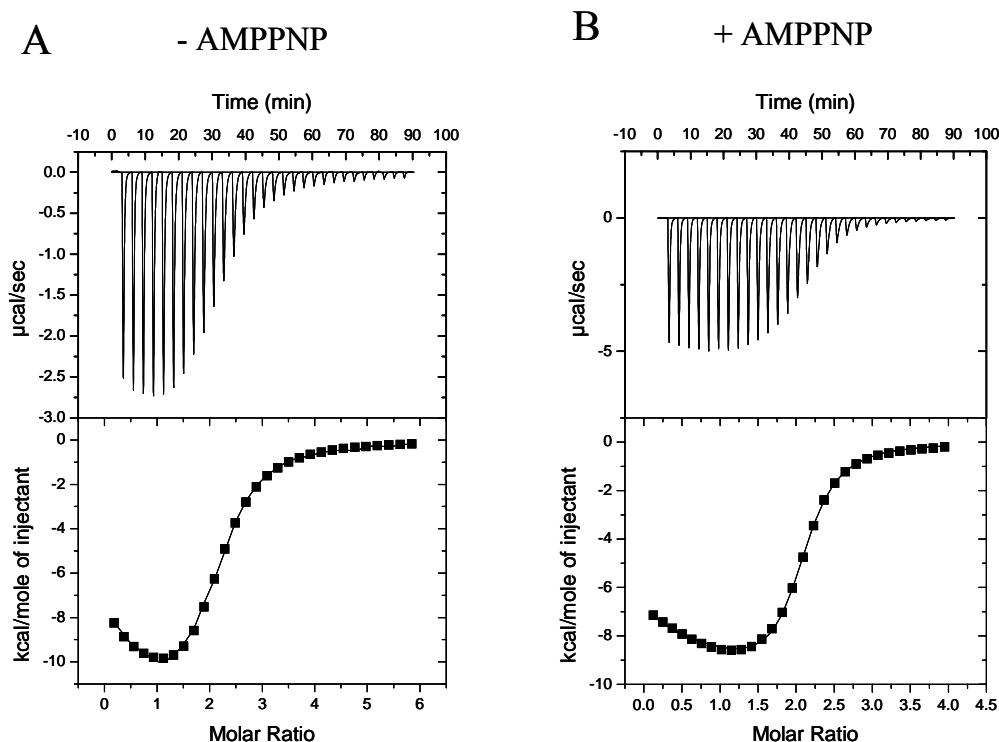


Figure 4.1 Calorimetric titration of PanK_{Tm} with substrates pantothenate in the absence or presence of AMPPNP

The upper panels of the figures show the heat changes elicited by successive injections of pantothenate into the solution containing PanK_{Tm}. The lower panels show the binding isotherms as a function of the molar ratio of substrate to the enzyme dimer. The theoretical curves were fitted to the integrated data. A sequential two site binding model was used. The concentration of the injectant pantothenate and PanK_{Tm} was 1 mM and 0.05 mM, respectively (**A**) and 2.1 mM and 0.12 mM, respectively (**B**, in the presence of 30 mM AMPPNP). K_d values in the absence of AMPPNP (**A**) are: $K_{d1} = 2.7 \mu\text{M}$, $K_{d2} = 6.4 \mu\text{M}$ and in the presence of 30 mM AMPPNP (**B**): $K_{d1} = 2.6 \mu\text{M}$, $K_{d2} = 6.3 \mu\text{M}$.

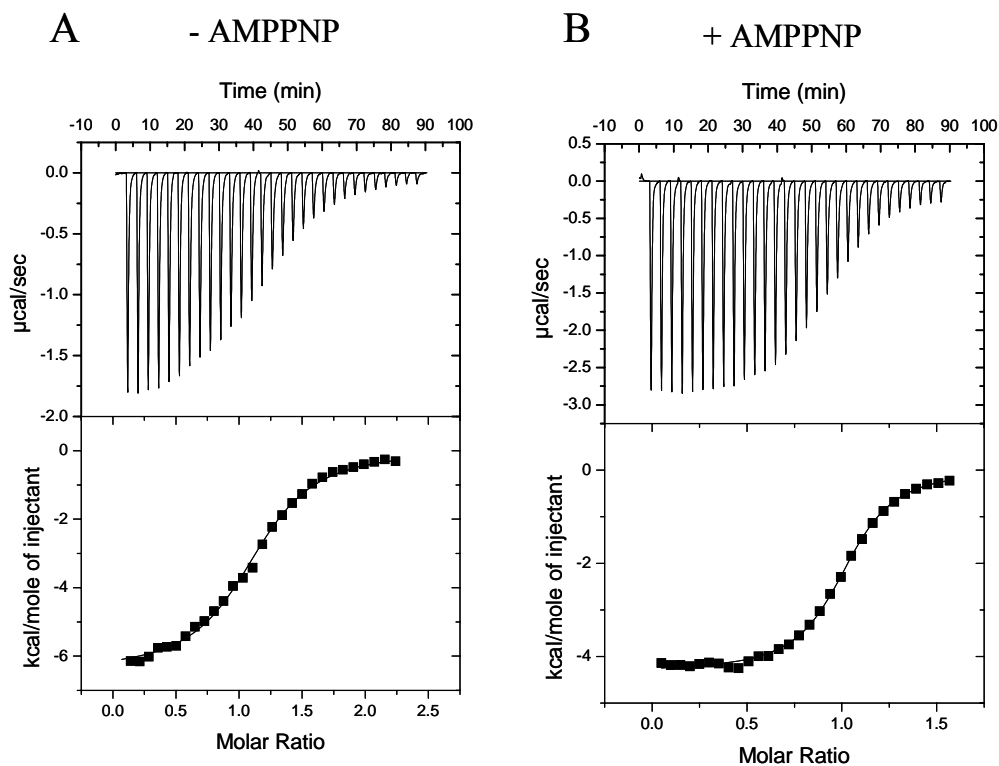


Figure 4.2 Calorimetric titration of PanK_{Hp} with substrates pantothenate in the absence or presence of AMPPNP

The concentration of the injectant pantothenate and PanK_{Hp} was 1 mM and 0.1 mM, respectively (**A**) and 2.1 mM and 0.3 mM, respectively (**B**, in the presence of 30 mM AMPPNP). One site binding model was used to fit the data. K_d values are 5.6 μ M (**A**) and 5.0 μ M (**B**).

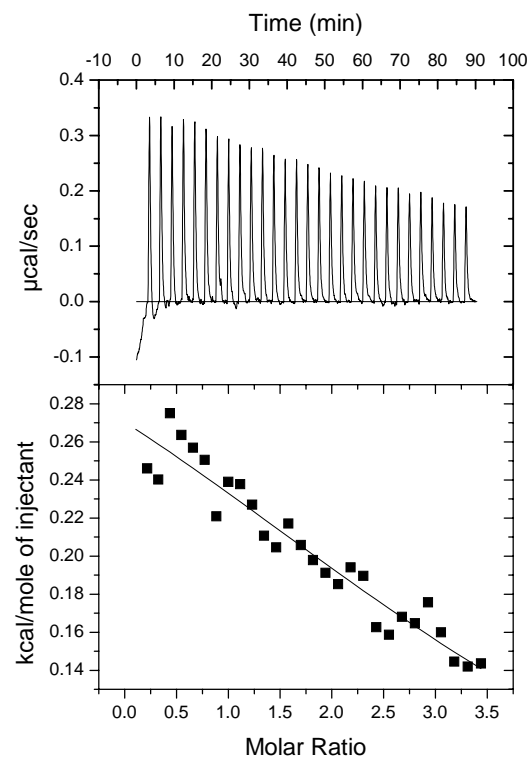


Figure 4.3 Calorimetric titration of PanK_{Tm} with AMPPNP

The concentration of AMPPNP was 4.2 mM, and that of PanK_{Tm} 0.28 mM. One site binding model was used to fit the data. K_d value is determined to be 3 mM.

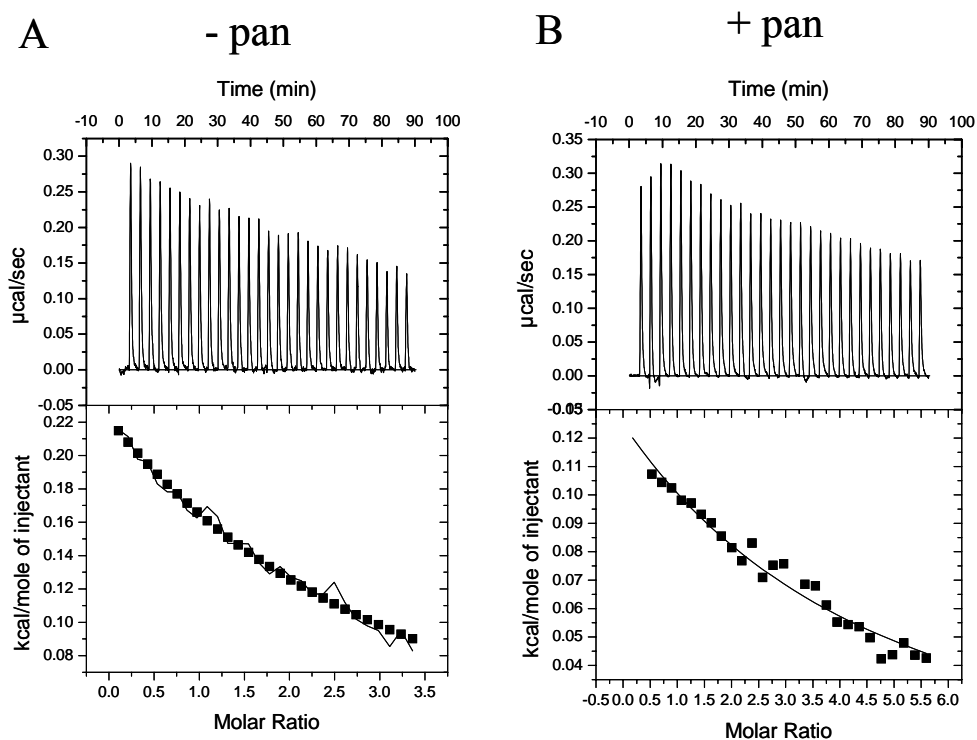


Figure 4.4 Calorimetric titration of PanK_{Hp} with substrates AMPPNP in the absence or presence of pantothenate

The concentration of the injectant AMPPNP and PanK_{Hp} was 4.2 mM and 0.28 mM, respectively (A) and 6.9 mM and 0.28 mM, respectively (B, in the presence of 1 mM pantothenate). One site binding model was used to fit the data. K_d values are 2.3 mM (A) and 2.1 mM (B).

CHAPTER V

Steady State Properties and Nucleotide Specificity of PanK-III

Introduction

One of the unusual enzymatic properties of PanK-III is the high K_m for ATP (~3-10 mM) (Brand and Strauss, 2005). Here we accurately determined the steady state kinetic parameters for PanK-III. The kinetics assays and inhibition studies indicate that the reaction catalyzed by PanK-III is likely to proceed via an ordered Bi-Bi mechanism. In addition, the high K_m for ATP had also raised possibility that PanK-III may utilize alternative phosphoryl donors. Nucleotide specificity of PanK-III was also tested here.

Materials and Methods

Enzyme activity and kinetics assays

The enzyme activity assay and steady state kinetics were carried out using a coupled enzyme assay (Brand and Strauss, 2005; Singh et al., 2004) (Appendix B). Briefly, a continuous spectrophotometric assay that couples the production of ADP to the reactions catalyzed by pyruvate kinase and lactate dehydrogenase was used and the consumption of NADH was monitored by changes in absorption at 340nm. Each 600 μ l reaction mixture contained 100 mM HEPES, pH 7.6, 20 mM

KCl, 10 mM MgCl₂, 2 mM phosphoenolpyruvate (PEP), 0.3 mM NADH, 5 units of lactate dehydrogenase, 2.5 units of pyruvate kinase and pantothenate and ATP as needed. Pantothenate concentration was varied from 0 to 80 μ M while ATP varied from 0 to 15 mM. 2 different enzyme concentrations were used in most experiments: 0.135 μ M and 0.27 μ M. 5 μ l of PanK-III protein was placed in cuvette and the reactions were initiated by the addition of reaction mixture. Apparent values of K_m and k_{cat} were calculated fitting initial rates to a standard Michaelis-Menten model using software SigmaPlot.

Results and Discussions

Steady state kinetic parameters of PanK_{Hp}-III

The K_d values for pantothenate obtained by ITC assays ($\sim 3 - 6$ μ M) are quite different from the previously determined K_m of pantothenate for PanK_{Hp}-III which is at ~ 100 μ M range (Brand and Strauss, 2005). To investigate the cause of this difference, we revisited the steady state kinetic experiment. To avoid potential complication that may arise from the apparent negative cooperativity of PanK_{Tm}-III, the *H. pylori* enzyme was used in these assays. Care was taken to minimize the time elapsed between the mixing of all reaction ingredients and the starting of the measurement, in order to ensure that less than 10% of substrate (pantothenate) was consumed when the initial rate was obtained. With the new

procedure, the $K_{m, \text{Pan}}$ was determined to be 6.3 μM for PanK_{Hp}-III (Figure 5.1), which is very closed to the K_d value of Pan determined with ITC experiment (5.6 μM). The k_{cat} is determined to be 3.5 min^{-1} , similar to the previous reported value (Brand and Strauss, 2005). A similar K_m value for ATP (~10 mM vs. 9.6 mM) was also obtained with the same procedure.

The consistency between the K_d values obtained from ITC binding assay and the steady state parameter K_m values for both substrates suggests that substrate binding instead of catalysis is likely to be the rate-limiting step of the enzyme. As PanK-III is no longer regulated by the end product of the pathway CoA, its low affinity for ATP may be of important physiological consequence for CoA biosynthesis and homeostasis in organisms harboring PanK-III.

Kinetics analysis and inhibition studies

Next, kinetics analysis was carried out by varying pantothenate concentration ranging from 5 to 80 μM , and ATP concentration from 4 to 10 mM (Figure 5.2A). The double-reciprocal plots of initial velocity *versus* pantothenate concentration at different fixed ATP concentrations were linear and not parallel (Figure 5.2B). Intersecting lines suggest that the reaction occurred by a sequential mechanism (either random or ordered bi-bi mechanism).

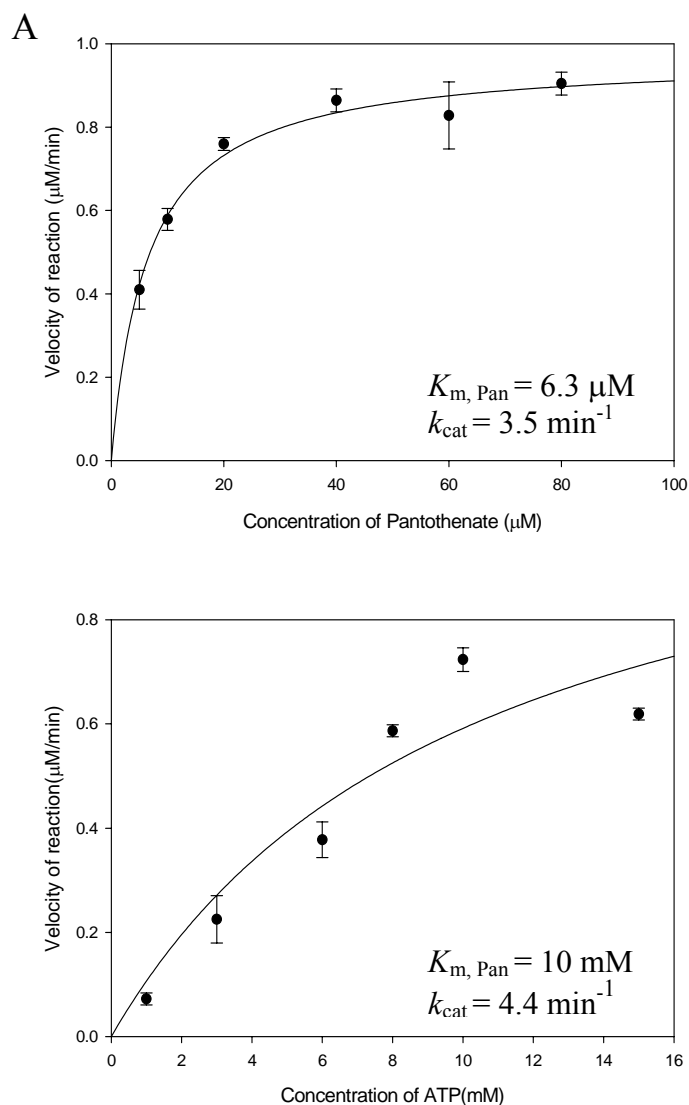


Figure 5.1 Steady state kinetic parameters of PanK_{Hp}

Data was fit to the Michaelis-Menten equation. **(A)** The pantothenate concentration varied from 0 to 80 μM , and the ATP concentration used is 12 mM. **(B)** The ATP concentration varied from 0 to 15 mM, and the pantothenate concentration is 200 μM .

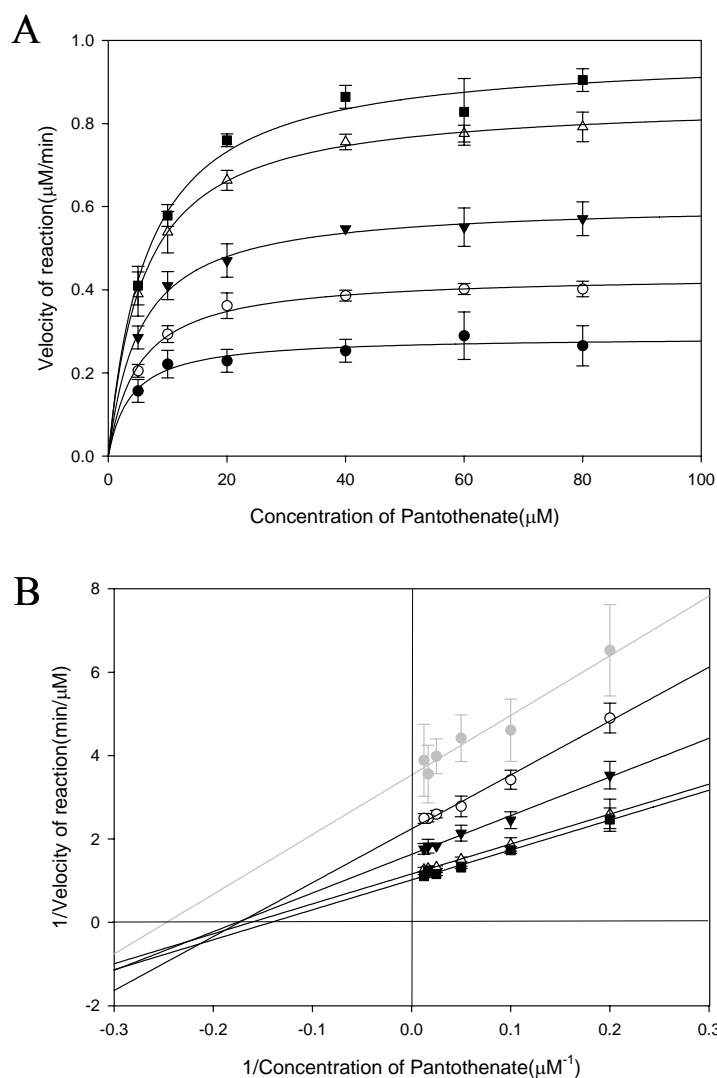


Figure 5.2 Initial rate plot of PanK_{Hp} versus pantothenate concentration

A. Data was fit to the Michaelis-Menten equation. The ATP concentration used is 4 mM (●), 6mM (○), 8 mM (▼), 10mM ATP(△), 12 mM ATP(■). **B.** Double-reciprocal plots of initial velocity versus pantothenate concentration at different fixed concentrations of ATP. The ATP concentration used is 4 mM (●, line was colored grey due to the large error), 6mM (○), 8 mM (▼), 10mM ATP(△), 12 mM ATP(■).

To fully resolve the random or ordered bi-bi mechanism, a series of product and substrate-analog inhibition experiments are required. However, the product inhibition assay is not feasible for PanK-III with the continuously assay method, because one of the product ADP is used in the subsequent coupled reaction in the assay (see Appendix B), and the second product phosphopantothenate provided to us by our collaborator is contaminated with large amount of substrate pantothenate. Additionally, substrate pantothenate analog that would inhibit PanK-III activity has not been identified so far. Therefore, the only inhibition assay that I am able to carry out at this point is with ATP analog AMPPNP (Figure 5.3).

In both cases, V_{max} clearly decreased when AMPPNP was added, which indicates that AMPPNP likely does not inhibit either ATP or Pan competitively. The effects of AMPPNP on the K_m values of ATP and pantothenate are unusual and may not have been determined accurately due to the difficulty in working with the very low $K_{m,Pan}$ and the very high $K_{m,ATP}$. In any event, combined with the structural considerations (as discussed in chapter VI), the reaction catalyzed by PanK_{Hp} is probably not a random but an ordered Bi-Bi mechanism with pantothenate binds first to the enzyme and followed by ATP.

Nucleotide specificity for PanK-III

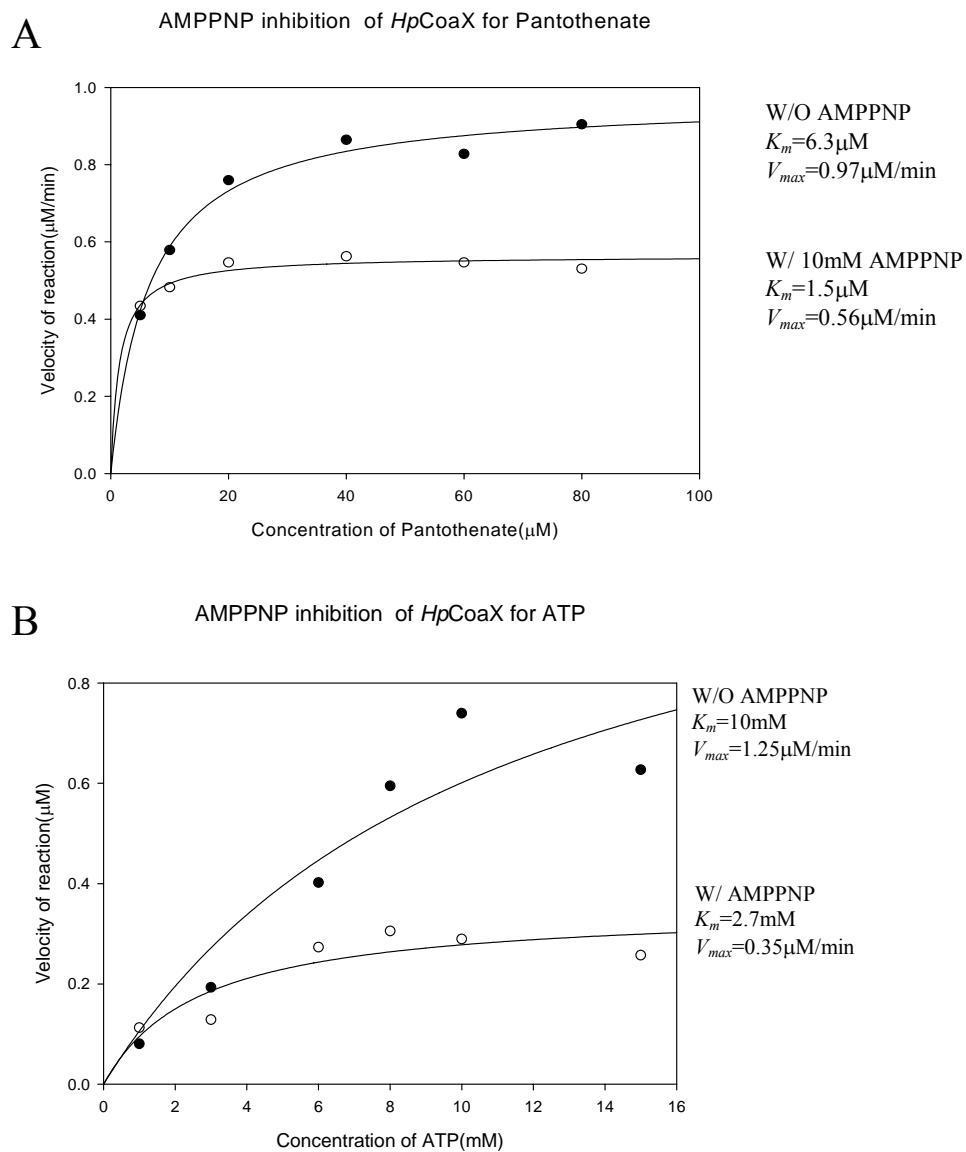


Figure 5.3 Kinetics of inhibition of PanK_{Hp} by AMPPNP

Effect of 10 mM AMPPNP on the steady state kinetic parameters with regard to pantothenate (**A**) and ATP (**B**).

One of the unusual enzymatic properties of PanK-III is the high K_m for ATP (~3-10 mM) (Brand and Strauss, 2005), which had raised possibility that PanK-III may utilize alternative phosphoryl donors. Experiments done with PanK-III from *P. aeruginosa* (PanK_{Pa}-III) showed that PanK_{Pa}-III can utilize various different NTP and dNTP as phosphoryl donors, and is most active with purinenucleotides (ATP, GTP, dATP and dGTP) (Hong et al., 2006). A nucleotide specificity assay was also carried out with PanK_{Hp}-III using 1 mM ATP, CTP, GTP and UTP as the phosphate group donor (Figure 5.4). The result showed a similar trend as the *P. aeruginosa* enzyme in that PanK_{Hp}-III is active with all nucleotides tested, and its activity is highest when using ATP as the substrate. Pyridine nucleotides UTP and CTP appear to be less efficient substrate than purine nucleotides. The broad substrate specificity of PanK-III is consistent with the structural observation and will be discussed later in Chapter VI.

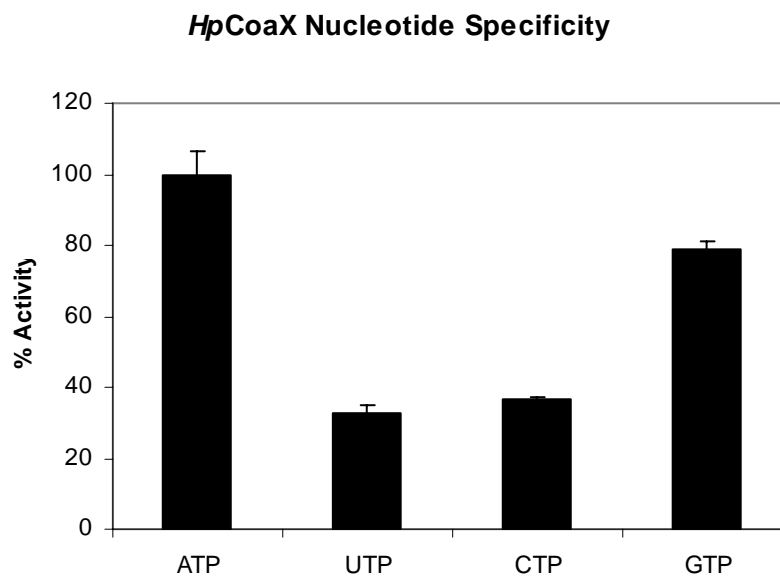


Figure 5.4 Nucleotide specificity of PanK-III

The activity of PanK_{Hp} was assayed using the indicated nucleotides (1 mM) as the phosphate group donor with the coupled kinase assay.

CHAPTER VI

Structures of PanK-III Substrate / Product Complexes

Introduction

Coenzyme A (CoA) is a ubiquitous and essential cofactor in all living organisms, where it functions as the major acyl group carrier in many crucial cellular processes, notably the tricarboxylic acid (TCA) cycle and fatty acid metabolism. In the five-step universal CoA biosynthesis pathway, pantothenate kinase (PanK) catalyzes the first rate-limiting step, transferring a phosphoryl group from ATP to pantothenate (Pan) (Begley et al., 2001; Jackowski, 1996; Leonardi et al., 2005b). Because of the essentiality of this pathway for the survival of the organism, the five CoA biosynthesis enzymes have been highlighted as new antimicrobial targets of important human pathogens (Becker et al., 2006). Due to its key role as the “gate-keeper” of the pathway and the lack of homology between bacterial PanK and their human counterpart, development of inhibitors targeting PanK is being actively pursued (Hong et al., 2006; Yang et al., 2006).

Three types of PanK have been characterized so far. The type I prokaryotic PanK (PanK-I) exemplified by *E. coli* CoaA protein and the predominantly eukaryotic type II PanK (PanK-II) have long been under intensive

investigations and their three dimensional structures have been determined (Calder et al., 1999; Raman, 2004; Rock et al., 2000; Rock et al., 2002). Both types of PanK, though evolutionarily unrelated, are feedback inhibited by the end-product of the pathway CoA and its thioesters, and play a key regulatory role in CoA biosynthesis and homeostasis in the cell (Rock et al., 2000; Rock et al., 2002; Vallari et al., 1987; Yun et al., 2000). Both PanK-I and PanK-II are also able to use pantothenate derived pantothenamides as alternative substrates, an observation that set off the pursuit of pantothenamides as potential antibiotics or antivitamin drugs (Hong, 2006; Ivey et al., 2004; Strauss and Begley, 2002; Virga et al., 2006; Yun et al., 2000). Notably, eukaryotic PanK-like enzyme is also found in a few gram-positive *bacilli* and *Staphylococci* species, including *Bacillus anthracis* and *Staphylococcus aureus* (Hong et al., 2006; Leonardi et al., 2005a; Nicely et al., 2007). Recent biochemical and structural characterizations of both *S. aureus* and human PanK-II enzymes revealed that bacterial PanK-II shares significant sequence and structural similarity with its eukaryotic counterpart, though they differs in certain catalytic properties (Hong et al., 2006; Leonardi et al., 2005a; Rock et al., 2000). In particular, PanK-II from *S. aureus* is not inhibited either by CoA or its thioesters (Leonardi et al., 2005a), a property that is likely related to the additional functions of CoA as the major low-molecular weight thiol in this organism (delCardayre et al., 1998; Luba et al., 1999; Nicely et al., 2007).

Recently a third type of PanK (PanK-III, encoded by gene *coaX*, to be distinguished from *coaA*), unrelated to bacterial PanK-I, was first identified in *Bacillus subtilis* (Brand and Strauss, 2005). Subsequently the biochemical properties and structures of PanK-III from several bacterial species were characterized (Brand and Strauss, 2005; Hong et al., 2006; Yang et al., 2006). Different from PanK-I and eukaryotic PanK-II, PanK-III is not feedback inhibited by CoA, neither can it use pantothenamide as substrate (Brand and Strauss, 2005). Additionally, PanK-III has an unusually high K_m for ATP (in the mM range) and requires a monovalent cation to for catalysis (Yang et al., 2006). A comparative genome-wide subsystem analysis of CoA biosynthesis pathway in ~400 bacteria with completely sequenced genomes using THE SEED tool (<http://theseed.uchicago.edu/FIG/index.cgi>) revealed that PanK-III has a wider phylogenetic distribution than the long known PanK-I, and is present in most major bacteria divisions, including many pathogenic bacteria (Yang et al., 2006).

The newly reported structures of PanK-II and PanK-III confirmed a previous prediction that both enzymes adopt the actin-like fold and belong to the ASKHA (abbreviation from “acetate and sugar kinase/Hsp70/actin”) protein superfamily (Cheek et al., 2005; Cheek et al., 2002). Notwithstanding the shared common fold and conserved key catalytic residues, PanK-II and PanK-III differ in dimer formation and in detailed substrate binding site configurations (Hong et al., 2006), which underlie their different substrate specificity and kinetic parameters.

These differences are particularly relevant to the inhibitor development effort targeting PanKs. Currently, a class of *N*-substituted alkylpantothenamides has shown greatest promise as growth inhibitors of both *E. coli* and *S. aureus* (Choudhry et al., 2003; Leonardi et al., 2005a; Strauss and Begley, 2002). The mechanism of the action of these pantothenate analogs lies in the promiscuity of bacterial PanK-I and PanK-II as well as the rest of CoA synthesis pathway to use these compounds as alternative substrate to generate nonfunctional CoA analogs that inhibit CoA- and acetyl-CoA-utilizing enzymes (Strauss and Begley, 2002). Given the wide phylogenetic distribution of PanK-III, and the inability of these pantothenamides to either inhibit PanK-III activity or to serve as alternative substrate, new strategies for developing inhibitors specifically targeting PanK-III will be needed.

Here we report the crystal structures of PanK-III from *Thermotoga maritima* (PanK_{Tm}-III) complexed with substrate pantothenate and product phosphopantothenate, respectively, as well as a ternary complex of PanK_{Tm}-III with pantothenate and ADP. These structures revealed the detailed interactions between both substrates and the enzyme, shed new light into the catalysis and the unique kinetic properties of PanK-III, and should facilitate the structural based approach for developing specific inhibitors targeting PanK-III.

Materials and Methods

Crystallization, data collection, and refinement

Cocrystals of PanK_{Tm} complexes were grown using hanging drop vapor-diffusion method in conditions similar to that of the native proteins. Prior to crystallization, 20 mg/ml PanK_{Tm} protein, in 20 mM HEPES, pH 8.0, 200 mM NaCl and 1 mM DTT, was incubated with either 5 mM pantothenic acid (Pan), 5mM phosphor-pantothenate (P-Pan), or with both 5mM pantothenic acid and 30mM ADP for the growth of respective binary and ternary complexes (Figure 6.1, 6.2 and 6.3). 1.5 μ l complex solution was then mixed with an equal volume of reservoir solution containing 16% PEG-3350, and were equilibrated against the reservoir over a period of several days at 4°C. Diffraction data for ternary complex (PanK_{Tm}-III•Pan•ADP) was collected in-house on an RAXIS-IV⁺⁺ image plate detector equipped with a Rigaku FR-E SuperBright X-ray generator and VarimaxTM HF mirrors. Data from two binary complexes, PanK_{Tm}-III•Pan and PanK_{Tm}-III•P-Pan, were collected at beamline 19-BM at the Advance Photon Source, Argonne National Laboratory (Argonne, IL). The diffraction data were indexed, integrated, and scaled using HKL2000 program package (Otwinowski and Minor, 1997). All complexes are isomorphous to the uncomplexed crystal and belong to the space group P2₁, with unit cell dimensions $a = 74.92 \text{ \AA}$, $b = 136.99 \text{ \AA}$, $c = 74.97 \text{ \AA}$, and $\beta = 109.62^\circ$.

Refinement and model building of all three PanK_{Tm}-III complexes were

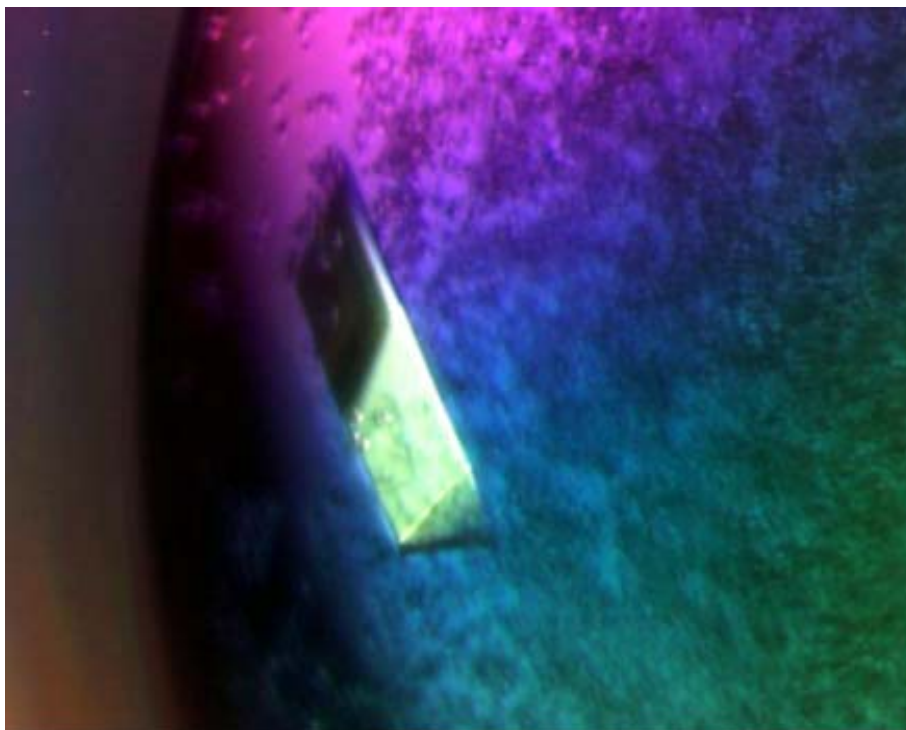


Figure 6.1 Binary complex crystal structure of PanK_{Tm} • pantothenate

The crystals were grown using sitting drop vapor diffusion method at 4°C. The optimized condition contains 16% PEG 3350. The protein concentration was 18 mg/ml.

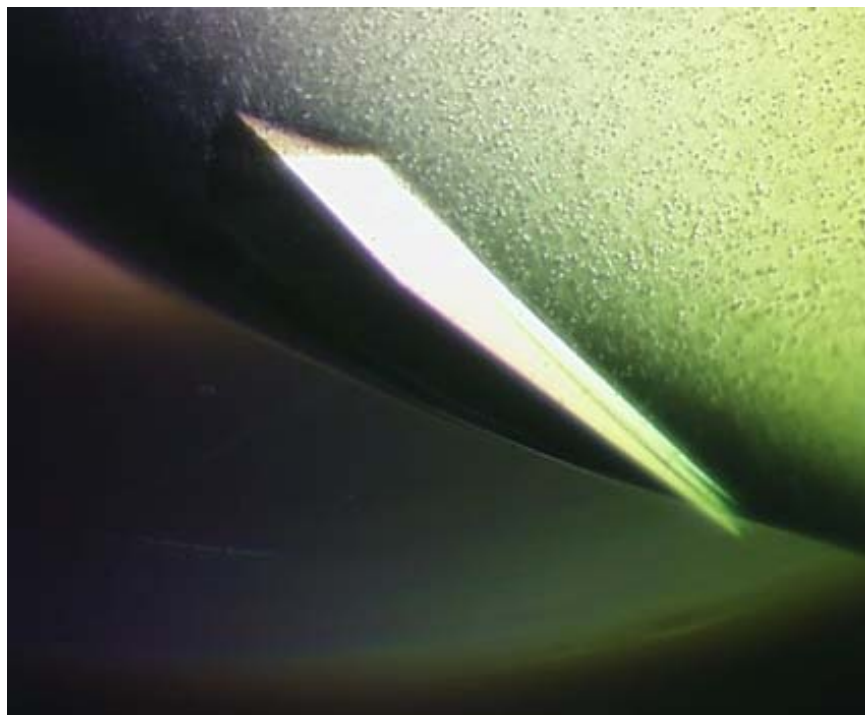


Figure 6.2 Ternary complex crystal structure of PanK_{Tm} • pantothenate • ADP

The crystals were grown using sitting drop vapor diffusion method at 4°C. The optimized condition contains 16% PEG 3350. The protein concentration was 18 mg/ml.



Figure 6.3 Binary complex crystal structure of PanK_{Tm} • phosphor-pantothenate

The crystals were grown using sitting drop vapor diffusion method at 4°C. The optimized condition contains 16% PEG 3350. The protein concentration was 18 mg/ml.

carried out using Refmac (Murshudov et al., 1997) of CCP4 package (Collaborative Computational Project, 1994) and coot (Emsley and Cowtan, 2004) starting with the dimer of the apo-PanK_{Tm}-III model determined previously (Yang et al., 2006). After first round of refinement, the *Fo-Fc* difference map revealed clear density for the bound ligands. These ligand molecules, Pan, ADP, and P-Pan, were built in the model based on the difference electron densities. The solvent molecules were added subsequently using coot (Emsley and Cowtan, 2004). The program PROCHECK (Laskowski et al, 1993) was used to evaluate the quality of the structures. All figures are made by PyMOL program (DeLano, 2002). The data collection and final refinement statistics are given in Table 6.1.

Protein structure accession number

Coordinates of PanK_{Tm}-III•Pan, PanK_{Tm}-III•Pan•ADP, and PanK_{Tm} -III•P-pan have been deposited in the Research Collaboratory for Structural Bioinformatics (RCSB) Protein Data Bank under accession code 2XXX, 2XXX, 2XXX respectively.

Results and Discussions

The overall structures of PanK_{Tm}-III complexes

Table 6.1 Data collection and refinement statistics for complex crystals

	PanK _{Tm} •Pan binary complex	PanK _{Tm} •PiPan binary complex	PanK _{Tm} •Pan•ADP ternary complex
General			
Wavelength (Å)	0.9790	0.9790	1.5418
Resolution (Å)	50-1.51 Å	50-1.63 Å	50-2.30 Å
Total no. of observations	220,185	176,024	62,103
No. of unique reflections	56,457	45,134	15,147
% Completeness (% in outer shell)	99.7 (98.4)	99.4 (94.0)	98.9 (97.6)
R_{sym} (outer shell)	0.061 (0.361)	0.057 (0.356)	0.066 (0.251)
I/σ (outer shell)	11.3 (5.0)	18.9 (9.2)	44 (7.9)
Refinement			
Resolution range (Å)	30.00-1.51 Å	30.00-1.63 Å	30.00-2.30 Å
R_{work}	18.4 (%)	18.5 (%)	19.0 (%)
R_{free}	21.3 (%)	22.9 (%)	25.7 (%)
Protein atoms (Avg. Bfactor) (Å ²)	11574 (19.3)	11528 (20.3)	11574 (20.9)
Solvent atoms (Avg. Bfactor)(Å ²)	1807 (34.1)	1652 (34.2)	648 (23.9)
Pan atoms (Avg. Bfactor) (Å ²)	15 (13.7)		15 (14.5)
P-pan atoms (Avg. Bfactor) (Å ²)		19 (19.0)	
ADP atoms (Avg. Bfactor) (Å ²)			27 (43.4)
R.m.s.d. bond length (Å)	0.009	0.010	0.009
R.m.s.d. bond angle (°)	1.37	1.49	1.41
Ramachandran Plot			
% in most favored region	92.0	92.5	92.5
% in additional allowed region	8.0	7.4	7.5
% in disallowed region	0	0.1	0.1

$$^a R_{\text{sym}} = \sum_{hkl} \{ (\sum_j |I_j| - \langle I \rangle) \sum_j |I_j| \}$$

$$^b R_{\text{work}} = \sum_{hkl} |F_o - F_c| / \sum_{hkl} |F_o|, \text{ where } F_o \text{ and } F_c \text{ are the observed and calculated structure factors, respectively.}$$

^c Five percent of the reflections were used in the calculation of R_{free} .

Previously, the structures of apo-PanK_{Tm}-III (Yang et al., 2006) and the structure of PanK_{Pa}-III complexed with pantothenate were reported (Hong et al., 2006). Here we have also obtained crystal structures of two binary complexes of PanK_{Tm}-III with substrate pantothenate and product phospho-pantothenate, respectively, and a ternary complex PanK_{Tm}-III•Pan•ADP. The overall structures of these complexes are very similar to the apo enzyme (Yang et al., 2006), with average root mean square deviation (rmsd) between C_α positions of ~0.3 Å. There appear to be no significant conformation differences in PanK_{Tm}-III complexes compared to the apo-enzyme under the crystallization condition.

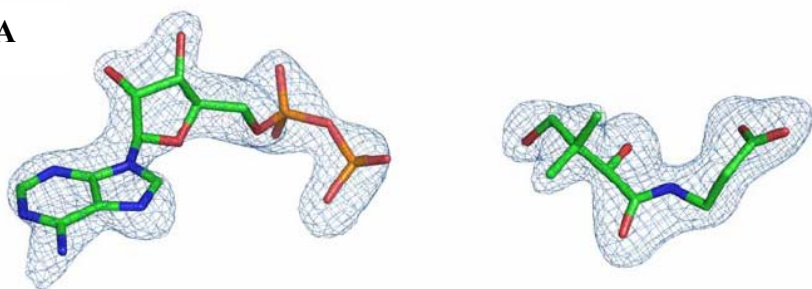
The structure of PanK_{Tm}-III is also very similar to that of PanK_{Pa}-III. The two enzymes share ~23% sequence identity and the average rmsd between their C_α atoms is 1.88 Å. In contrast, *Sa*CoaA, a type-II PanK, though also of ASHKA protein superfamily is much more divergent from PanK-III, with a large rmsd (2.87 Å) ~180 superimposable C_α atoms. PanK-II and PanK-III share a common core of the same ASKHA fold that contains a duplication of two domains with each domain consisting of a five-stranded mixed β-sheet and three α helices. The substrate binding and catalytic site is located at the domain interface. The signature motifs of the superfamily such as ADENOSINE 1, PHOSPHATE 1 and PHOSPHATE 2 are conserved in both PanK-II and PanK-III (Cheek et al., 2005; Cheek et al., 2002; Nicely et al., 2007; Yang et al., 2006). However, there are significant differences in the periphery structure elements and the organization of

the dimer in two enzymes, which result in substantially different substrate binding modes (See below).

Pantothenate binding site

In the PanK_{Tm}-III•Pan complex, as well as the ternary PanK_{Tm}-III•Pan•ADP complex, the conformation of the bound pantothenate molecule is well defined by the unambiguous electron density (Figure 6.4A). The bound pantothenate is located in a small cavity at the domain interface and forms extensive and highly specific interactions with protein residues from both monomers of PanK_{Tm}-III dimer (Figure 6.5). At one end of the molecule, the C1 carboxyl group forms two hydrogen bonds with the side chain of Arg106 and Thr179B of the second monomer, while at the other end, C2' and the C4' hydroxyl groups are hydrogen bonded to the side chain of Asp105. The 3' dimethyl groups of pantothenate make van der Waals contact with the side chains of Ile145 and Leu163B from the second monomer. Two structural motifs unique to PanK-III contribute to shape the Pan binding pocket. The first is the “Pan motif” identified recently by Nicely *et al.* (Nicely et al., 2007) that comprised the loop connecting the last β -strand of the N-terminal domain (β 5) to α 3 helix, including residues D105 and R106 (Figure 6.5). This loop is a unique insertion element to the ASKHA core and not present in PanK-II. Together with the

A



B

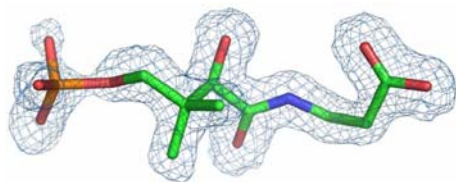


Figure 6.4 Electron density of the bound substrate and product

A. *Fo-Fc* omit map for ADP and pantothenate in PanK_{Tm}-III ternary complex, contoured at 3.5 σ .

B. *Fo-Fc* omit map for phosphopantothenate in PanK_{Tm}-III product complex, contoured at 2.5 σ .

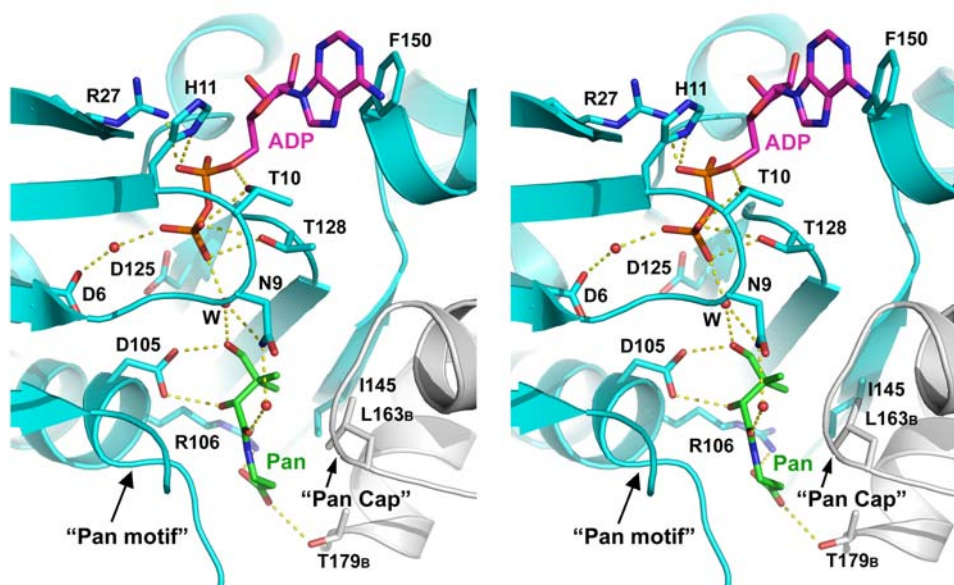


Figure 6.5 Stereoview of pantothenate and ATP binding sites of PanK_{Tm}-III

Pantothenate molecule (Pan) is colored green and ADP magenta. The two monomers in PanK_{Tm}-III dimer are colored cyan and gray, respectively. Protein side chains that interact with substrates are drawn as sticks. Active site water molecules are shown small red spheres. Hydrogen bonds are represented by dotted lines.

second “Pan Cap” motif from the second monomer of the dimer, they complete the enclosed Pan-binding pocket.

Superimposition of the pantothenate binding pocket of PanK_{Tm}-III with that of PanK_{Pa}-III revealed that the position and conformation of Pan in the two structures are essentially identical (Figure 6.6). Most of the Pan-interacting residues are conserved in both enzymes. These include Asn9, Asp105, Arg106, Ile145, Ile160 and Thr179 of PanK_{Tm}-III, corresponding to Asn9, Asp101, Arg102, Ile142, Ile160, and Thr180 of PanK_{Pa}-III, respectively. A notable difference in Pan binding sites of the two enzymes is the presence of Tyr92 in PanK_{Pa}-III, which forms a hydrogen bond to the carboxylate group of Pan. The corresponding residue in PanK_{Tm}-III is Val90. Therefore PanK_{Pa}-III appears to have a more enclosed Pan binding pocket than PanK_{Tm}-III, and contains an additional specific H-bond between enzyme and the substrate (Figure 6.6). Nevertheless, it is likely that the general pantothenate binding mode is highly conserved in all Type III PanKs given the high degree of conservation of most of pantothenate interacting residues among PanK-III.

ATP binding site

Despite a very low affinity between PanK-III and nucleotide ($K_d \sim 3$ mM), we were able to obtain a PanK_{Tm}-III•Pan•ADP ternary complex structure, which

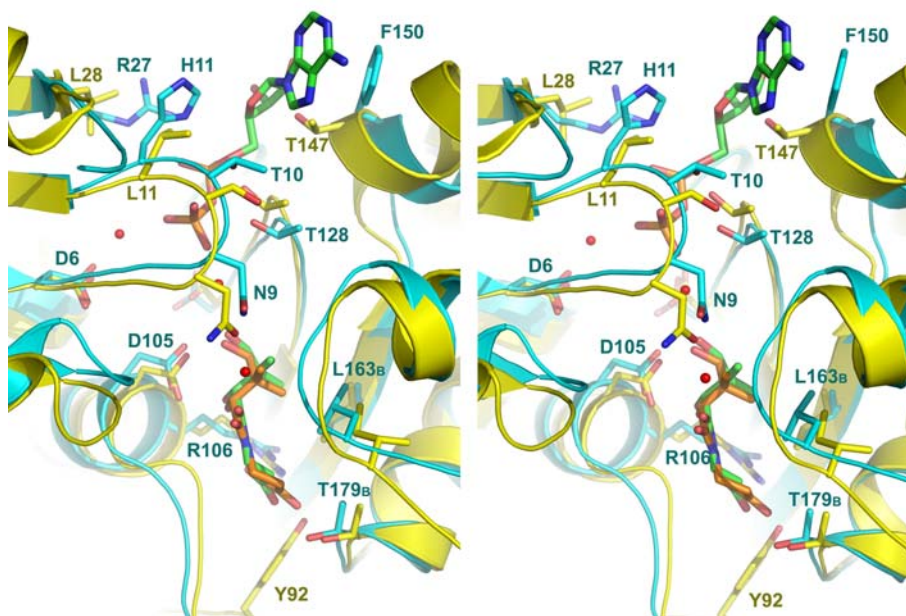


Figure 6.6 Superposition of substrate binding site of PanK_{Tm}-III (Cyan) and PanK_{Pa}-III (yellow)

The pantothenate and ADP bound to PanK_{Tm}-III are shown as a stick representation with green carbons. The pantothenate bound to PanK_{Pa}-III is shown as a stick representation with orange carbons.

offers a first view of the interactions between the nucleotide and enzyme. Notably, the B-factors of the bound ADP molecule are higher than that of surrounding protein atoms (43 \AA^2 vs. 20 \AA^2), indicating partial occupancy or disordering. Nevertheless the electron density for ADP is clear enough to allow unambiguous assignment of the adenine base, ribose and the two phosphate groups (Figure 6.4A).

The ADP molecule binds in the cleft between the two domains of the enzyme (Figure 6.5), in a position generally conserved for nucleotide binding in the ASKHA superfamily (Cheek et al., 2005; Cheek et al., 2002). Most interactions between ADP and the enzyme are formed through the two phosphate groups of ADP. Here the β -phosphate of the bound ADP is hydrogen bonded to the side chains of Thr10 and Thr 128 directly, and to Asp6 and Asn9 (both are part of the “PHOSPHATE 1” motif) indirectly through water molecules. The β -phosphate of ADP also interacts with the C4' hydroxyl group of Pan via a water molecule (labeled “W” in Figure 6.5), which is coordinated by the side chain of Asn9. This water molecule appears to occupy a position corresponding to that of the γ -phosphate of the ATP. It is likely that Asn9 would directly interact with the transferring γ -phosphate of ATP and thus play an important role in catalysis. The α -phosphate of ADP also forms several hydrogen bonded with the enzyme, directly to the side chains of Thr10, His11, and Arg27. There are essentially no specific interactions observed between the enzyme and the ribose group of ADP

(Figure 6.5). The interaction between enzyme and the adenine ring of ADP is also very limited, involving primarily a weak stacking interaction with the side chain of Phe150. The ATP binding pocket of PanK_{Tm}-III opens up considerably to the solvent and does not support tight interactions with the adenosyl moiety of ATP (Figure 6.5).

As shown in the nucleotide specificity assays (Figure 5.4 and (Hong et al., 2006)), PanK-III from either *T. maritima* or *P. aeruginosa* does not discriminate strongly among various NTPs or dNTPs, though a general trend is observed that both enzymes appear to work better with purine nucleotides than with pyrimidine nucleotides (Figure 5.4 and (Hong et al., 2006)). It is possible that since adenine and guanine bases have bigger surface areas than the pyridine bases, they may form a slightly stronger van der Waals interaction with the enzyme and thus may have a somewhat higher binding affinity. Since there are few specific direct interactions between PanK-III and the ribose moiety of the bound nucleotide, PanK-III is nearly equally as active when using dNTP as phosphoryl donor (Hong et al., 2006).

Currently, no experimentally determined structure of PanK_{Pa}-III complex with nucleotide is reported. Superposition of PanK_{Tm}-III•Pan•ADP with PanK_{Pa}-III brings ADP into PanK_{Pa}-III active site without any steric hindrance, thus affords a model of ADP binding in PanK_{Pa}-III. In this model, Ser10 and Thr124 of PanK_{Pa}-III, corresponding to Thr10 and Thr128 of PanK_{Tm}-III, are well

conserved. However, His11, Arg27, and Phe150 in PanK_{Tm}-III, all involved in ATP binding, are changed to Leu11, Leu28, and Thr147 in PanK_{pa}-III, respectively, which are no longer able to form the same specific interactions with ATP as in PanK_{Tm}-III. Therefore, the ATP binding modes in the two closely related type III PanK, PanK_{pa}-III and PanK_{Tm}-III, may yet to be considerably different.

PanK_{Tm}-III product complex structure and catalytic mechanism of PanK-III

We have determined the complex structure of PanK_{Tm}-III with product phospho-pantothenate (P-Pan, Figure 6.4B and 6.7). The less than optimal density for the terminal phosphate of the bound product indicates partial occupancy of this phosphate group, probably due to the decomposition of P-pan to pantothenate and phosphate. In this complex structure, the side chain of Asn9, previously interacting with a water molecule bridging C4' hydroxyl and β -phosphate of ADP in PanK_{Tm}-III•ADP•Pan ternary complex structure, is now hydrogen bonded to the phosphate group of P-Pan, which replaces the above mentioned water molecule. When an ADP molecule is modeled according to PanK_{Tm}-III•Pan•ADP complex structure to generate a model for the ternary product complex, the distance between one of the oxygens of ADP β -phosphate and the phosphorus of P-Pan is only ~ 3.0 Å and it is collinear with the P-O4' bond (Figure 6.7). This

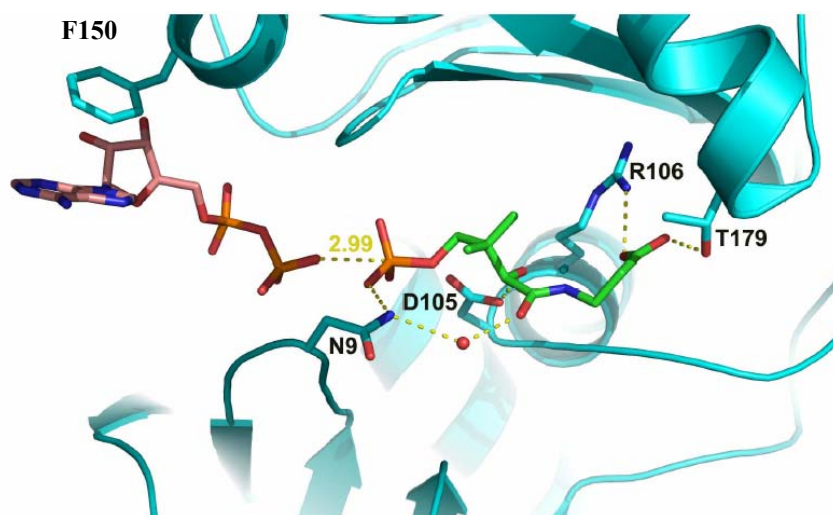


Figure 6.7 Interactions with product phospho-pantothenate (P-Pan) in PanK_{Tm}-III•P-Pan product complex

ADP is model from PanK_{Tm}-III•Pan•ADP ternary complex.

Hydrogen bonds are represented with dotted lines.

active site configuration is consistent with a mechanism in which the phosphoryl transfer proceeds via a nucleophilic attack by pantothenate C4' hydroxyl group on the γ -phosphorus of ATP, followed by direct in-line transfer of the terminal phosphate to C4' hydroxyl. The phosphoryl transfer reaction is likely to be associative in nature given the short distance between the two oxygen atoms where phosphoryl transfer occurs. Residue Asp105 is positioned ideally to act as a catalytic base activating pantothenate C4' hydroxyl group for the nucleophilic attack (Figure 6.5). Additional catalytic residues include two Asp residues of the PHOSPHATE 1 and PHOSPHATE 2 motifs, Asp6 and Asp125, respectively, which would coordinate a Mg^{2+} ion that interact with ATP β - and γ -phosphate and stabilize the reaction intermediate. The proposed roles for these active site residues have been supported by the mutagenesis analysis of both PanK_{Tm}-III and PanK_{Hp}-III (Brand and Strauss, 2005; Yang et al., 2006). Similar mechanism is thought to be prevalent in proteins of ASKHA superfamily (Cheek et al., 2005; Cheek et al., 2002) (Blattler and Knowles, 1979; Hurley, 1996).

Comparison of PanK-III and PanK-II

Although type II and type III PanKs both belong to the ASHKA superfamily, they share only limited sequence identity (~13%) that is restricted to the signature motifs of the superfamily (Cheek et al., 2005; Cheek et al., 2002),

and have distinct kinetic properties and substrate preferences (Hong et al., 2006). Structural comparison of PanK-III and recently published PanK-II structures (from both *S. aureus* and human) reveals that, while the key catalytic residues are conserved, the substrate binding sites for both ATP nucleotide and pantothenate are significantly different in the two enzymes. First, a superposition between the structures of AMNPNP bound type II PanK from *S. aureus* (PanK_{Sa}-II) (Hong et al., 2006) and the ADP bound PanK_{Tm}-III structure (Figure 6.8) shows that the conformations of the bound nucleotide in the two enzyme active sites are significantly different mostly in the adenine ribose region (Figure 6.8A), while the β - and γ - phosphates of the nucleotide are likely to occupy the same position and interact with the same conserved set of protein residues such as those in PHOSPHATE 1 (containing Asp6 and Asn9 of PanK_{Tm}-III) and PHOSPHATE 2 (containing Asp 125) motifs. Different sets of protein residues are involved in the interactions with the adenosyl moiety of the nucleotide in PanK-II and PanK-III (Figure 6.8B and 6.8C). In PanK_{Sa}-II, the adenine ring is sandwiched between side chains of Tyr137 and Leu11 (Figure 6.8C), while in type III PanK, the side chains of Phe150 and His11 (corresponding to Leu 11 of PanK_{Sa}-II) do not form sandwich-like interactions with the adenine ring (Figure 6.8B). There are also additional specific hydrogen bonds and hydrophobic interactions between PanK_{Sa}-II and the nucleotide. These include Lys13 (corresponding to Val 13 in PanK_{Tm}-

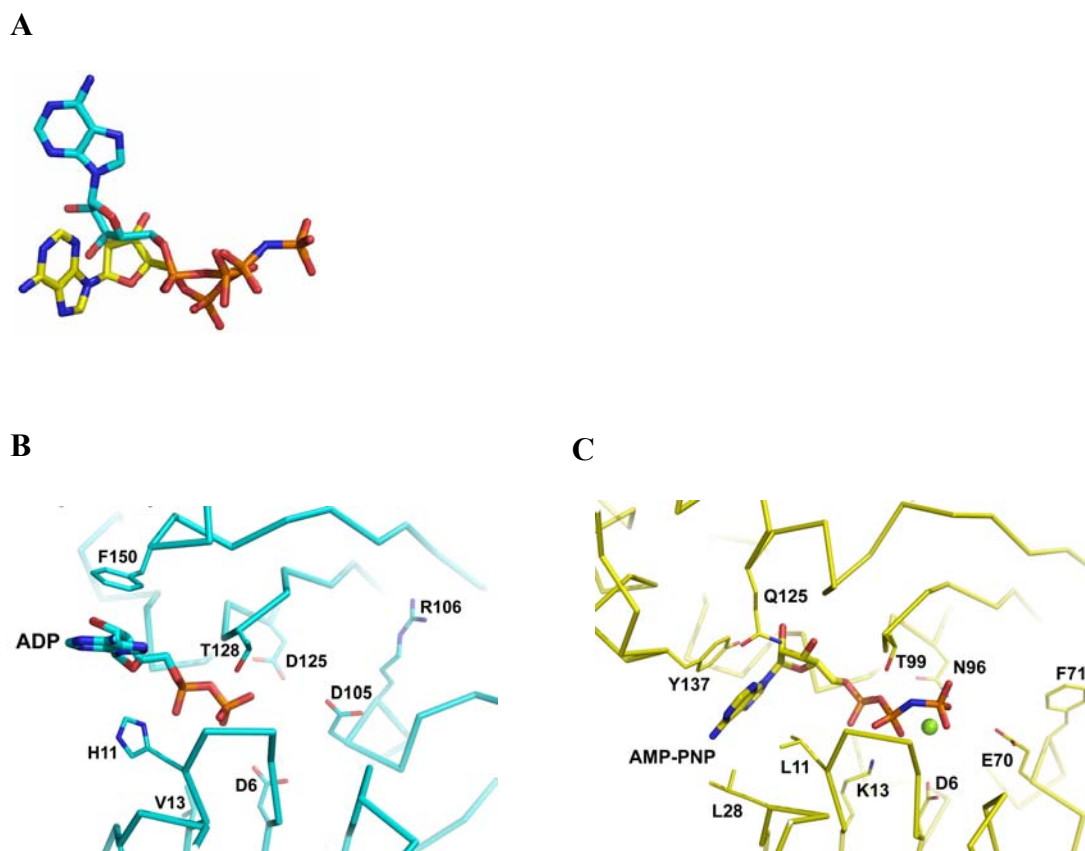


Figure 6.8 Comparison of the nucleotide binding in PanK-III and PanK-II

A. Superposition of PanK_{Tm}-III bound ADP (cyan) and PanK_{Sa}-II bound AMPPNP (yellow). This superposition is achieved by superimposed the conserved active site residues of the two enzyme structures.

B-C. A side-by-side view of the nucleotide binding site of PanK-III from *Thermotoga maritima* (**B**) and PanK-II from *Staphylococcus aureus*, SaCoA (**C**). The proteins are shown as C_α trace. The Mg²⁺ ion is shown as a green sphere. Selected protein side chains are also shown.

III) with α and β -phosphates, Gln125 and Leu28 with the adenine base (Figure 6.8C). Overall, type II PanK, as represented here by PanK_{Sa}, has a better fitted ATP binding pocket that enable more specific hydrogen bonds and van der Waals interactions with the nucleotide than type III PanK.

At present, a complex structure of type II PanK with pantothenate substrate has not been reported. In the absence of an experimentally determined PanK-II•Pan complex structure, the structures of two human PanK isoforms (PANK1 and PANK2) complexed with a physiological feedback inhibitor acetyl-CoA (pdb code 2i7n and 2i7p) may allow a model for pantothenate bound in PanK-II, assuming that Pan substrate would occupy the same position as that of the pantothenate moiety of acetyl-CoA (Figure 6.9). This pantothenate binding mode in human PanK is drastically different from that in PanK-III (Figure 6.9). When the active site residues of PanK_{Tm}-III and human PanK are superimposed, the pantothenate molecules in the two structures are orientated very differently and interact with different set of protein residues. A closer inspection of PanK-II complex structure shows that its pantothenate binding site is also located near the dimer interface. However, unlike PanK-III where the dimer interface resulted in a small and enclosed binding pocket, the PanK-II dimer interface leaves a long and largely solvent accessible channel that could accommodate both pantothenate and much longer CoA, as well as the N-substituted pantothenamides.

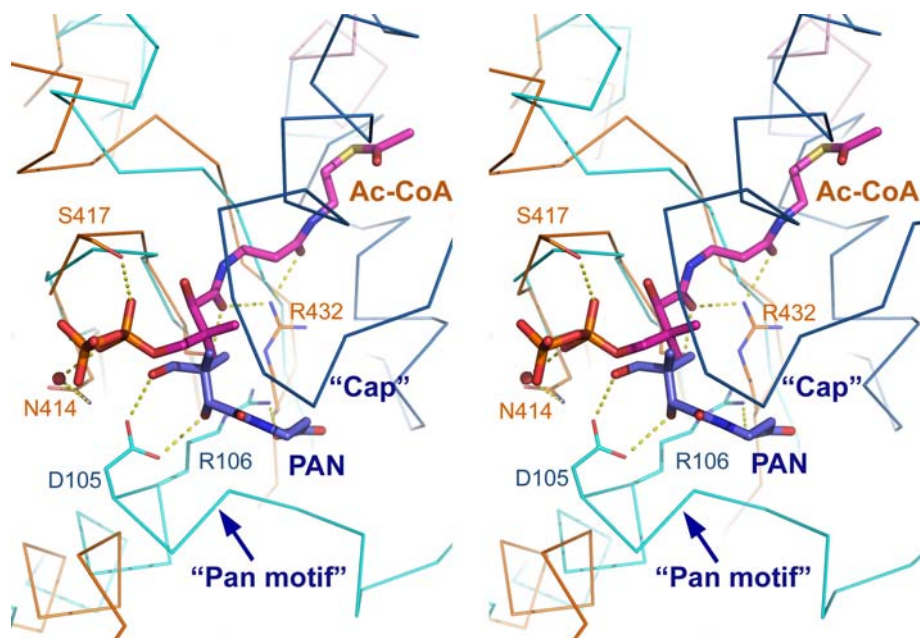


Figure 6.9 Stereoview of the superposition of the pantothenate binding sites in PanK_{Tm}-III (light and dark blue) and human PanK-II (orange)

Only C_a traces of proteins are shown. PanK_{Tm}-III Pantothenate (blue) bound to PanK_{Tm} and acetyl-CoA (magenta) in *hs*PanK-II•acetyl-CoA complex are shown as thick sticks. Selected protein side chains are shown in thin sticks. Dotted lines represent hydrogen bonds. The water molecule is shown as a small red sphere.

In summary, the present work including high-resolution three-dimensional structures of PanK_{Tm}-III complexes investigated substrate-enzyme interactions for both ATP and pantothenate and provided a structural explanation for several unique enzymatic properties of the enzyme, such as the unusually low affinity between ATP nucleotide and enzyme. Comparison of substrate binding and catalytic sites of PanK-III with that of eukaryotic PanK-II revealed drastic differences in the binding modes for both ATP and pantothenate. These differences may be exploited to design highly specific inhibitors that target PanK-III.

CHAPTER VII

Conclusions and Future Directions

Pantothenate kinase catalyzes the first committed step in the five-step universal pathway of CoA biosynthesis. Three types of PanK have been characterized so far. Prokaryotic PanK (PanK-I) and eukaryotic PanK (PanK-II) were identified previously, and most of them are feedback regulated by CoA, the end product of the pathway, and its thioesters. Recently, a novel PanK protein encoded by the gene *coaX* was identified by genetic complementation. This new PanK was named type III PanK (PanK-III). Compared with the previously characterized PanK-I (exemplified by the *E. coli coaA*-encoded protein) and eukaryotic PanK-II, PanK-III has a wider phylogenetic distribution and exists in many pathogenic species. It also displays other unique enzymatic characteristics like resistance to the inhibition by CoA or its thioesters, high K_m value (about 10 mM) for ATP and inability to utilize pantothenate antimetabolite as substrate. The focus of my research is to unravel the underlying structural basis for these unique enzymatic properties of PanK-III through crystallography and other biochemical methods.

We solved the first crystal structure of type III PanK from *Thermotoga maritima* (PanK_{Tm}) at 2.0 Å resolution. As the structure reveals, PanK-III belongs to the acetate and sugar kinase/heat shocks protein 70/actin (ASKHA)

protein superfamily, like type II PanK, whereas type I PanK belongs to the P-loop kinase superfamily. Comparative structural analysis of the PanK_{Tm} active site configuration and mutagenesis of three highly conserved active site aspartates identify these residues as critical for PanK-III catalysis. Furthermore, the analysis also provides an explanation for the lack of CoA feedback inhibition by the enzyme. Since PanK-III adopts a different structural fold from that of the *E. coli* PanK -- which is a member of the "P-loop kinase" superfamily -- this finding represents yet another example of convergent evolution of the same biological function from a different protein ancestor.

In order to gain more insight into the structural mechanism and function of PanK-III, we also solved the crystal structures of two binary complexes of PanK-III with substrate pantothenate and product phospho-pantothenate, respectively, as well as a ternary complex of PanK-III with pantothenate and ADP. Combined with isothermal titration calorimetry, we present a detailed structural and thermodynamic characterization of the interactions between PanK-III and its substrates ATP and pantothenate. Comparison of substrate binding and catalytic sites of PanK-III with that of eukaryotic PanK-II revealed drastic differences in the binding modes of both ATP and pantothenate, even though both PanK-II and PanK-III belong to the same ASKHA superfamily and may share a common catalytic mechanism.

In contrast to the long known prokaryotic PanK-I and eukaryotic PanK-II, the PanK-III is refractory to feedback inhibition by CoA and its thioesters, do not accept the pantothenate antimetabolite N-substitute pantothenamide as a substrate, nor are inhibited by pantothenamide. These are the main differences between PanK-III and the other PanKs classified so far. Studies with type I PanK_{Ec} and type II PanK_{Sa} have shown that *N*-pentylpantothenamide act as antimicrobial agents through their function as CoA antimetabolites. These compounds act as substrates of the CoA biosynthetic enzymes in both *E. coli* and *S. aureus*, and they are converted to inactive CoA analogues, and inhibit CoA-dependent cellular processes. Structural comparison of the active site of PanK_{Tm} and that of PanK_{Ec} provided a structural explanation for the lack of feedback inhibition by CoA and the inability to phosphorylate the N-alkylpantothenamide antimetabolites by PanK-III.

Despite extensive studies on inhibitors against PanK-I and PanK-II, no specific PanK-III inhibitor has been reported. As many bacteria harboring PanK-III are pathogenic and have no alternative pantothenate kinase enzymes in their genomes, finding inhibitors that specifically target this group of pantothenate kinase with new strategies would be invaluable to the formulation of new drugs against pathogenic bacteria. In order to design drugs against the PanK-III-harboring pathogenic bacteria, several future directions may be pursued.

First, high-throughput chemical compound screening may be employed to identify potent and specific PanK-III inhibitors (Baldwin et al, 2005). Efforts should be made to select the most appropriate library for screening, such as a scaffold library that is a collection of drug-like compounds with maximum common substructures (scaffolds), like a substrate (pantothenate) analogue scaffold library. If such a library is not available, a synthetic library of pantothenate derivatives should be helpful. An ATP analogue scaffold library should be avoided since it might not be specific due to the fact that most kinases use ATP as substrate. Steady state kinetics analysis needs to be performed to determine the inhibition properties of the hit compounds, followed by crystal structural analysis of PanK-III in complex with the potent and selective inhibitors. Information provided by the complex crystal structure will be valuable for understanding the mechanism of the inhibition and provide molecular basis for future improvement of the inhibitors as potential drugs.

Secondly, a relatively new hit-generating method of *in silico* or virtual screening may be pursued since the crystal structures of PanK-III from several organisms are available (Kitchen et al, 2004; Tautz and Mustelin, 2007). Through virtual screening, very large libraries of compounds can be automatically evaluated using computer programs and reduced to a manageable size for direct *in vitro* testing. The basic technique for virtual screening is called “docking”, which is a computational method that samples conformations of compounds in substrate-

binding sites and applies scoring functions to rank molecules by a free-energy-like term. Many commercial software packages with different docking algorithms are available, among which the most successfully applied algorithms are DOCK, FlexX, GOLD, AutoDock, GLIDE, QXP, and ICM.

Thirdly, structure-based lead optimization may be explored. PanK-III inhibitors may be optimized starting from a co-crystal structure of PanK-III with pantothenate (Kitchen et al, 2004). Modeling of analogues will result in some compounds with high potency, which will be experimentally confirmed. In addition, since the N-substituted pantothenamides show inhibition of type I and type II PanKs, and the enclosed pantothenate binding pocket of type III PanK can not accommodate the long tails of pantothenamides, modifications could be introduced at the other parts of pantothenate molecule to exploit the specific interactions with the active site of PanK-III. Molecular docking can be used to predict whether PanK-III could interact with such optimized inhibitors. The resulting potent compounds will be further tested *in vitro*.

Finally, new molecules that could bind to the active sites of PanK-III may be generated with a new algorithm, CombiSMoG for "combinatorial small molecule growth". The advantage of this method has been demonstrated by its application in the design of picomolar inhibitors for human carbonic anhydrase (Grzybowski et al, 2002). CombiSMoG should be applicable for PanK-III since its ligand bound conformation has now been very well characterized.

Pantothenate will be chosen as the starting fragment for CombiSMoG design. The growth algorithm will generate thousands of candidate ligands, from which the top candidates will be chosen for further analysis, including synthesizing the compounds, testing the binding affinities and analyzing the X-ray structures of the complexes.

APPENDIX A

Phylogenetic distribution of Pantothenate Kinase Types

Organisms are colored according to the following legend:

Archaeal PanK
PanK-I
PanK-II
PanK-III
PanK-I and PanK-II
PanK-II and PanK-III

	Phylogeny	Organism	PANK-I	PANK-II	PANK-III
ARCHAEA					
1	Crenarchaeota	<i>Aeropyrum pernix</i> K1 [A]			
2		<i>Sulfolobus acidocaldarius</i> DSM 639 [A]			
3		<i>Sulfolobus solfataricus</i> P2 [A]			
4		<i>Sulfolobus tokodaii</i> str. 7 [A]			
5		<i>Pyrobaculum aerophilum</i> str. IM2 [A]			
6	Euryarchaeota	<i>Archaeoglobus fulgidus</i> DSM 4304 [A]			
7		<i>Haloarcula marismortui</i> ATCC 43049 [A]			
8		<i>Halobacterium</i> sp. NRC-1 [A]			
9		<i>Methanothermobacter thermautotrophicus</i> str. Delta H [A]			
10		<i>Methanocaldococcus jannaschii</i> DSM 2661 [A]			
11		<i>Methanococcoides burtonii</i> DSM 6242 [A]			
12		<i>Methanosarcina acetivorans</i> C2A [A]			
13		<i>Methanosarcina barkeri</i> str. fusaro [A]			
14		<i>Methanosarcina mazei</i> Go1 [A]			
15		<i>Methanopyrus kandleri</i> AV19 [A]			
16		<i>Pyrococcus abyssi</i> GE5 [A]			
17		<i>Pyrococcus furiosus</i> DSM 3638 [A]			
18		<i>Pyrococcus horikoshii</i> OT3 [A]			
19		<i>Thermococcus kodakaraensis</i> [A]			
20		<i>Ferroplasma acidarmanus</i> [A]	gi68140748		
21		<i>Thermoplasma volcanium</i> GSSI [A]	uniQ97CG3		
BACTERIA					
22	Actinobacteria	<i>Corynebacterium diphtheriae</i> NCTC 13129 [B]	uniQ6NI48		
23		<i>Corynebacterium efficiens</i> YS-314 [B]	uniQ8FOR2		
24		<i>Corynebacterium glutamicum</i> ATCC 13032 [B]	gi19552218		
25		<i>Corynebacterium glutamicum</i> ATCC 13032 [B]	gi41325219		
26		<i>Corynebacterium jeikeium</i> K411 [B]	uniQ4JU68		
27		<i>Mycobacterium avium</i> subsp. <i>paratuberculosis</i> str. k10 [B]	uniQ73WG0		uniQ743Y3
28		<i>Mycobacterium leprae</i> TN [B]	uniQ9X795		uniQ9CD56
29		<i>Mycobacterium marinum</i> M [B]	739		74
30		<i>Mycobacterium smegmatis</i> str. MC2 155 [B]	236		6059
31		<i>Mycobacterium bovis</i> AF2122/97 [B]	uniP63810		uniQ06282
32		<i>Mycobacterium microti</i> OV254 [B]	fig1806.1.pcg.333		uniQ06282
33		<i>Mycobacterium tuberculosis</i> CDC1551 [B]	uniP63810		uniQ06282
34		<i>Mycobacterium tuberculosis</i> H37Rv [B]	uniP63810		uniQ06282
35		<i>Nocardia farcinica</i> IFM 10152 [B]	uniQ5YQ72		uniQ5Z2U1
36		<i>Kineococcus radiotolerans</i> SRS30216 [B]	gi46364458		
37		<i>Tropheryma whippelii</i> TW08/27 [B]	uniQ83IA2		
38		<i>Tropheryma whippelii</i> str. Twist [B]	uniQ83GU4		
39		<i>Leifsonia xyli</i> subsp. <i>xyli</i> str. CTCB07 [B]	uniQ6AD31		

40		<i>Propionibacterium acnes</i> KPA171202 [B]	uni Q6A6T7		
41		<i>Streptomyces avermitilis</i> MA-4680 [B]	uni Q82DL5		uni Q82EC5
42		<i>Streptomyces coelicolor</i> A3(2) [B]	uni Q86779		uni Q9X8N6
43		<i>Thermobifida fusca</i> YX [B]	uni Q47LM9		uni Q47KV7
44		<i>Bifidobacterium longum</i> DJO10A [B]			uni Q8G558
45		<i>Bifidobacterium longum</i> NCC2705 [B]			uni Q8G558
46		<i>Rubrobacter xylanophilus</i> DSM 9941 [B]			gi 45547863
47		<i>Symbiobacterium thermophilum</i> IAM 14863 [B]			uni Q67JI4
48	Aquificae	<i>Aquifex aeolicus</i> VF5 [B]			uni Q67753
49	Bacteroidetes/	<i>Bacteroides fragilis</i> 638R [B]			fig 817.1.pcg.91_6
50	Chlorobi	<i>Bacteroides fragilis</i> NCTC9343 [B]			uni Q5LGL2
51		<i>Bacteroides fragilis</i> ATCC 25285 [B]			uni Q5LGL2
52		<i>Bacteroides fragilis</i> YCH46 [B]			uni Q64XF4
53		<i>Bacteroides thetaiotaomicron</i> VPI-5482 [B]			uni Q89ZL1
54		<i>Porphyromonas gingivalis</i> W83 [B]			uni Q7MWY2
55		<i>Cytophaga hutchinsonii</i> [B]			gi 48855427
56		<i>Chlorobium tepidum</i> TLS [B]			uni Q8KCK7
57		<i>Pelodictyon luteolum</i> DSM 273 [B]			gi 78187267
58		uncultured <i>Chlorobi</i> bacterium [B]			_325
59	Chlamydiae	<i>Chlamydia muridarum</i> Nigg [B]			
60		<i>Chlamydia trachomatis</i> A/HAR-13 [B]			
61		<i>Chlamydia trachomatis</i> D/UW-3/CX [B]			
62		<i>Chlamydia abortus</i> [B]			
63		<i>Chlamydia abortus</i> S26/3 [B]			
64		<i>Chlamydia caviae</i> GPIC [B]			
65		<i>Chlamydia pneumoniae</i> AR39 [B]			
66		<i>Chlamydia pneumoniae</i> CWL029 [B]			
67		<i>Chlamydia pneumoniae</i> J138 [B]			
68		<i>Chlamydia pneumoniae</i> TW-183 [B]			
69		<i>Parachlamydia</i> sp. UWE25 [B]			
70	Chloroflexi	<i>Chloroflexus aurantiacus</i> [B]	gi:76166188		
71		<i>Dehalococcoides ethenogenes</i> 195 [B]			gi 57234779
72		<i>Dehalococcoides</i> sp. CBDB1 [B]			gi 73748270
73	Cyanobacteria	<i>Synechococcus elongatus</i> PCC 6301 [B]			uni Q5MZIP5
74		<i>Synechococcus elongatus</i> PCC 7942 [B]			gi 46130103
75		<i>Synechococcus</i> sp. CC9605 [B]			gi 78213816
76		<i>Synechococcus</i> sp. CC9902 [B]			gi 78183964
77		<i>Synechococcus</i> sp. WH 8102 [B]			uni Q7U4A8
78		<i>Synechocystis</i> sp. PCC 6803 [B]			uni P74045
79		<i>Thermosynechococcus elongatus</i> BP-1 [B]			uni Q8DJS3
80		<i>Gloeobacter violaceus</i> PCC 7421 [B]			uni Q7NCI5
81		<i>Anabaena variabilis</i> ATCC 29413 [B]			gi 75908022
82		<i>Nostoc punctiforme</i> PCC 73102 [B]			gi 23123588
83		<i>Nostoc</i> sp. PCC 7120 [B]			uni Q8YQD7
84		<i>Prochlorococcus marinus</i> str. MIT 9312 [B]			gi 78778468
85		<i>Prochlorococcus marinus</i> str. MIT 9313 [B]			uni Q7VSE1
86		<i>Prochlorococcus marinus</i> str. NATL2A [B]			uni Q46HU6
87		<i>Prochlorococcus marinus</i> subsp. <i>marinus</i> str. CCMP1375 [B]			uni Q7VEB9
88		<i>Prochlorococcus marinus</i> subsp. <i>pastoris</i> str. CCMP1986 [B]			uni Q7V3J5
89	Deinococcus/	<i>Deinococcus geothermalis</i> DSM11300 [B]			uni Q4H9S6
90	Thermus	<i>Deinococcus radiodurans</i> RI [B]			uni Q9RX54
91		<i>Thermus thermophilus</i> HB27 [B]			uni Q72IX3
92		<i>Thermus thermophilus</i> HB8 [B]			uni Q5SIJ5
93	Firmicutes	<i>Bacillus anthracis</i> str. 'Ames Ancestor' [B]		uni Q81PB2	uni Q63HD3
94		<i>Bacillus anthracis</i> str. Ames [B]		uni Q81PB2	uni Q63HD3
95		<i>Bacillus anthracis</i> str. Sterne [B]		uni Q81PB2	uni Q63HD3

96	<i>Bacillus cereus</i> ATCC 10987 [B]		uni Q736G1	uni Q63HD3
97	<i>Bacillus cereus</i> ATCC 14579 [B]		uni Q81C81	uni Q81J81
98	<i>Bacillus cereus</i> ZK [B]		uni Q63A51	uni Q63HD3
99	<i>Bacillus cereus</i> G9241 [B]		uni Q4MPF0	uni Q4MH84
100	<i>Bacillus thuringiensis</i> serovar konkukian str. 97-27 [B]		uni Q6HHK0	uni Q63HD3
101	<i>Oceanobacillus iheyensis</i> HTE831 [B]		uni Q8EN08	uni Q8EU15
102	<i>Bacillus clausii</i> KSM-K16 [B]	uni Q5WEY6		uni Q5WLV5
103	<i>Bacillus halodurans</i> C-125 [B]	uni Q9K8X7		uni Q9KGH5
104	<i>Bacillus licheniformis</i> ATCC 14580 [B]	uni Q62T53		uni Q62ZU0
105	<i>Bacillus subtilis</i> subsp. subtilis str. 168 [B]	uni P54556		uni P37564
106	<i>Exiguobacterium</i> sp. 255-15 [B]			gi 68055909
107	<i>Geobacillus kaustophilus</i> HTA426 [B]			uni Q5L3T0
108	<i>Geobacillus stearothermophilus</i> 10 [B]			uni Q5L3T0
109	<i>Listeria innocua</i> Clip11262 [B]	uni Q92D94		uni Q92F54
110	<i>Listeria monocytogenes</i> EGD-e [B]	uni Q4EKB4		uni Q4ENP5
111	<i>Listeria monocytogenes</i> str. 1/2a F6854 [B]	uni Q4EKB4		uni Q4ENP5
112	<i>Listeria monocytogenes</i> str. 4b F2365 [B]	uni Q4EKB4		uni Q4EHG3
113	<i>Listeria monocytogenes</i> str. 4b H7858 [B]	uni Q4EKB4		uni Q4EHG3
114	<i>Staphylococcus aureus</i> subsp. aureus COL [B]		uni Q5HE70	
115	<i>Staphylococcus aureus</i> subsp. aureus MRSA252 [B]		uni Q6GEU5	
116	<i>Staphylococcus aureus</i> subsp. aureus MSSA476 [B]		uni Q5HE70	
117	<i>Staphylococcus aureus</i> subsp. aureus MW2 [B]		uni Q5HE70	
118	<i>Staphylococcus aureus</i> subsp. aureus Mu50 [B]		uni Q6GEU5	
119	<i>Staphylococcus aureus</i> subsp. aureus N315 [B]		uni Q6GEU5	
120	<i>Staphylococcus epidermidis</i> ATCC 12228 [B]		uni Q8CRM3	
121	<i>Staphylococcus haemolyticus</i> JCSC1435 [B]		uni Q4L811	
122	ATCC 15305 [B]		uni Q49Z76	
123	<i>Staphylococcus aureus</i> subsp. aureus NCTC 8325 [B]		uni Q5HE70	
124	<i>Lactobacillus delbrueckii</i> subsp. bulgaricus ATCC BAA-365 (JGI) [B]	fig 1585.2.peg.32		
125	<i>Lactococcus lactis</i> subsp. cremoris SK11 (JGI) [B]	gi 62462361		
126	<i>Streptococcus equi</i> subsp. zooepidemicus [B]	9		
127	<i>Streptococcus pneumoniae</i> 23F [B]	371		
128	<i>Streptococcus suis</i> [B]	gi 50591270		
129	<i>Streptococcus thermophilus</i> ACTT BAA-491 [B]	gi 62527278		
130	<i>Streptococcus uberis</i> [B]	fig 1349.1.peg.246		
131	<i>Clostridium acetobutylicum</i> ATCC 824 [B]			uni Q97EB4
132	<i>Clostridium botulinum</i> ATCC 3502 [B]			254
133	<i>Clostridium difficile</i> 630 [B]			811
134	<i>Clostridium perfringens</i> str. 13 [B]			uni Q8XHL5
135	<i>Clostridium tetani</i> E88 [B]			uni Q899H1
136	<i>Clostridium thermocellum</i> ATCC 27405 [B]			uni Q4CJC5
137	<i>Carboxydotherrmus hydrogenoformans</i> Z-2901 [B]			gi 78043564
138	<i>Moorella thermoacetica</i> ATCC 39073 [B]			gi 68270621
139	<i>Thermoanaerobacter tengcongensis</i> MB4 [B]			uni Q8R7M2
140	<i>Enterococcus faecalis</i> V583 [B]	uni Q839J7		
141	<i>Lactobacillus acidophilus</i> NCFM [B]	uni Q5FIJ3		
142	<i>Lactobacillus gasseri</i> [B]	gi 23003514		
143	<i>Lactobacillus johnsonii</i> NCC 533 [B]	uni Q74HT8		
144	<i>Lactobacillus plantarum</i> WCFS1 [B]	uni Q88Y75		
145	<i>Pediococcus pentosaceus</i> [B]	gi 48869851		
146	<i>Leuconostoc mesenteroides</i> subsp. mesenteroides ATCC 8293 [B]	gi 23023857		
147	<i>Oenococcus oeni</i> PSU-1 [B]	gi 48865550		

148		<i>Lactococcus lactis</i> subsp. <i>lactis</i> II1403 [B]	uni Q9CFM3		
149		<i>Streptococcus agalactiae</i> 2603V/R [B]	uni P63812		
150		<i>Streptococcus agalactiae</i> A909 [B]	uni P63812		
151		<i>Streptococcus agalactiae</i> NEM316 [B]	uni P63812		
152		<i>Streptococcus mitis</i> NCTC 12261 [B]	476		
153		<i>Streptococcus mutans</i> UA159 [B]	uni Q8DU31		
154		<i>Streptococcus pneumoniae</i> R6 [B]	uni Q8DOC7		
155		<i>Streptococcus pneumoniae</i> TIGR4 [B]	uni Q97RH6		
156		<i>Streptococcus pyogenes</i> M5 [B]	2		
157		<i>Streptococcus pyogenes</i> M1 GAS [B]	uni Q48TD0		
158		<i>Streptococcus pyogenes</i> MGAS10394 [B]	uni Q48TD0		
159		<i>Streptococcus pyogenes</i> MGAS315 [B]	uni Q8K7C7		
160		<i>Streptococcus pyogenes</i> MGAS5005 [B]	uni Q48TD0		
161		<i>Streptococcus pyogenes</i> MGAS6180 [B]	uni Q48TD0		
162		<i>Streptococcus pyogenes</i> MGAS8232 [B]	uni Q8P0V9		
163		<i>Streptococcus pyogenes</i> SSI-1 [B]	uni Q8K7C7		
164		<i>Streptococcus thermophilus</i> CNRZ1066 [B]	uni Q5M079		
165		<i>Streptococcus thermophilus</i> LMG 18311 [B]	uni Q5M079		
166		<i>Spiroplasma kunkelii</i> CR2-3x [B]			.202
167		<i>Mycoplasma mobile</i> 163K [B]			uni Q6KH54
168		<i>Mycoplasma penetrans</i> HF-2 [B]			uni Q8EUB0
169		<i>Mycoplasma pulmonis</i> UAB CTIP [B]			uni Q98Q93
170		<i>Mycoplasma synoviae</i> 53 [B]			uni Q4A743
171	Fusobacteria	<i>Fusobacterium nucleatum</i> subsp. <i>nucleatum</i> ATCC 25586 [B]			uni Q8RFE4
172		<i>Fusobacterium nucleatum</i> subsp. <i>vincentii</i> ATCC 49256 [B]			uni Q7P6A2
173	Planctomycetes	<i>Pirellula</i> sp. 1 [B]			uni Q7UVH5
174	Proteobacteria- alpha	<i>Caulobacter crescentus</i> CB15 [B]			uni Q9A6Z1
175		<i>Bartonella henselae</i> str. Houston-1 [B]	uni Q6G5G3		
176		<i>Bartonella quintana</i> str. Toulouse [B]	uni Q6G0Q2		
177		<i>Bradyrhizobium japonicum</i> USDA 110 [B]	uni Q89WM2		
178		<i>Nitrobacter winogradskyi</i> Nb-255 [B]	gi 75674319		
179		<i>Rhodopseudomonas palustris</i> CGA009 [B]	uni Q6ND01		
180		<i>Brucella abortus</i> biovar 1 str. 9-941 [B]	uni Q57AH1		
181		<i>Brucella melitensis</i> 16M [B]	uni P63808		
182		<i>Brucella abortus</i> [B]	uni Q57AH1		
183		<i>Brucella suis</i> 1330 [B]	uni P63808		
184		<i>Mesorhizobium loti</i> MAFF303099 [B]	uni Q98CS9		
185		<i>Mesorhizobium</i> sp. BNC1 [B]	gi 45916707		
186		<i>Agrobacterium tumefaciens</i> str. C58 [B]	uni Q8UJ92		
187		<i>Rhizobium leguminosarum</i> bv. <i>viciae</i> 3841 [B]	2		
188		<i>Sinorhizobium meliloti</i> 1021 [B]	uni Q92TB5		
189		<i>Rhodobacter sphaeroides</i> 2.4.1 [B]			gi 22956528
190		<i>Silicibacter pomeroyi</i> DSS-3 [B]			uni Q5LPT8
191		<i>Silicibacter</i> sp. TM1040 [B]			gi 69298188
192		<i>Gluconobacter oxydans</i> 621H [B]			uni Q5FUV0
193		<i>Magnetospirillum magnetotacticum</i> [B]			gi 23014038
194		<i>Rhodospirillum rubrum</i> [B]			gi 48764632
195		<i>Anaplasma marginale</i> str. St. Maries [B]			uni Q5PBF7
196		<i>Ehrlichia canis</i> str. Jake [B]			gi 73666810
197		<i>Ehrlichia ruminantium</i> str. Gardel [B]			uni Q5FFD0
198		<i>Ehrlichia ruminantium</i> str. Welgevonden [B]			uni Q5HBZ7
199		<i>Candidatus Pelagibacter ubique</i> HTCC1062 [B]			uni Q4FM76
200		<i>Novosphingobium aromaticivorans</i> [B]			gi 23107278
201		<i>Zymomonas mobilis</i> subsp. <i>mobilis</i> ZM4 [B]			uni Q5NLB9
202	Proteobacteria - beta	<i>Bordetella avium</i> [B]			fig 521.1.psg.12
203		<i>Bordetella bronchiseptica</i> RB50 [B]			52 uni O88005

204		<i>Bordetella parapertussis</i> 12822 [B]		4191
205		<i>Bordetella pertussis</i> Tohama I [B]		uniQ45338
206		<i>Burkholderia cenocepacia</i> J2315 [B]		3767
207		<i>Burkholderia cepacia</i> R18194 [B]		gi46315372
208		<i>Burkholderia cepacia</i> R1808 [B]		uniQ4BG89
209		<i>Burkholderia fungorum</i> [B]		gi22987459
210		<i>Burkholderia mallei</i> ATCC 23344 [B]		uniQ62MZ8
211		<i>Burkholderia xenovorans</i> LB400 [B]		gi48786399
212		<i>Burkholderia pseudomallei</i> K96243 [B]		uniQ63XZ2
213		<i>Ralstonia metallidurans</i> CH34 [B]		gi22978823
214		<i>Ralstonia eutropha</i> JMP134 [B]		uniQ477E8
215		<i>Ralstonia solanacearum</i> GMI1000 [B]		uniQ8Y2M4
216		<i>Polaromonas</i> sp. JS666 [B]		uniQ4B0D5
217		<i>Thiobacillus denitrificans</i> ATCC 25259 [B]		gi52006082
218		<i>Methylobacillus flagellatus</i> KT [B]		gi46120804
219		<i>Chromobacterium violaceum</i> ATCC 12472 [B]		uniQ7P0S9
220		<i>Neisseria gonorrhoeae</i> FA 1090 [B]		uniQ5F5C8
221		<i>Neisseria lactamica</i> ST-640 [B]		fig486.1.pdg.406
222		<i>Neisseria meningitidis</i> FAM18 [B]		fig487.2.pdg.1365
223		<i>Neisseria meningitidis</i> MC58 [B]		uniQ9JXF1
224		<i>Neisseria meningitidis</i> Z2491 [B]		uniQ9JW17
225		<i>Nitrosomonas europaea</i> ATCC 19718 [B]		uniQ82S93
226		<i>Azoarcus</i> sp. EbN1 [B]		uniQ5P2L1
227		<i>Dechloromonas aromatica</i> RCB [B]		uniQ479J6
228	Proteobacteria - delta	<i>Bdellovibrio bacteriovorus</i> HD100 [B]		uniQ6MH10
229		<i>Desulfotalea psychrophila</i> L5v54 [B]		uniQ6AKW8
230		<i>Desulfovibrio desulfuricans</i> G20 [B]		gi23474277
231		<i>Desulfovibrio vulgaris</i> subsp. <i>vulgaris</i> str. Hildenborough [B]		uniQ72F93
232		<i>Geobacter metallireducens</i> GS-15 [B]		gi68004582
233		<i>Geobacter sulfurreducens</i> PCA [B]		uniQ74BU2
234	Proteobacteria - epsilon	<i>Campylobacter coli</i> RM2228 [B]		uniQ4HFS9
235		<i>Campylobacter jejuni</i> RM1221 [B]		uniQ5HW73
236		<i>Campylobacter jejuni</i> subsp. <i>jejuni</i> NCTC 11168 [B]		uniQ9PIA9
237		<i>Campylobacter lari</i> RM2100 [B]		uniQ4HIV0
238		<i>Campylobacter upsaliensis</i> RM3195 [B]		uniQ4HNC4
239		<i>Helicobacter hepaticus</i> ATCC 51449 [B]		uniQ7VJU3
240		<i>Helicobacter mustelae</i> 43772 [B]		fig217.1.pdg.527
241		<i>Helicobacter pylori</i> 26695 [B]		uniQ25533
242		<i>Helicobacter pylori</i> J99 [B]		uniQ9ZKY6
243		<i>Wolinella succinogenes</i> DSM 1740 [B]		uniQ7MS38
244	Proteobacteria - gamma	<i>Microbulbifer degradans</i> 2-40 [B]		gi48860574
245		<i>Colwellia psychrerythraea</i> 34H [B]		uniQ47UY4
246		<i>Idiomarina loihiensis</i> L2TR [B]		uniQ5QY78
247		<i>Shewanella amazonensis</i> SB2B [B]	gi68548554	
248		<i>Shewanella baltica</i> OS155 [B]	gi68544815	
249		<i>Shewanella denitrificans</i> OS-217 [B]	gi69942914	
250		<i>Shewanella frigidimarina</i> NCIMB 400 [B]	gi69954189	
251		<i>Shewanella oneidensis</i> MR-1 [B]	uniQ8EK83	
252		<i>Shewanella putrefaciens</i> CN-32 [B]	826	
253		<i>Shewanella</i> sp. PV-4 [B]	532	
254		<i>Nitrosococcus oceani</i> ATCC 19707 [B]		gi77163727
255		<i>Escherichia coli</i> 042 [B]	259	
256		<i>Escherichia coli</i> CFT073 [B]	gi26250747	
257		<i>Escherichia coli</i> E2348/69 [B]	820	

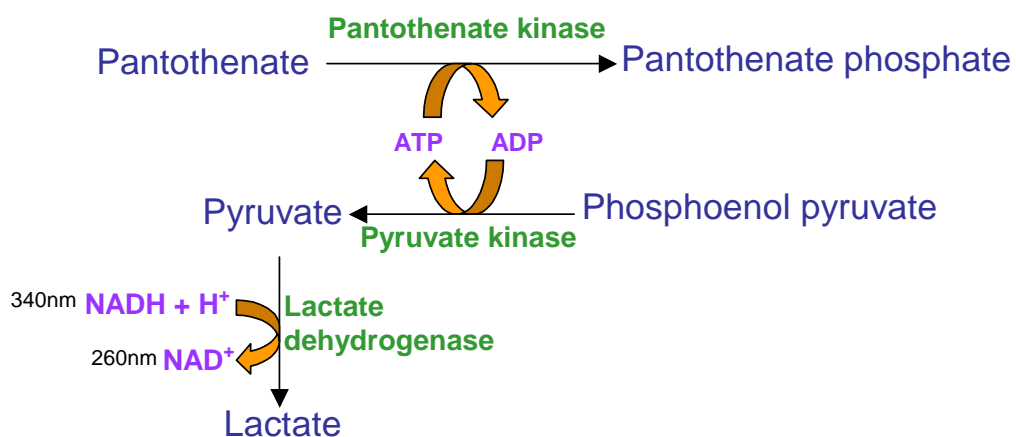
258	<i>Escherichia coli</i> K12 [B]	uni P0A613		
259	<i>Escherichia coli</i> O157:H7 [B]	uni P0A613		
260	<i>Escherichia coli</i> O157:H7 EDL933 [B]	gil15804568		
261	<i>Escherichia coli</i> DH10B [B]	uni P0A613		
262	<i>Klebsiella pneumoniae</i> MGH78578 [B]	fig 573.2.peg.283		
263	<i>Erwinia carotovora</i> subsp. <i>atroseptica</i> SCRI1043 [B]	uni Q6DAN8		
264	<i>Photorhabdus asymbiotica</i> subsp. <i>asymbiotica</i> [B]	69		
265	<i>Photorhabdus luminescens</i> subsp. <i>laumondii</i> TTO1 [B]	uni Q7MYE7		
266	<i>Proteus mirabilis</i> HI4320 [B]	fig 584.1.peg.238		
267	<i>Choleraesuis</i> str. SC-B67 [B]	uni Q57H79		
268	<i>Salmonella enterica</i> subsp. <i>enterica</i> serovar <i>Gallinarum</i> [B]	fig 594.1.peg.3939		
269	str. ATCC 9150 [B]	uni Q5PK80		
270	<i>Salmonella enterica</i> subsp. <i>enterica</i> serovar <i>Typhi</i> Ty2 [B]	uni Q5PK80		
271	<i>Salmonella enterica</i> subsp. <i>enterica</i> serovar <i>Typhi</i> str. CT18 [B]	uni Q5PK80		
272	<i>Salmonella dublin</i> [B]	uni Q57H79		
273	<i>Salmonella paratyphi</i> [B]	20		
274	<i>Salmonella typhimurium</i> LT2 [B]	uni Q9L9K3		
275	<i>Serratia marcescens</i> Db11 [B]	fig 615.1.peg.3484		
276	<i>Shigella dysenteriae</i> M131649 [B]	97		
277	<i>Shigella flexneri</i> 2a str. 2457T [B]	gil30064739		
278	<i>Shigella flexneri</i> 2a str. 301 [B]	gil24115264		
279	<i>Shigella sonnei</i> 53G [B]	807		
280	<i>Shigella sonnei</i> Ss046 [B]	uni P0A613		
281	<i>brevipalpis</i> [B]	uni Q8D241		
282	<i>Yersinia enterocolitica</i> 8081 [B]	fig 630.2.peg.271		
283	<i>Yersinia pestis</i> CO92 [B]	uni Q66FR1		
284	<i>Yersinia pestis</i> KIM [B]	uni Q66FR1		
285	<i>Yersinia pestis</i> biovar <i>Medievalis</i> str. 91001 [B]	uni Q66FR1		
286	<i>Yersinia pseudotuberculosis</i> IP 32953 [B]	uni Q66FR1		
287	<i>Yersinia pseudotuberculosis</i> [B]	fig 633.1.peg.1701		
288	<i>Candidatus Blochmannia pennsylvanicus</i> str. BPEN [B]	uni Q493L1		
289	<i>Coxiella burnetii</i> RSA 493 [B]	uni Q83EV9		
290	<i>Legionella pneumophila</i> str. <i>Lens</i> [B]			uni Q5WXZ6
291	<i>Legionella pneumophila</i> str. <i>Paris</i> [B]			uni Q5X6J2
292	<i>Legionella pneumophila</i> subsp. <i>pneumophila</i> str. <i>Philadelphia 1</i> [B]			uni Q5X6J2
293	<i>Methylococcus capsulatus</i> str. <i>Bath</i> [B]			uni Q607I0
294	<i>Actinobacillus actinomycetemcomitans</i> HK1651 [B]	fig 714.2.peg.470		
295	<i>Actinobacillus pleuropneumoniae</i> serovar 1 str. 4074 [B]	gil32034918		
296	<i>Haemophilus ducreyi</i> 35000HP [B]	uni Q7VPK9		
297	<i>Haemophilus influenzae</i> 86-028NP [B]	uni Q4QMW5		
298	<i>Haemophilus influenzae</i> R2866 [B]	uni Q4QMW5		
299	<i>Haemophilus influenzae</i> Rd KW20 [B]	uni P44793		
300	<i>Haemophilus somnus</i> 2336 [B]	gil46156238		
301	<i>Mannheimia haemolytica</i> [B]	2		
302	<i>Mannheimia succiniciproducens</i> MBEL55E [B]	uni Q65QG5		
303	<i>Pasteurella multocida</i> subsp. <i>multocida</i> str. Pm70 [B]	uni P57967		
304	<i>Acinetobacter</i> sp. ADP1 [B]			uni Q6FDW7
305	<i>Psychrobacter</i> sp. 273-4 [B]			uni Q4FUX4

306		<i>Psychrobacter</i> sp. 273-4 [B]			uni Q4FUX4
307		<i>Pseudomonas aeruginosa</i> PAO1 [B]			uni Q9HWC1
308		<i>Pseudomonas aeruginosa</i> UCBPP-PA14 [B]			gi 46164770
309		<i>Pseudomonas fluorescens</i> PfO-1 [B]			gi 77461316
310		<i>Pseudomonas fluorescens</i> SBW25 [B]			_6895
311		<i>Pseudomonas putida</i> KT2440 [B]			uni Q88QQ0
312		<i>Pseudomonas syringae</i> pv. phaseolicola 1448A [B]			uni Q48D21
313		<i>Pseudomonas syringae</i> pv. syringae B728a [B]			uni Q4ZMM9
314		<i>Pseudomonas syringae</i> pv. tomato str. DC3000 [B]			uni Q889Y6
315		<i>Francisella tularensis</i> (Livermore) [B]			uni Q5NF55
316		<i>Francisella tularensis</i> subsp. tularensis Schu 4 [B]			uni Q5NF55
317		<i>Vibrio cholerae</i> O1 biovar eltor str. N16961 [B]	uni Q9KV38		
318		<i>Vibrio fischeri</i> ES114 [B]	uni Q5E227		
319		<i>Vibrio parahaemolyticus</i> RIMD 2210633 [B]	849		
320		<i>Vibrio vulnificus</i> CMCP6 [B]	uni Q8DD30		
321		<i>Vibrio vulnificus</i> YJ016 [B]	uni Q7MGQ9		
322		<i>Stenotrophomonas maltophilia</i> K279a [B]			_1559
323		<i>Xanthomonas axonopodis</i> pv. citri str. 306 [B]			uni Q8PFG5
324		<i>Xanthomonas campestris</i> pv. campestris str. 8004 [B]			uni Q4UPF9
325		<i>Xanthomonas campestris</i> pv. campestris ATCC 33913 [B]			uni Q4UPF9
326		<i>Xanthomonas campestris</i> pv. vesicatoria str. 85-10 [B]			gi 78049665
327		<i>Xanthomonas oryzae</i> pv. oryzae KACC10331 [B]			uni Q5H5T7
328		<i>Xylella fastidiosa</i> 9a5c [B]			uni Q9PCI4
329		<i>Xylella fastidiosa</i> Ann-1 [B]			uni Q9PCI4
330		<i>Xylella fastidiosa</i> Temecula1 [B]			uni Q87CJ5
331		<i>Xylella fastidiosa</i> Dixon [B]			gi 22993998
332	Unclassified	<i>Magnetococcus</i> sp. MC-1 [B]			gi 68245538
	proteobacteria				
333	Spirochaetes	<i>Fiocruz</i> LI-130 [B]			uni Q72NP0
334		<i>Leptospira interrogans</i> serovar Lai str. 56601 [B]			uni Q72NP0
335		<i>Borrelia burgdorferi</i> B31 [B]			uni Q51477
336		<i>Borrelia garinii</i> PBi [B]			uni Q660Z6
337		<i>Treponema pallidum</i> subsp. pallidum str. Nichols [B]			uni Q83446
338	Thermotogae	<i>Thermotoga maritima</i> MSB8 [B]			uni Q9WZY5
339	EUKARYOTA				
340		<i>Plasmodium falciparum</i> 3D7 [E] - All possible isozyme forms		uni Q8ILP4 uni Q8IL92	
341		<i>Neurospora crassa</i> [E]		uni Q7SEX7	
342		<i>Saccharomyces cerevisiae</i> (baker's yeast) [E]		uni Q04430	
343		<i>Schizosaccharomyces pombe</i> [E]		uni Q74962	
344		<i>Anopheles gambiae</i> str. PEST [E]		gi 31240767	
345		<i>Mus musculus</i> (House mouse) [E] - All possible isozyme forms		uni Q8BQG7 fig 10090.3.pe g.4975 uni Q7M753 uni Q80YV4 uni Q8BQE9 uni Q5SXA7	
346		<i>Homo sapiens</i> (Human) [E] - All possible isozyme forms		uni Q5TA84 uni Q7RTX4 uni Q8TE04 gi 55957270 gi 55957269	

347	<i>Arabidopsis thaliana</i> [E] - All possible isozyme forms		uni Q8L5Y9 uni O80765 uni Q949P3 gi 42569099 uni Q8W4Q1	
-----	--	--	---	--

APPENDIX B

Continuous Coupled Kinase Assay



The assay couples the production of ADP to the reactions catalyzed by pyruvate kinase and lactate dehydrogenase (Huo and Viola, 1996; Singh et al., 2004; Strauss and Begley, 2002). The consumption of NADH was monitored by changes in absorption at 340 nm.

REFERENCES

- Aceti, D.J., and Ferry, J.G. (1988). Purification and characterization of acetate kinase from acetate-grown *Methanosarcina thermophila*. Evidence for regulation of synthesis. *J Biol Chem* 263, 15444-15448.
- Aleshin, A.E., Kirby, C., Liu, X., Bourenkov, G.P., Bartunik, H.D., Fromm, H.J., and Honzatko, R.B. (2000). Crystal structures of mutant monomeric hexokinase I reveal multiple ADP binding sites and conformational changes relevant to allosteric regulation. *J Mol Biol* 296, 1001-1015.
- Andreeva, A., Howorth, D., Brenner, S.E., Hubbard, T.J., Chothia, C., and Murzin, A.G. (2004). SCOP database in 2004: refinements integrate structure and sequence family data. *Nucleic Acids Res* 32, D226-229.
- Arora, K.K., Filburn, C.R., and Pedersen, P.L. (1991). Glucose phosphorylation. Site-directed mutations which impair the catalytic function of hexokinase. *J Biol Chem* 266, 5359-5362.
- Baddiley, J., Thain, E.M., Novelli, G.D., and Lipmann, F. (1953). Structure of coenzyme A. *Nature* 171, 76.
- Baldwin, J., Michnoff, C.H., Malmquist, N.A., White, J., Roth, M.G., Rathod, P.K., and Phillips, M.A. (2005). High-throughput screening for potent and selective inhibitors of *Plasmodium falciparum* dihydroorotate dehydrogenase. *J Biol Chem* 280, 21847-21853.
- Becker, D., Selbach, M., Rollenhagen, C., Ballmaier, M., Meyer, T.F., Mann, M., and Bumann, D. (2006). Robust *Salmonella* metabolism limits possibilities for new antimicrobials. *Nature* 440, 303-307.
- Begley, T.P., Kinsland, C., and Strauss, E. (2001). The biosynthesis of coenzyme A in bacteria. *Vitam Horm* 61, 157-171.

- Blattler, W.A., and Knowles, J.R. (1979). Stereochemical course of phosphokinases. The use of adenosine [γ -(S)-16O,17O,18O] triphosphate and the mechanistic consequences for the reactions catalyzed by glycerol kinase, hexokinase, pyruvate kinase, and acetate kinase. *Biochemistry* 18, 3927-3933.
- Bock, A., Bantscheff, M., Perraud, A.L., Rippe, K., Weiss, V., Glocker, M.O., and Gross, R. (2001). Rational design and molecular characterization of a chimaeric response regulator protein. *J Mol Biol* 310, 283-290.
- Bork, P., Sander, C., and Valencia, A. (1992). An ATPase domain common to prokaryotic cell cycle proteins, sugar kinases, actin, and hsp70 heat shock proteins. *Proc Natl Acad Sci U S A* 89, 7290-7294.
- Bradford, M.M. (1976). A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal Biochem* 72, 248-254.
- Brand, L.A., and Strauss, E. (2005). Characterization of a new pantothenate kinase isoform from *Helicobacter pylori*. *J Biol Chem* 280, 20185-20188.
- Bremer, J., Wojtczak, A., and Skrede, S. (1972). The leakage and destruction of CoA in isolated mitochondria. *Eur J Biochem* 25, 190-197.
- 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., *et al.* (1998). Crystallography & NMR system: A new software suite for macromolecular structure determination. *Acta Crystallogr D Biol Crystallogr* 54, 905-921.
- Buss, K.A., Cooper, D.R., Ingram-Smith, C., Ferry, J.G., Sanders, D.A., and Hasson, M.S. (2001). Urkinase: structure of acetate kinase, a member of the ASKHA superfamily of phosphotransferases. *J Bacteriol* 183, 680-686.

- Calder, R.B., Williams, R.S., Ramaswamy, G., Rock, C.O., Campbell, E., Unkles, S.E., Kinghorn, J.R., and Jackowski, S. (1999). Cloning and characterization of a eukaryotic pantothenate kinase gene (panK) from *Aspergillus nidulans*. *J Biol Chem* 274, 2014-2020.
- Cheek, S., Ginalski, K., Zhang, H., and Grishin, N.V. (2005). A comprehensive update of the sequence and structure classification of kinases. *BMC Struct Biol* 5, 6.
- Cheek, S., Zhang, H., and Grishin, N.V. (2002). Sequence and structure classification of kinases. *J Mol Biol* 320, 855-881.
- Choudhry, A.E., Mandichak, T.L., Broskey, J.P., Egolf, R.W., Kinsland, C., Begley, T.P., Seefeld, M.A., Ku, T.W., Brown, J.R., Zalacain, M., *et al.* (2003). Inhibitors of pantothenate kinase: novel antibiotics for staphylococcal infections. *Antimicrob Agents Chemother* 47, 2051-2055.
- Collaborative Computational Project, n. (1994) The CCP4 suite: programs for protein crystallography, *Acta Crystallogr D Biol Crystallogr* 50, 760-3.
- Cronan, J.E., Jr. (1980). Beta-alanine synthesis in *Escherichia coli*. *J Bacteriol* 141, 1291-1297.
- Cronan, J.E., Jr., Littel, K.J., and Jackowski, S. (1982). Genetic and biochemical analyses of pantothenate biosynthesis in *Escherichia coli* and *Salmonella typhimurium*. *J Bacteriol* 149, 916-922.
- Daugherty, M., Polanuyer, B., Farrell, M., Scholle, M., Lykidis, A., de Crecy-Lagard, V., and Osterman, A. (2002). Complete reconstitution of the human coenzyme A biosynthetic pathway via comparative genomics. *J Biol Chem* 277, 21431-21439.
- DeLano, W.L. (2002). The PyMOL Molecular Graphics System. www.pymol.org.
- delCardayre, S.B., Stock, K.P., Newton, G.L., Fahey, R.C., and Davies, J.E. (1998). Coenzyme A disulfide reductase, the primary low molecular

- weight disulfide reductase from *Staphylococcus aureus*. Purification and characterization of the native enzyme. *J Biol Chem* 273, 5744-5751.
- DeShazer, D., Wood, G.E., and Friedman, R.L. (1995). Identification of a *Bordetella pertussis* regulatory factor required for transcription of the pertussis toxin operon in *Escherichia coli*. *J Bacteriol* 177, 3801-3807.
- Doublet, S. (1997). Preparation of selenomethionyl proteins for phase determination. *Methods Enzymol* 276, 523-530.
- Dunn, S.D., and Snell, E.E. (1979). Isolation of temperature-sensitive pantothenate kinase mutants of *Salmonella typhimurium* and mapping of the *coaA* gene. *J Bacteriol* 140, 805-808.
- Emsley, P., and Cowtan, K. (2004). Coot: model-building tools for molecular graphics. *Acta Crystallogr D Biol Crystallogr* 60, 2126-2132.
- Fenstermacher, D.K., and Rose, R.C. (1986). Absorption of pantothenic acid in rat and chick intestine. *Am J Physiol* 250, G155-160.
- Flaherty, K.M., DeLuca-Flaherty, C., and McKay, D.B. (1990). Three-dimensional structure of the ATPase fragment of a 70K heat-shock cognate protein. *Nature* 346, 623-628.
- Garrett, R.H., and Grisham, C. M. (1999). *Biochemistry*, Second edition ed. edn.
- Geerlof, A., Lewendon, A., and Shaw, W.V. (1999). Purification and characterization of phosphopantetheine adenylyltransferase from *Escherichia coli*. *J Biol Chem* 274, 27105-27111.
- Genschel, U. (2004). Coenzyme A biosynthesis: reconstruction of the pathway in archaea and an evolutionary scenario based on comparative genomics. *Mol Biol Evol* 21, 1242-1251.
- Gerdes, S.Y., Scholle, M.D., D'Souza, M., Bernal, A., Baev, M.V., Farrell, M., Kurnasov, O.V., Daugherty, M.D., Mseeh, F., Polanuyer, B.M., *et al.* (2002). From genetic footprinting to antimicrobial drug targets: examples in cofactor biosynthetic pathways. *J Bacteriol* 184, 4555-4572.

- Grassl, S.M. (1992). Human placental brush-border membrane Na(+)-pantothenate cotransport. *J Biol Chem* 267, 22902-22906.
- Grzybowski, B.A., Ishchenko, A.V., Shimada, J., and Shakhnovich, E.I. (2002). From knowledge-based potentials to combinatorial lead design in silico. *Accounts of chemical research* 35, 261-269.
- Hirabayashi, Y., Kanamori, A., Nomura, K.H., and Nomura, K. (2004). The acetyl-CoA transporter family SLC33. *Pflugers Arch* 447, 760-762.
- Holm, L., and Sander, C. (1995). Dali: a network tool for protein structure comparison. *Trends Biochem Sci* 20, 478-480.
- Holmes, K.C., Popp, D., Gebhard, W., and Kabsch, W. (1990). Atomic model of the actin filament. *Nature* 347, 44-49.
- Hong, B., Wang, L., Tempel, W., Loppnau, P., Allali-Hassani, A., Arrowsmith, C.H., Edwards, A.M., Sundstrom, M., Weigelt, J., Bochkarev, A., Park, H. (2007). Crystal structure of human pantothenate kinase alpha: the catalytic core domain in complex with feedback regulator. To be Published.
- Hong, B.S., Yun, M.K., Zhang, Y.M., Chohnan, S., Rock, C.O., White, S.W., Jackowski, S., Park, H.W., and Leonardi, R. (2006). Prokaryotic type II and type III pantothenate kinases: The same monomer fold creates dimers with distinct catalytic properties. *Structure* 14, 1251-1261.
- Huo, X., and Viola, R.E. (1996). Functional group characterization of homoserine kinase from *Escherichia coli*. *Arch Biochem Biophys* 330, 373-379.
- Hurley, J.H. (1996). The sugar kinase/heat shock protein 70/actin superfamily: implications of conserved structure for mechanism. *Annu Rev Biophys Biomol Struct* 25, 137-162.
- Hurley, J.H., Faber, H.R., Worthylake, D., Meadow, N.D., Roseman, S., Pettigrew, D.W., and Remington, S.J. (1993). Structure of the regulatory

- complex of *Escherichia coli* III Glc with glycerol kinase. *Science* 259, 673-677.
- Ito, S., Fushinobu, S., Jeong, J.J., Yoshioka, I., Koga, S., Shoun, H., and Wakagi, T. (2003). Crystal structure of an ADP-dependent glucokinase from *Pyrococcus furiosus*: implications for a sugar-induced conformational change in ADP-dependent kinase. *J Mol Biol* 331, 871-883.
- Ivey, R.A., Zhang, Y.M., Virga, K.G., Hevener, K., Lee, R.E., Rock, C.O., Jackowski, S., and Park, H.W. (2004). The structure of the pantothenate kinase-ADP-pantothenate ternary complex reveals the relationship between the binding sites for substrate, allosteric regulator, and antimetabolites. *J Biol Chem* 279, 35622-35629.
- Izard, T. (2003). A novel adenylate binding site confers phosphopantetheine adenylyltransferase interactions with coenzyme A. *J Bacteriol* 185, 4074-4080.
- Jackowski, S. (1996). Biosynthesis of pantothenic acid and coenzyme A. *Escherichia coli and Salmonella: Cellular and Molecular Biology*, edited by F C NEIDHARDT, R CURTISS III, J INGRAHAM, E LIN, K LOW, B MAGASANIK, W REZNIKOFF, M RILEY, M SCHAECHTER and H UMBARGER, 7.
- Jackowski, S., and Alix, J.H. (1990). Cloning, sequence, and expression of the pantothenate permease (panF) gene of *Escherichia coli*. *J Bacteriol* 172, 3842-3848.
- Jackowski, S., and Rock, C.O. (1981). Regulation of coenzyme A biosynthesis. *J Bacteriol* 148, 926-932.
- Jackowski, S., and Rock, C.O. (1984). Metabolism of 4'-phosphopantetheine in *Escherichia coli*. *J Bacteriol* 158, 115-120.
- Jackowski, S., and Rock, C.O. (1986). Consequences of reduced intracellular coenzyme A content in *Escherichia coli*. *J Bacteriol* 166, 866-871.

- Jones, C.E., Brook, J.M., 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 (Pt 2), 110-119.
- Kabsch, W., Mannherz, H.G., Suck, D., Pai, E.F., and Holmes, K.C. (1990). Atomic structure of the actin:DNase I complex. *Nature* 347, 37-44.
- Kapust, R.B., and Waugh, D.S. (1999). *Escherichia coli* maltose-binding protein is uncommonly effective at promoting the solubility of polypeptides to which it is fused. *Protein Sci* 8, 1668-1674.
- Kitchen, D.B., Decornez, H., Furr, J.R., and Bajorath, J. (2004). Docking and scoring in virtual screening for drug discovery: methods and applications. *Nature reviews* 3, 935-949.
- Kotzbauer, P.T., Truax, A.C., Trojanowski, J.Q., and Lee, V.M. (2005). Altered neuronal mitochondrial coenzyme A synthesis in neurodegeneration with brain iron accumulation caused by abnormal processing, stability, and catalytic activity of mutant pantothenate kinase 2. *J Neurosci* 25, 689-698.
- Kupke, T., Uebele, M., Schmid, D., Jung, G., Blaesse, M., and Steinbacher, S. (2000). Molecular characterization of lantibiotic-synthesizing enzyme EpiD reveals a function for bacterial Dfp proteins in coenzyme A biosynthesis. *J Biol Chem* 275, 31838-31846.
- Lambalot, R.H., and Walsh, C.T. (1995). Cloning, overproduction, and characterization of the *Escherichia coli* holo-acyl carrier protein synthase. *J Biol Chem* 270, 24658-24661.

- 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. Cryst.* *26*, 283-291.
- Leonardi, R., Chohnan, S., Zhang, Y.M., Virga, K.G., Lee, R.E., Rock, C.O., and
Jackowski, S. (2005a). A pantothenate kinase from *Staphylococcus aureus*
refractory to feedback regulation by coenzyme A. *J Biol Chem* *280*, 3314-
3322.
- Leonardi, R., Zhang, Y.M., Rock, C.O., and Jackowski, S. (2005b). Coenzyme A:
back in action. *Prog Lipid Res* *44*, 125-153.
- Leonardi, R., Rock, C.O., Jackowski, S., and Zhang, Y.M. (2007). Activation of
human mitochondrial pantothenate kinase 2 by palmitoylcarnitine. *Proc
Natl Acad Sci U S A* *104*, 1494-1499.
- Lipmann, F., Kaplan, N.O., Novelli, G.D., Tuttle, L.C., and Guirard, B.M. (1950).
Isolation of coenzyme A. *J Biol Chem* *186*, 235-243.
- Locher, K.P., Hans, M., Yeh, A.P., Schmid, B., Buckel, W., and Rees, D.C.
(2001). Crystal structure of the *Acidaminococcus fermentans* 2-
hydroxyglutaryl-CoA dehydratase component A. *J Mol Biol* *307*, 297-308.
- Luba, J., Charrier, V., and Claiborne, A. (1999). Coenzyme A-disulfide reductase
from *Staphylococcus aureus*: evidence for asymmetric behavior on
interaction with pyridine nucleotides. *Biochemistry* *38*, 2725-2737.
- Lunin, V.V., Li, Y., Schrag, J.D., Iannuzzi, P., Cygler, M., and Matte, A. (2004).
Crystal structures of *Escherichia coli* ATP-dependent glucokinase and its
complex with glucose. *J Bacteriol* *186*, 6915-6927.
- Maas, W.K. (1952). Pantothenate studies. III. Description of the extracted
pantothenate-synthesizing enzyme of *Escherichia coli*. *J Biol Chem* *198*,
23-32.

- Maas, W.K., and Davis, B.D. (1950). Pantothenate studies. I. Interference by D-serine and L-aspartic acid with pantothenate synthesis in *Escherichia coli*. *J Bacteriol* *60*, 733-745.
- Martin, D.P., Bibart, R.T. and Drueckhammer, D.G. (1994). Synthesis of novel analogs of acetyl coenzyme A: mimics of enzyme reaction intermediates. *J Amer Chem Soc* *116*, 4660–4668.
- Matarin, M.M., Singleton, A.B., and Houlden, H. (2006). PANK2 gene analysis confirms genetic heterogeneity in neurodegeneration with brain iron accumulation (NBIA) but mutations are rare in other types of adult neurodegenerative disease. *Neurosci Lett* *407*, 162-165.
- Merkel, W.K., and Nichols, B.P. (1996). Characterization and sequence of the *Escherichia coli* panBCD gene cluster. *FEMS Microbiol Lett* *143*, 247-252.
- Mishra, P.K., and Drueckhammer, D.G. (2000). Coenzyme A Analogues and Derivatives: Synthesis and Applications as Mechanistic Probes of Coenzyme A Ester-Utilizing Enzymes. *Chem Rev* *100*, 3283-3310.
- Mishra, P., Park, P.K., and Drueckhammer, D.G. (2001). Identification of yacE (coaE) as the structural gene for dephosphocoenzyme A kinase in *Escherichia coli* K-12. *J Bacteriol* *183*, 2774-2778.
- Miyatake, K., Nakano, Y., and Kitaoka, S. (1978). Enzymological properties of pantothenate synthetase from *Escherichia coli* B. *J Nutr Sci Vitaminol (Tokyo)* *24*, 243-253.
- Murray, R.K., D.K.G., Rodwell, V.W. (2006). *Harper's Illustrated Biochemistry*, 27th Edition (The McGraw-Hill Companies, Inc).
- Murshudov, G.N., Vagin, A.A., and Dodson, E.J. (1997). Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr D Biol Crystallogr* *53*, 240-255.

- 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.
- Nakamura, H., and Tamura, Z. (1973). Pantothenate uptake in *Escherichia coli* K-12. *J Nutr Sci Vitaminol (Tokyo)* 19, 389-400.
- Nicely, N.I., Parsonage, D., Paige, C., Newton, G.L., Fahey, R.C., Leonardi, R., Jackowski, S., Mallett, T.C., and Claiborne, A. (2007). Structure of the type III pantothenate kinase from *Bacillus anthracis* at 2.0 Å resolution: implications for coenzyme A-dependent redox biology. *Biochemistry* 46, 3234-3245.
- Osterman, A., and Overbeek, R. (2003). Missing genes in metabolic pathways: a comparative genomics approach. *Curr Opin Chem Biol* 7, 238-251.
- Ottenhof, H.H., Ashurst, J.L., Whitney, H.M., Saldanha, S.A., Schmitzberger, F., Gweon, H.S., Blundell, T.L., Abell, C., and Smith, A.G. (2004). Organisation of the pantothenate (vitamin B5) biosynthesis pathway in higher plants. *Plant J* 37, 61-72.
- Otwinowski, Z., and Minor, W. (1997). Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol* 276, 307-326.
- Overbeek, R., Begley, T., Butler, R.M., Choudhuri, J.V., Chuang, H.Y., Cohoon, M., de Crecy-Lagard, V., Diaz, N., Disz, T., Edwards, R., *et al.* (2005). The subsystems approach to genome annotation and its use in the project to annotate 1000 genomes. *Nucleic Acids Res* 33, 5691-5702.
- Prasad, P.D., Ramamoorthy, S., Leibach, F.H., and Ganapathy, V. (1997). Characterization of a sodium-dependent vitamin transporter mediating the uptake of pantothenate, biotin and lipoate in human placental choriocarcinoma cells. *Placenta* 18, 527-533.
- Prasad, P.D., Wang, H., Huang, W., Fei, Y.J., Leibach, F.H., Devoe, L.D., and Ganapathy, V. (1999). Molecular and functional characterization of the

- intestinal Na⁺-dependent multivitamin transporter. *Arch Biochem Biophys* 366, 95-106.
- Prasad, P.D., Wang, H., Kekuda, R., Fujita, T., Fei, Y.J., Devoe, L.D., Leibach, F.H., and Ganapathy, V. (1998). Cloning and functional expression of a cDNA encoding a mammalian sodium-dependent vitamin transporter mediating the uptake of pantothenate, biotin, and lipoate. *J Biol Chem* 273, 7501-7506.
- Raman, S.B., and B. Rathinasabapathi (2004). Pantothenate synthesis in plants. *Plant Sci* 167, 7.
- Rathinasabapathi, B., and Raman, S.B. (2005). Exogenous supply of pantoyl lactone to excised leaves increases their pantothenate levels. *Annals of botany* 95, 1033-1037.
- Rawlings, M., and Cronan, J.E., Jr. (1992). The gene encoding Escherichia coli acyl carrier protein lies within a cluster of fatty acid biosynthetic genes. *J Biol Chem* 267, 5751-5754.
- Reibel, D.K., Uboh, C.E., and Kent, R.L. (1983). Altered coenzyme A and carnitine metabolism in pressure-overload hypertrophied hearts. *Am J Physiol* 244, H839-843.
- Reibel, D.K., Wyse, B.W., Berkich, D.A., Palko, W.M., and Neely, J.R. (1981). Effects of diabetes and fasting on pantothenic acid metabolism in rats. *Am J Physiol* 240, E597-601.
- Rock, C.O., Calder, R.B., Karim, M.A., and Jackowski, S. (2000). Pantothenate kinase regulation of the intracellular concentration of coenzyme A. *J Biol Chem* 275, 1377-1383.
- Rock, C.O., Karim, M.A., Zhang, Y.M., and Jackowski, S. (2002). The murine pantothenate kinase (Pank1) gene encodes two differentially regulated pantothenate kinase isozymes. *Gene* 291, 35-43.

- Rock, C.O., Park, H.W., and Jackowski, S. (2003). Role of feedback regulation of pantothenate kinase (CoaA) in control of coenzyme A levels in *Escherichia coli*. *J Bacteriol* *185*, 3410-3415.
- Rosano, C., Sabini, E., Rizzi, M., Deriu, D., Murshudov, G., Bianchi, M., Serafini, G., Magnani, M., and Bolognesi, M. (1999). Binding of non-catalytic ATP to human hexokinase I highlights the structural components for enzyme-membrane association control. *Structure* *7*, 1427-1437.
- Saraste, M., Sibbald, P.R., and Wittinghofer, A. (1990). The P-loop--a common motif in ATP- and GTP-binding proteins. *Trends Biochem Sci* *15*, 430-434.
- Schneider, T.R., and Sheldrick, G.M. (2002). Substructure solution with SHELXD. *Acta Crystallogr D Biol Crystallogr* *58*, 1772-1779.
- Shibata, K., Gross, C.J., and Henderson, L.M. (1983). Hydrolysis and absorption of pantothenate and its coenzymes in the rat small intestine. *J Nutr* *113*, 2107-2115.
- Singh, S.K., Yang, K., Karthikeyan, S., Huynh, T., Zhang, X., Phillips, M.A., and Zhang, H. (2004). The thrH gene product of *Pseudomonas aeruginosa* is a dual activity enzyme with a novel phosphoserine:homoserine phosphotransferase activity. *J Biol Chem* *279*, 13166-13173.
- Song, W.J., and Jackowski, S. (1992). Cloning, sequencing, and expression of the pantothenate kinase (coaA) gene of *Escherichia coli*. *J Bacteriol* *174*, 6411-6417.
- Song, W.J., and Jackowski, S. (1994). Kinetics and regulation of pantothenate kinase from *Escherichia coli*. *J Biol Chem* *269*, 27051-27058.
- Steitz, T.A., Fletterick, R.J., Anderson, W.F., and Anderson, C.M. (1976). High resolution x-ray structure of yeast hexokinase, an allosteric protein exhibiting a non-symmetric arrangement of subunits. *J Mol Biol* *104*, 197-122.

- Strauss, E., and Begley, T.P. (2002). The antibiotic activity of N-pentylpantothenamide results from its conversion to ethyldethia-coenzyme A, a coenzyme A antimetabolite. *J Biol Chem* 277, 48205-48209.
- Strauss, E., Kinsland, C., Ge, Y., McLafferty, F.W., and Begley, T.P. (2001). Phosphopantothenoylcysteine synthetase from *Escherichia coli*. Identification and characterization of the last unidentified coenzyme A biosynthetic enzyme in bacteria. *J Biol Chem* 276, 13513-13516.
- Tahiliani, A.G., and Beinlich, C.J. (1991). Pantothenic acid in health and disease. *Vitam Horm* 46, 165-228.
- Tautz, L., and Mustelin, T. (2007). Strategies for developing protein tyrosine phosphatase inhibitors. *Methods (San Diego, Calif)* 42, 250-260.
- Terwilliger, T.C. (2000). Maximum-likelihood density modification. *Acta Crystallogr D Biol Crystallogr* 56, 965-972.
- Terwilliger, T.C. (2003). SOLVE and RESOLVE: automated structure solution and density modification. *Methods Enzymol* 374, 22-37.
- Vallari, D.S., and Jackowski, S. (1988). Biosynthesis and degradation both contribute to the regulation of coenzyme A content in *Escherichia coli*. *J Bacteriol* 170, 3961-3966.
- Vallari, D.S., Jackowski, S., and Rock, C.O. (1987). Regulation of pantothenate kinase by coenzyme A and its thioesters. *J Biol Chem* 262, 2468-2471.
- Vallari, D.S., and Rock, C.O. (1985). Pantothenate transport in *Escherichia coli*. *J Bacteriol* 162, 1156-1161.
- Vallari, D.S., and Rock, C.O. (1987). Isolation and characterization of temperature-sensitive pantothenate kinase (coaA) mutants of *Escherichia coli*. *J Bacteriol* 169, 5795-5800.
- van den Ent, F., and Lowe, J. (2000). Crystal structure of the cell division protein FtsA from *Thermotoga maritima*. *Embo J* 19, 5300-5307.

- Virga, K.G., Zhang, Y.M., Leonardi, R., Ivey, R.A., Hevener, K., Park, H.W., Jackowski, S., Rock, C.O., and Lee, R.E. (2006). Structure-activity relationships and enzyme inhibition of pantothenamide-type pantothenate kinase inhibitors. *Bioorg Med Chem* 14, 1007-1020.
- 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.
- Wang, H., Huang, W., Fei, Y.J., Xia, H., Yang-Feng, T.L., Leibach, F.H., Devoe, L.D., Ganapathy, V., and Prasad, P.D. (1999). Human placental Na⁺-dependent multivitamin transporter. Cloning, functional expression, gene structure, and chromosomal localization. *J Biol Chem* 274, 14875-14883.
- Wang, H.Y., Baxter, C.F., Jr., and Schulz, H. (1991). Regulation of fatty acid beta-oxidation in rat heart mitochondria. *Arch Biochem Biophys* 289, 274-280.
- Wharton, C.W., and Eisenthal, R. (1981). *Molecular enzymology* (Blackie and Son Limited, London.).
- Wilbanks, S.M., DeLuca-Flaherty, C., and McKay, D.B. (1994). Structural basis of the 70-kilodalton heat shock cognate protein ATP hydrolytic activity. I. Kinetic analyses of active site mutants. *J Biol Chem* 269, 12893-12898.
- Wilbanks, S.M., and McKay, D.B. (1998). Structural replacement of active site monovalent cations by the epsilon-amino group of lysine in the ATPase fragment of bovine Hsc70. *Biochemistry* 37, 7456-7462.
- Williamson, J.M., and Brown, G.M. (1979). Purification and properties of L-Aspartate-alpha-decarboxylase, an enzyme that catalyzes the formation of beta-alanine in *Escherichia coli*. *J Biol Chem* 254, 8074-8082.

- Williamson, J.R., and Corkey, B.E. (1979). Assay of citric acid cycle intermediates and related compounds--update with tissue metabolite levels and intracellular distribution. *Methods Enzymol* 55, 200-222.
- Wimmer, M.J., and Rose, I.A. (1978). Mechanisms of enzyme-catalyzed group transfer reactions. *Annu Rev Biochem* 47, 1031-1078.
- Wittwer, C.T., Beck, S., Peterson, M., Davidson, R., Wilson, D.E., and Hansen, R.G. (1990). Mild pantothenate deficiency in rats elevates serum triglyceride and free fatty acid levels. *J Nutr* 120, 719-725.
- Wood, G.E., and Friedman, R.L. (2000). The Bvg accessory factor (Baf) enhances pertussis toxin expression in *Escherichia coli* and is essential for *Bordetella pertussis* viability. *FEMS Microbiol Lett* 193, 25-30.
- Yang, K., Eyobo, Y., Brand, L.A., Martynowski, D., Tomchick, D., Strauss, E., and Zhang, H. (2006). Crystal structure of a type III pantothenate kinase: insight into the mechanism of an essential coenzyme A biosynthetic enzyme universally distributed in bacteria. *J Bacteriol* 188, 5532-5540.
- Ye, Y., Osterman, A., Overbeek, R., and Godzik, A. (2005). Automatic detection of subsystem/pathway variants in genome analysis. *Bioinformatics* 21 Suppl 1, i478-486.
- Yocum, R.R., and Patterson T.A. (2002). Microorganisms and assays for the identification of antibiotics acting on the pantothenate kinase encoded by the *coaX* gene., I. Omnigene Bioproducts, USA, ed. (USA).
- Yun, M., Park, C.G., Kim, J.Y., Rock, C.O., Jackowski, S., and Park, H.W. (2000). Structural basis for the feedback regulation of *Escherichia coli* pantothenate kinase by coenzyme A. *J Biol Chem* 275, 28093-28099.
- Zhang, Y.M., Frank, M.W., Virga, K.G., Lee, R.E., Rock, C.O., and Jackowski, S. (2004). Acyl carrier protein is a cellular target for the antibacterial action of the pantothenamide class of pantothenate antimetabolites. *J Biol Chem* 279, 50969-50975.

- Zhao, L., Allanson, N.M., Thomson, S.P., Maclean, J.K., Barker, J.J., Primrose, W.U., Tyler, P.D., and Lewendon, A. (2003). Inhibitors of phosphopantetheine adenylyltransferase. *Eur J Med Chem* 38, 345-349.
- Zhou, B., Westaway, S.K., Levinson, B., Johnson, M.A., Gitschier, J., and Hayflick, S.J. (2001). A novel pantothenate kinase gene (PANK2) is defective in Hallervorden-Spatz syndrome. *Nat Genet* 28, 345-349.
- Zhyvoloup, A., Nemazanyy, I., Babich, A., Panasyuk, G., Pobigailo, N., Vudmaska, M., Naidenov, V., Kukhareno, O., Palchevskii, S., Savinska, L., *et al.* (2002). Molecular cloning of CoA Synthase. The missing link in CoA biosynthesis. *J Biol Chem* 277, 22107-22110.

VITAE

Kun Yang was born in Shandong, China, on Feb 27, 1976, the son of Shuxin Yang and Yanjun Li. After completing his studies at Jinan No. 1 High School, Shandong, China in 1994, he joined to the department of Microbiology at Shandong University at Jinan, Shandong. After graduating from Shandong University, he conducted his graduate study for one year at the Graduate School of the Chinese Academy of Science, Beijing, China. In August 2003, he was accepted to the Division of Cellular and Molecular Biology at UT Southwestern Medical Center as a graduate student. He did his thesis work under the supervision of Dr. Hong Zhang in the department of Biochemistry. Kun and his wife Jing Wang were married in May 2001. Their son, Max Yang, was born in 2002.

Permanent Address: 2330 Tan Oak Drive
Dallas, TX 75212