

2-8-2011

Divergence of thioesterase function : human BFIT2, Escherichia coli EntH, and YDII

CHEN DANQI

Follow this and additional works at: https://digitalrepository.unm.edu/chem_etds

Recommended Citation

DANQI, CHEN. "Divergence of thioesterase function : human BFIT2, Escherichia coli EntH, and YDII." (2011).
https://digitalrepository.unm.edu/chem_etds/10

This Dissertation is brought to you for free and open access by the Electronic Theses and Dissertations at UNM Digital Repository. It has been accepted for inclusion in Chemistry ETDs by an authorized administrator of UNM Digital Repository. For more information, please contact disc@unm.edu.

DANQI CHEN

Candidate

Department of Chemistry and Chemical Biology

Department

This dissertation is approved, and it is acceptable in quality and form for publication:

Approved by the Dissertation Committee:



Debra Dunaway-Mariano, Chairperson



Patrick S. Mariano



Charles E. Melancon



Karen N. Allen

**DIVERGENCE OF THIOESTERASE FUNCTION: HUMAN BFIT2,
ESCHERICHIA COLI ENTH, AND YDII**

BY

DANQI CHEN

B.A., Applied Chemistry, Huaqiao University, China, 2000
M.S., Polymer Science and Physics, Sichuan University, China, 2005

DISSERTATION

Submitted in Partial Fulfillment of the
Requirements for the Degree of

Doctor of Philosophy

Chemistry

The University of New Mexico
Albuquerque, New Mexico

December, 2010

DEDICATION

To

My wife, Hong Zhao

My parents, Bojian Chen and Shaoying Rong

My sister and brother in law, Subi Chen and Shaotang Lin

My niece, Zixin Lin

**DIVERGENCE OF THIOESTERASE FUNCTION: HUMAN BFIT2,
ESCHERICHIA COLI ENTH, AND YDII**

BY

DANQI CHEN

ABSTRACT OF DISSERTATION

Submitted in Partial Fulfillment of the
Requirements for the Degree of
Doctor of Philosophy

Chemistry

The University of New Mexico
Albuquerque, New Mexico

December, 2010

**DIVERGENCE OF THIOESTERASE FUNCTION: HUMAN BFIT2,
ESCHERICHIA COLI ENTH, AND YDII**

by

DANQI CHEN

B.A., Applied Chemistry, Huaqiao University, China, 2000

M.S., Polymer Science and Physics, Sichuan University, China, 2005

Ph.D., Chemistry, University of New Mexico, USA, 2010

ABSTRACT

My doctoral research primarily focuses on two hotdog-fold thioesterases, EntH (also known as YbdB) from *E. coli*, and BFIT2 from *Homo sapiens*.

The *entH* (*ybdB*) gene is included in a large gene cluster that encodes the enzymes of the biosynthetic pathway leading to enterobactin. Building on the hypothesis that EntH might function in a “house-keeping” role by liberating misacylated EntB, two potential pathways to EntB misacylation were identified, one involving the phosphopantetheinyl transferase EntD and the other involving 2,3-DHB-AMP ligase EntE. EntH displays thioesterase activity towards a variety of acyl and aryl-holo EntB adducts. Lastly, It was shown that EntF acts on the 2,3-DHB-holo-EntB quickly, but not quickly on misacylated EntB adducts.

X-ray structures of liganded EntH revealed a solvated acyl-binding pocket which explains the broad substrate range of EntH. Likewise, the nucleotide moiety of the ligand CoA unit appears to sit on the enzyme surface. Only the pantetheine arm and the thioester C=O are engaged in bonding interactions with the enzyme.

Human thioesterase BFIT2 contains two tandem hotdog-fold thioesterase domains and a C-terminal steroidogenic acute regulatory protein related lipid transfer (START) domain. The expression of BFIT2 is induced during the thermogenesis transition of brown fat tissue. The expression of the recombinant BFIT2 in transfected HEK cells was confirmed by Western blot analysis. The recombinant BFIT2 contains a N-terminal His₆-tag and epitope, which was found to be susceptible to posttranslational removal. The recombinant N-terminal (minus residues 1-34) truncated mutant was found not to undergo posttranslational cleavage, thus suggesting that the N-terminal region is a signal sequence. A chimeric protein BFIT2 N(1-42)-GFP was shown by confocal microscopy to co-locate with the mitochondria. The BFTI2 precursor was shown to be taken up by freshly isolated HEK cell mitochondria and cleaved to the mature form. These results confirmed that the N-terminal region of BFIT2 functions as MTS.

During the thermogenesis transition of brown fat tissue, BFIT2 might function to restore the balance between free CoA and fatty acyl-CoA by hydrolyzing the long to medium chain fatty acyl-CoAs. Consistent with this hypothesis, BFIT2 was found to be much more active towards palmitoyl-CoA, myristoyl-CoA and lauroyl-CoA.

TABLE OF CONTENTS

LIST OF FIGURES	XII
LIST OF TABLES	XXV
CHAPTER 1. DIVERGENCE OF BIOLOGICAL FUNCTION WITHIN THE HOTDOG FOLD THIOESTERASE FAMILY.....	1
1.1 Overview of the Hot Dog Fold Protein Superfamily	1
1.2 Divergence of Hotdog Fold Thioesterases.....	7
1.2.1 Divergence in Cellular functions	8
1.2.1.1 4-Hydroxybenzoyl-CoA Thioesterase (4-HBT) in Dehalogenation of 4-Chlorobenzoate.....	8
1.2.1.2 E. coli PaaI in Phenylacetic Acid Catabolism Pathway.....	8
1.2.1.3 CalE7 in Calicheamicin Biosynthesis.....	10
1.2.1.4 FIK in Fluoroacetate Biosynthesis.....	12
1.2.1.5 Acot8, A HIV-1 Nef Binding Protein.....	14
1.2.1.6 Cytosolic Acetyl-CoA Hydrolase (CACH)	15
1.2.2 Divergence in Protein Folding of Hotdog Thioesterase	16
1.2.2.1 Tertiary Structure	16
1.2.2.2 Quaternary Structure.....	20
1.2.2.3 Catalytic Mechanisms.....	23
1.3 Conclusions.....	29
1.4 References.....	30

CHAPTER 2. IN VITRO KINETIC ANALYSIS OF SUBSTRATE SPECIFICITY IN ENTEROBACTIN BIOSYNTHETIC LOWER PATHWAY ENZYMES PROVIDES INSIGHT INTO THE BIOCHEMICAL FUNCTION OF ENTH34

2.1 Introduction.....	34
2.2 Experimental	38
2.2.1 Materials.	38
2.2.2 Growth monitoring of type BW2513, $\Delta entH$ /BW2513 and $\Delta entB$ /BW2513 strains under iron-rich or iron-depleting conditions.	38
2.2.3 Electrospray Ionization Mass Spectroscopy.....	39
2.2.4 <i>EntH</i> (<i>ybdB</i>) Cloning and Expression and EntH Purification.....	39
2.2.5 Preparation of Recombinant Apo-EntB, EntD, EntE and Apo-EntF.....	40
2.2.6 EntD Catalyzed Formation of Misacylated-holo-EntB with Aroyl- and Acyl-CoAs.	40
2.2.7 Formation of Misacylated-holo-EntB by EntE Catalyzed Loading of Benzoic Acid Derivatives onto Holo-EntB.	41
2.2.8 Steady-state Kinetic Constant Determinations for EntH Thioesterase-catalyzed Hydrolysis of Acyl-CoA and Acyl-holo-EntB.	41
2.2.9 HPLC Analysis of Formation of 2,4-DHB by Hydrolysis of 2,4-DHB-holo-EntB by EntH.....	42
2.2.10 ESI-MS Analysis of the Hydrolysis of 2,4-DHB-holo-EntB by EntH.	42
2.2.11 ESI-MS Analysis of the Processivity of Acylated- or Aroylated-holo-EntB by EntF.....	43
2.3 Results and Discussion.....	43
2.3.1 Correlation of YbdB (EntH) Function and Enterobactin Biosynthesis.	43
2.3.2 Expression and Purification of EntH and Enzymes of the Lower Pathway of Enterobactin Pathway.	44
2.3.3 Misacylation of Holo-EntB by EntD and Acyl-CoA.....	46
2.3.4 EntE Misacylation of Holo-EntB with Aromatic Acids Plus ATP.....	48
2.3.5 Steady-state Kinetic Constant Determinations for EntH Catalyzed Thioester	

Hydrolysis of Acyl-CoAs and Acyl-holo-EntBs.	51
2.3.6 ESI-MS Analysis of the Processivity of Acylated- or Aroylated-holo-EntB by EntF.....	56
2.4 Conclusion	58
2.5 References	59

CHAPTER 3. STRUCTURE-FUNCTION ANALYSIS OF THE *E. COLI* ENTEROBACTIN PATHWAY THIOESTERASE ENTH (Ybdb).....61

3.1 Introduction	61
3.2 Experimental	62
3.2.1 EntH(Ybdb) Crystallization and X-ray Diffraction Data Collection. (work carried out by Dr. Rui Wu at Boston University).....	62
3.2.2 EntH(2,4-Dihydroxyphenacyl-CoA) or EntH(Phenacyl-CoA) Complex Structure Refinement. (work carried out by Dr. Rui Wu at Boston University) ..	64
3.2.3 Synthesis of Acyl-CoA analog inhibitors	64
3.2.4 Enzymatic Assay.....	65
3.2.5 Hotdog Thioesterase Sequence and Structure Alignments, and Graphics...66	
3.2.6 ³¹ P-NMR Spectrum of HS-CoA and Phenacyl-CoA in the Presence and in the Absence of EntH.	66
3.3 Results and Discussion.....	66
3.3.1 Synthesis and Evaluation of Inert Substrate Analogs for Co-crystallization with EntH.	66
3.3.2 Quaternary and Tertiary Structure of the Apo and Liganded EntH.....	68
3.3.3 An Overview of the EntH Substrate Binding Site	72
3.3.4. Comparative Analysis of the EntH ACP/CoA Binding Site.....	75
3.4 References	88

CHAPTER 4. KINETIC AND STRUCTURAL STUDIES OF HOTDOG FOLD THIOESTERASE, E. COLI YDI I90

4.1 Introduction	90
4.2 Experimental	91
4.2.1 YdiI Crystallization and X-ray Diffraction Data Collection	91
4.2.2 YdiI (2,4-Dihydroxyphenacyl-CoA, Phenacyl-CoA or Undecan-2-one-CoA) Complex Structure Refinement.....	92
4.2.3 Synthesis of Acyl-CoA analog inhibitors	92
4.2.4 Enzymatic Assay.....	93
4.2.5 Sequence Aligment of YdiI, EntH, HThem2, PaaI and Arthrobacter sp. SU 4-HBA-CoA Thioesterase.....	94
4.3 Results	94
4.3.1 Steady-state Kinetic Constants of YdiI Catalyzed Hydrolysis of Acyl-CoA and Acyl-holo-EntB.....	94
4.3.2 Overall fold and oligomeric association of the YdiI-inhibitor complexes. ...	96
4.3.3 The 2,4-Dihydroxyphenacyl-CoA, Phenacyl-CoA and Undecan-2-one-CoA ligands.....	99
4.3.4 The Substrate Binding Site of YdiI.....	100
4.3.5 The Substrate Model.....	104
4.4 Discussion	107
4.5 References	109

CHAPTER 5 HUMAN BROWN FAT INDUCIBLE THIOESTERASE VARIANT 2 (BFIT2): CELLULAR TARGETING AND CATALYTIC FUNCTION.....111

5.1 Introduction	111
5.2 Experimental	113
5.2.1 Materials	113

5.2.2 Expression Vectors and Transfection of HEK293T/17 Cells.....	114
5.2.3 Immunoblot Analysis.....	114
5.2.4 Immunoprecipitation of Recombinant BFIT2 from HEK293T/17 Stably Transfected Cells.	115
5.2.5 Confocal Imaging of GFP Fusion Protein.	116
5.2.6 Mitochondrial Fraction Isolation from HEK293T/17 Cells.	116
5.2.7 In vitro Processing and Importing of BFIT2 by Freshly Isolated Mitochondria from Untransfected HEK293T/17 cells.	116
5.2.8 In silico Analysis of N terminal Region of BFIT2.	117
5.2.9 MALDI-TOF Mass Spectrometry Analysis of BFIT2.....	117
5.2.10 HPLC Analysis of Thioesterase Activities of BFIT2 towards Acyl-CoAs.	118
5.2.11 Digitonin Treatment of Mitochondrial Fraction Freshly Isolated from BFIT2 Stably Transfected HEK293T/17 Cell.	118
5.3 Results.....	119
5.3.1 Immunoblot of Recombinant BFIT2.	119
5.3.2 Isolation of BFIT2 by Immunoprecipitation and Subsequent MALDI-TOF Mass Spectrometry Analysis.	120
5.3.3 Posttranslational Modification of Precursor BFIT2 Occurs at Its N-terminal Region.	122
5.3.4 In silico Analysis of N terminal Region of BFIT2 reveals the presence of amphipathic helix.....	123
5.3.5 Cellular Localization of BFIT2 N-terminal region-GFP Expression Constructs.	126
5.3.6 Mitochondrial Processing and Import of Precursor BFIT2.	127
5.3.7 BFIT2 Localizes in the Matrix of Mitochondria.	130

LIST OF FIGURES

- Figure 1.1** X-ray crystal structure of apo-FabA. **(A)** monomeric structure of FabA and **(B)** dimeric structures of FabA. Cartoon representation is colored in rainbow ranging from blue (N-terminus) to red (C-terminus). **(C)** catalytic site. Structure elements within a single catalytic site carrying catalytically relevant residues are represented in cartoon, remaining structure in ribbon. Protomer A is highlighted in white with His70 shown in sticks and protomer B in green with Asp84 shown in sticks. Atomic coloring is as follows, carbon in white or green, oxygen in red and nitrogen in blue.....2
- Figure 1.2** X-ray crystal structure of apo-4-HBA-CoA thioesterase from *Arthrobacter*. *sp* strain SU. **(A)** monomeric and **(B)** dimeric structures. Cartoon representation is colored in rainbow ranging from blue (N-terminus) to red (C-terminus). **(C)** Catalytic site Structure elements within a single catalytic site carrying catalytically relevant residues are represented in cartoon, remaining structure in ribbon. Protomer A is highlighted in white with His64 and Gln58 shown in sticks and protomer B in green with Glu73 shown in sticks. Atomic coloring is as follows, carbon in white or green, oxygen in red and nitrogen in blue.....3
- Figure 1.3** Illustration of the active site **(A)** and catalytic roles of Asp84 and His70 of FabA **(B)**.....5
- Figure 1.4** Snapshot of active site **(A)** of *Arthro* 4-HBT and proposed catalytic role of a conserved Asp/Glu residue **(B)** in Hotdog thioesterase as base or

nucleophile.	6
Figure 1.5 4-Chlorobenzoate dehalogenation pathway catalyzed by (A) 4-CBA-CoA ligase (CBAL), (B) 4-CBA-CoA dehalogenase, and (C) 4-HBA-CoA thioesterase.	8
Figure 1.6 Genetic organization of paa cluster of catabolism of phenylacetic acid in <i>E. Coli</i> (A) and putative role of PaaI as an thioesterase in phenylacetic acid pathway (B).	10
Figure 1.7 The calicheamicin structure (A) and the catalytic role of CalE8 and hotdog thioesterase CalE7 in calicheamicin biosynthesis (B)	11
Figure 1.8 Gene cluster (A) and biosynthesis of fluoroacetate (B), and proposed function of Flk in the biosynthesis of fluoroacetate (C).	13
Figure 1.9 Sequence alignment of <i>Arthrobacter sp</i> 4-HBT, PaaI, CalE7 and Flk (A) and tertiary structure of monomers of <i>Arthrobacter sp</i> 4-HBT, PaaI, CalE7 and Flk. Secondary structures are color ramped starting with the blue N-terminus and extending to the red C-terminus (B).	18
Figure 1.10 Tertiary structures of functional units of <i>Arthro sp</i> 4-HBT, PaaI, Flk, CalE7, <i>E coli</i> TEII and CACH colored in rainbow, ramped from blue (N-terminus) to red (C-terminus).	21
Figure 1.11 Snapshots of the ligand binding site and the catalytic site (A) and the proposed catalytic mechanism (B) of <i>Arthrobacter sp.</i> 4-HBT.	25

Figure 1.12 Snapshots of the ligand binding site (A), and the catalytic site (B) and the proposed catalytic mechanism (C) of fluoroacetyl-CoA thioesterase.	25
Figure 1.13 Snapshots of the ligand binding site and the catalytic site (A) and the proposed catalytic mechanism (B) of the hotdog thioesterase CalE7.	29
Figure 2.1 Iron-chelating siderophore enterobactin and Fe (III) bound Enterobactin ...	34
Figure 2.3 Lower chemical pathway for biosynthesis of enterobactin.	37
Figure 2.4 Growth curves of wild type BW2513, $\Delta entH$ /BW2513 and $\Delta entB$ /BW2513 in iron-rich or iron-depleted LB medium: (●) wild type BW2513, (◆) $\Delta entB$ /BW2513 and (■) $\Delta entH$ /BW2513 grown in the absence of pyridyl (iron rich medium); (●) wild type BW2513, (◆) $\Delta entB$ /BW2513 and (■) $\Delta entH$ /BW2513 grown in the presence of 400 μ M pyridyl (iron-depleted conditions). Growth was monitored at 600 nm for up to 12 h following the inoculation of the respective culture flasks.	44
Figure 2.5 Exact molecular weight determination of EntH (YbdB) and enzymes in the lower pathway of enterobactin biosynthesis.	45
Figure 2.6 ESI-MS spectra of the reaction mixtures formed in the reaction of EntD with A) lauroyl-holo-EntB; B) palmitoyl-holo-EntB; C) hexanoyl-holo-EntB; D) isobutyryl-holo-EntB.	48
Figure 2.7 ESI mass spectra of A) apo-EntB, B) holo-EntB and of the reaction products C) 2,3-DHB-holo-EntB; D) 2,4-DHB-holo-EntB; E) 4-HBA-holo-EntB; F)	

2-HBA-holo-EntB; G) 3-HBA-holo-EntB; H) benzyol-holo-EntB formed from the reaction of holo-EntB, ATP and the corresponding aromatic acid in the presence of EntE.....	51
Figure 2.8 ESI-MS analysis of EntH catalyzed 2,4-DHB-holo-EntB hydrolysis at pH 7.5 and 25°C.....	53
Figure 2.9 ESI-MS analysis of the hydrolysis of 2,4-DHB-S-holo-EntB by incubation with 0.5 μ M, 2 μ M and 8 μ M of EntH for 15 min at pH 7.5 and 25 °C.....	54
Figure 2.10 HPLC analysis of formation of 2,4-DHB during the hydrolysis of 2,4-DHB-holo-EntB by EntH at 314 nm at 25°C. Lane 1(control): 2,4-DHB-holo- EntB incubated for 25 min in the absence of EntH; Lane 2 to 4: 2,4-DHB-holo-EntB incubated with EntH	55
Figure 2.11 ESI-MS spectra measured as a function of incubation time for reactions of holo-EntF with A) 2,3-DHB-holo-EntB, B) 2,4-DHB-holo-EntB, C) 2,4-DHB-holo-EntB without ATP and L-ser and D) lauroyl-holo-EntB.....	57
Figure 3.1 Chemical structure of the inert substrate analogs, 2,4-dihydroxyphenacyl-CoA and phenacyl-CoA (A). Initial velocity data measured for EntH (0.0645 μ M) catalyzed hydrolysis of benzoyl-CoA (15 to 200 μ M) in 50 mM HEPES (pH 7.5, 25 °C) containing 1 mM DTNB and 35, 70 and 105 μ M phenacyl-CoA (B) or 5 and 10 μ M 2,4-dihydroxyphenacyl-CoA (C)..	67
Figure 3.2 (A) The EntH dimer bound with 2,4-dihydroxyphenacyl-CoA (subunit A in green and subunit B in gray). (B) Illustration of the tertiary structure of the	

EntH monomer. **(C)** and **(D)** The EntH tetramer (bound with phenacyl-CoA) shown in two different orientations.....71

Figure 3.3 **(A)** The overlay of the *apo*-EntH (black) and EntH (2,4-dihydroxy phenacyl-CoA) complex (green). **(B)** The overlay of the *apo*-EntH (black) and EntH (phenacyl: CoA) complex (red). The ligands are colored blue in both figures. The small loop (His47-Gly51), which leads into the N-terminus of the central α helix and that flanks the ligand binding site, is indicated in both figures.....72

Figure 3.4 A cartoon representation of the structure of the EntH(phenacyl-CoA) complex showing the location of the phenacyl-CoA ligand relative to the three subunits that form the substrate binding site. The phenacyl-CoA ligand is shown in stick representation (carbon atoms light blue, nitrogen blue, oxygen red, sulfur yellow and phosphorus orange).73

Figure 3.5 The structure of the EntH(phenacyl-CoA) complex with the phenacyl-CoA ligand shown in stick representation (carbon atoms cyan, nitrogen blue, oxygen red, sulfur yellow and phosphorus orange) and the surfaces of three subunits that contribute to the substrate binding site colored yellow, light gray and dark gray. The panel on the left depicts the phenacyl-CoA nucleotide moiety at the protein surface. The panel to the right depicts that tunnel through which the phosphopantetheinyl arm threads (for clarity the nucleotide moiety is not shown).....74

Figure 3.6 Two snapshots of the aroyl thioester binding pocket observed in the structure of the EntH(phenacyl-CoA) complex. The phenacyl-CoA ligand shown in stick representation (carbon atoms magenta, nitrogen blue, oxygen red, and sulfur yellow).75

Figure 3.7 The EntH-EntB docking model. The cartoon representation of the EntB apo-ACP domain is colored cyan with the side chain of the Ser phosphopantetheinyl attachment residue shown in stick representation (colored carbon green, oxygen red). The EntH surface is colored red for oxygen, blue for nitrogen and white for carbon. The phenacyl-CoA ligand is shown in stick (carbon and phosphorus orange, oxygen red, nitrogen blue).76

Figure 3.8. Stereoscopic representation of the ligand-binding interactions observed in the structure of the EntH (phenacyl-CoA) complex. The coloring is as follows: gray for carbon atom, blue for nitrogen atom, red for oxygen atom, yellow for sulfur atom and brown for phosphorus atom. The two colors (gray and light green) of the ribbon trace correspond to the two molecules within the dimer. The light blue ribbon corresponds to a subunit of a neighboring dimer that flanks the nucleotide-binding site.....77

Figure 3.9 The composite omit map of **(A)** 2,4-dihydroxyphenacyl-CoA ligand in subunit A of EntH (2,4-dihydroxyphenacyl-CoA) complex, **(B)** phenacyl-CoA in subunit D of EntH (phenacyl-CoA) complex and **(C)** undecan-2-one-CoA in hTHEM2 (undecan-2-one-CoA) complex (PDB ID 3F5O). The density of the map (shown in blue) is contoured at the level of 1σ . The

ligand carbon atoms are colored black, the EntH carbon atoms are colored white, and all oxygen, nitrogen sulfur and phosphorus atoms are colored red, blue, yellow and magenta respectively.79

Figure 3.10 (Top) ^{31}P -NMR spectrum of a solution containing 1 mM phenacyl-CoA, 50 mM HEPES (pH 7.5) and 50 mM NaCl (singlet (3'-P) at 5.02 ppm and two doublets (5'-PP) at -10.02 to -9.60 ppm). **(Middle)** ^{31}P -NMR spectrum of a solution containing 1 mM phenacyl-CoA, 0.5 mM EntH, 50 mM HEPES (pH 7.5) and 50 mM NaCl. **(Bottom)** ^{31}P -NMR spectrum of a solution containing 1 mM phenacyl-CoA, 1.0 mM EntH, 50 mM HEPES (pH 7.5) and 50 mM NaCl.81

Figure 3.11 (Top) ^{31}P -NMR of a solution containing 1 mM CoA, 50 mM HEPES (pH 7.5) and 50 mM NaCl. **(Bottom)** ^{31}P -NMR of a solution containing 1 mM CoA, 1 mM EntH, 50 mM HEPES (pH 7.5) and 50 mM NaCl.82

Figure 3.12 Topological and electrostatic features of the substrate binding site of **(A)** and **(E)** EntH bound with phenacyl-CoA, **(C)** hTHEM2 bound with undecan-2-one-CoA (the adenine moiety is not shown for clarity), **(B)** and **(D)** *Arthrobacter* 4-HBA-CoA thioesterase bound with 4-hydroxyphenacyl-CoA. The electropositive area is represented in blue color, and the electron rich surface is represented in red color.84

Figure 3.13 (A) and **(B)** The catalytic sites defined by the X-ray structures of EntH bound **(A)** 2,4-dihydroxyphenacyl-CoA and **(B)** phenacyl-CoA bound to

EntH.	86
Figure 3.14 Stereoscopic representation of a model of 2,4-dihydroxybenzoyl-CoA bound to EntH.	87
Figure 4.1 Co-location of <i>ydiI</i> gene with <i>suf</i> operon in <i>Escherichia coli</i> K-12 genome..	90
Figure 4.2 Chemical structures of the inert substrate analogs (A) and competitive inhibition plots for 2,4-dihydroxyphenacyl-CoA (B) , phenacyl-CoA (C) and undecan-2-one-CoA (D) . YdiI catalyzed hydrolysis of benzoyl-CoA was monitored by DTNB assay using 0.0645 μM YdiI, (15 to 200 μM) benzoyl-CoA, (5 to 20 μM) phenacyl-CoA, or (2.5 to 10 μM) 2,4-dihydroxyphenacyl-CoA or (15 to 30 μM) undecan-2-one-CoA.	97
Figure 4.3 The asymmetric unit of YdiI bound with (A) 2,4-dihydroxyphenacyl-CoA (B) phenacyl-CoA and (C) undecan-2-one-CoA. Carbon atoms of ligands are colored white, and all oxygen, nitrogen, sulfur and phosphorus atoms are colored red, blue, yellow and magenta respectively	99
Figure 4.4 The electron density map of (shown in grey) of YdiI bound (A) 2,4-dihydroxy phenacyl-CoA, (B) phenacyl-CoA and (C) undecan-2-one-CoA ligands. The density maps are contoured at the level of 1σ	100
Figure 4.5 Stereoscopic representation of atomic interactions associated with the YdiI substrate-binding site. Key interactions of (A) 2,4-dihydroxyphenacyl-CoA and (B) undecan-2-one-CoA (ligand P in Figure 4.2 C) bound to YdiI. Atomic coloring is as follows: blue for nitrogen atom, red for oxygen atom,	

yellow for sulfur atom and brown for phosphorus atom. The two colors (white and light green) of the ribbon trace correspond to the two molecules within the dimer. The light blue ribbon corresponds to a subunit of a neighboring dimer that flanks the nucleotide binding site. The carbon atoms are colored according to which subunit the residue is located on, and the carbon atoms of the ligand are shown in cyan.103

Figure 4.6 Stereoscopic representation **(A)** 2,4-dihydroxyphenacyl-CoA (ligand D in Figure 4.4 C) and **(B)** undecan-2-one-CoA (ligand P in Figure 4.4 D) bound to YdiI. **(C)** Stereoscopic representation of a model of decanoyl-CoA bound to YdiI.106

Figure 4.7 (A) Stereoscopic view of the superimposed active site showing the amino acid residues of YdiI (black), EntH (green), *Arthrobacter* 4-HBA: CoA thioesterase (magenta), hTHEM2 (brown) and *E. coli* PaaI (cyan). **(B)** Structure based sequence alignment of YdiI with other hot dog fold thioesterases created with ESPript. Protein sequences were from YdiI (YdiI (undecan-2-one-CoA) complex chain E&F), EntH (1VH9), Arthr. sp (*Arthrobacter* sp 4HBT 1Q4T), *E. coli* (*E. coli* PaaI 2FS2) and hTHEM2 (human thioesterase superfamily member 2 3F5O). Sequence accession numbers of the Swiss-Prot database are P77781, POA8Y8, Q04416, P76084 and Q9NPJ3 respectively. The secondary structure of YdiI, which is mainly defined by the analysis of the structure using DSSP program, is indicated above the alignment. The identical residues were colored in red and similar

residues in yellow.....106

Figure 5.1 Immunoblot of untransfected and *BFIT2/pcDNA4 transfected* HEK cell lysates. Various BFIT2 construct utilized are illustrated in *panel A*. BFIT2 includes two hotdong thioesterase domains (HD-THIO), a steriodogenic acute regulatory protein related lipid transfer domain (START) and a N-terminal region (N-term). Full length (precursor) BFIT2 and BFIT2 lacking the 1-34 amino acids at its N-terminal region (Δ N33 BFIT2) were introduced 35 amino acids to their N-terminus by expression vector, which includes Xpress and HisG epitopes. Shown in *panel B* is Immunoblot of both untransfected and *BFIT2/pcDNA4 transfected* HEK cell lysates using anti-Xpress, anti-HisG and anti-BFIT2 antibody respectively.120

Figure 5.2 Immunoblot monitoring of isolation of precursor BFIT2 by immunoprecipitation. Immunoprecipitation was carried out using anti-Xpress antibody immobilized agarose beads (see experimental for detail). Eluents obtained from each steps were loaded to SDS-page and then nitrocellulose membrane, which subsequently subject to immunoblot using anti-Xpress antibody and anti-BFIT2 antibody for antigen detection respectively. Immunoblots confirmed that only the precursor BFIT2 was pulled down by anti-Xpress coupled beads.....122

Figure 5.3 Immunoblot of ΔN 34 *BFIT2/pcDNA4 HisMax* ransfected HEK cell lysates. Transfection yielded BFIT2 lacking 1-34 amino acid of its N-terminal region (ΔN -34 BFIT2). Immunoblot of transfected cell lysate only detect

one protein corresponding to Δ N-34 BFIT2 using either anti-Xpress or anti-BFIT2 antibody, indicating the abolishment of posttranslational modification after removal of part of the N-terminal region of BFIT2.123

Figure 5.4 Modeled tertiary structure of N-terminal region of BFIT2 and the electrostatic vacuum surface of the central helix. A. predicted tertiary structure of BFIT2 N-terminal region; B. electrostatic vacuum surface of central helix together with its adjacent region (residues 9-24). One face of the surface is exclusively positively charged, while the opposite face is mainly uncharged/hydrophobic.....124

Figure 5.5 Amphipathy of N-terminal region of BFIT2 and sequence alignment of the region from various species. A. Emboss pepwheel hexical wheel plot of residue 5-24 of the N-terminal region of BFIT2; B. Sequence alignment and estimated pI value of N-terminal region of BFIT2 from various species. Identical residues are shade in black, residues common to at least two of the three species in gray. Conserved residues with positively charged side chain are marked by asterisk.....125

Figure 5.6 Confocal imaging of BFIT2 N-terminal region-GFP reveals its mitochondrial localization. A. domain organization representation of BFIT2 N(1-42)-GFP and GFP; B. BFIT2 N(1-42)-GFP showed punctuate localization in HEK 293T/17 cells, whose nuclei were stained by Hoechst 33342 and plasma membrane by WGA/Alexa Fluor 594 conjugate; C. The localization of BFIT2 N(1-42)-GFP overlapped with mitotracker CMXRos,

giving a merged image with yellow color.....129

Figure 5.7 In vitro mitochondrial processing and importing of precursor BFIT2. **A.** isolated mitochondria fraction is free of cytosolic component and enriched with mitochondria. mHSP70 is a mitochondria marker (71 kDa) and α -tubulin, the cytosolic marker (50 kDa); **B.** Affinity purified precursor BFIT2 was incubated with mitochondria in mannitol buffer. Aliquots of reaction mixture were extracted at 5 min, 30 min and 60 min respectively and subject to immunoblot analysis. Precursor BFIT2, free of its mature form, was posttranslationally processed after incubation with mitochondria; **C.** Affinity purified precursor BFIT2 was incubated with mitochondria in mannitol buffer. Aliquots were extracted as described above. Supernatant and pellet with mitochondria were separated and subject to immunoblot analysis. BFIT2 were relocated to mitochondria and processed into its mature form.....130

Figure 5.8 Digitonin treatment of mitochondria isolated from BFIT2 stably transfected HEK293T/17 cell. Mitochondria fractions after treated with various concentrations of digitonin at 4°C for 45 mins were subject to western blot by anti-mHSP70, anti-AIF and anti-Xpress (for BFIT2 detection) antibodies. mHSP70 is mitochondrial matrix protein, while AIF is intermembrane space marker. Similar to mHSP70, BFIT2 is insensitive to digitonin treatment and released at high digitonin concentration at 0.6 mg/mL, suggesting its matrix submitochondrial location.....131

Figure 5.9 HPLC analysis of BFIT2 thioesterase activities to acyl-CoAs. Palmitoyl-CoA (**A**), myristoyl-CoA (**B**), lauroyl-CoA (**C**) and butyryl-CoA (**D**) were tested as substrates for affinity purified BFIT2. 50 μ M of one of the acyl-CoA in 50 mM HEPES buffer (pH 7.5) was incubated with 0.34 μ M BFIT2 (see experimental for details). Reactions were monitored by HPLC (UV detection wavelength at 260 nm) at 30 min and 60 min respectively at 37°C. Palmityl-, myristoyl-, lauroyl- and butyr- CoAs have retention time of 19.3 min, 16.9 min, 15.1 min and 6.4 min respectively. As control experiment, 50 μ M of one of the acyl-CoA was incubated in 50 mM HEPES in the absence of BFIT2 for 2 hours at 37°C, before subject to HPLC analysis. .133

LIST OF TABLES

Table 2.1 Exact molecular weight determination of His ₆ -EntH and His ₆ -enzymes of the lower Enterobactin pathway.	46
Table 2.2 Summary of ESI mass spectral data for the formation of misacylated holo-EntB by EntD catalyzed reactions between apo-EntB various acyl-CoAs. ...	48
Table 2.3 Summary of the ESI mass spectral data measured for the acyl-holo-EntB products formed from the reaction of holo-EntB, ATP and the corresponding aromatic acid in the presence of EntE and the formation of the misacylated holo-EntBs via thioesterification of benzoic acid derivatives to ArCP domain catalyzed by EntE.	49
Table 2.4 Steady-state kinetic constants for EntH catalyzed hydrolysis of acyl-CoA and acyl-holo-EntB thioesters pH 7.5 and 25 °C measured using the DTNB spectrophotometric assay.....	52
Table 3.1 Crystallographic and refinement statistics for wild-type EntH bound with 2,4-dihydroxyphenacyl-CoA or phenacyl-CoA.....	69
Table 3.2 Steady-state constants for EntH wild type and mutant catalyzed hydrolysis of benzoyl-CoA or salicyl-CoA (if otherwise indicated) in 50 mM K+HEPES (pH 7.5 and 25 °C) and 1 mM DTNB ^a	86
Table 4.1 Steady-state kinetic constants for YdiI catalyzed hydrolysis of acyl-CoA and acyl-S-holo-EntB thioesters measured at pH 7.5 and 25 °C using DTNB	

spectrophotometric assay.....	95
-------------------------------	----

CHAPTER 1

DIVERGENCE OF BIOLOGICAL FUNCTION WITHIN THE HOTDOG FOLD THIOESTERASE FAMILY

1.1 Overview of the Hot Dog Fold Protein Superfamily

The first reported X-ray crystal structure of a “Hotdog” fold protein is that of FabA, a β -hydroxydecanoyl thioester dehydratase (1). The tertiary structure of FabA shows a mixed α/β fold with a central α -helix (sausage) wrapped by 6 anti-parallel β -strands which is now recognized as characteristic of Hotdog fold superfamily (Figure 1.1 A). The typical secondary structure of a monomeric Hotdog FabA from N- to C-terminus is a short β sheet (β_0)- α_1 - β_1 - β_2 - α_2 - α_3 - β_3 - β_4 - β_5 - β_6 . Common to members of this protein family, the functional unit is a dimer, stabilized by hydrogen bonds formed between the respective β_3 -strands and the stacking of the central α_3 -helices of the two protomers (Figure 1.1 B). The catalytic site of FabA, as for other typical Hotdog proteins, is located at the interface of the dimer, delineated by the α_2 helix, the N-terminus of α_3 helices and the loop linking the two helices of one subunit and the middle section of α_3 helix from the other subunit (Figure 1.1 C).

The hotdog fold superfamily consists of three major chemical function classes: thioesterase, dehydratase and enoyl hydratase (2). The thioesterases (for example 4-hydroxybenzoyl-CoA thioesterases (4-HBT) and 4-HBT like-clade members (3-6),

including PaaI (7-8), human hotdog-fold thioesterase family member 2 (hThem2) (9-10) are the most well defined in a structure-function level.

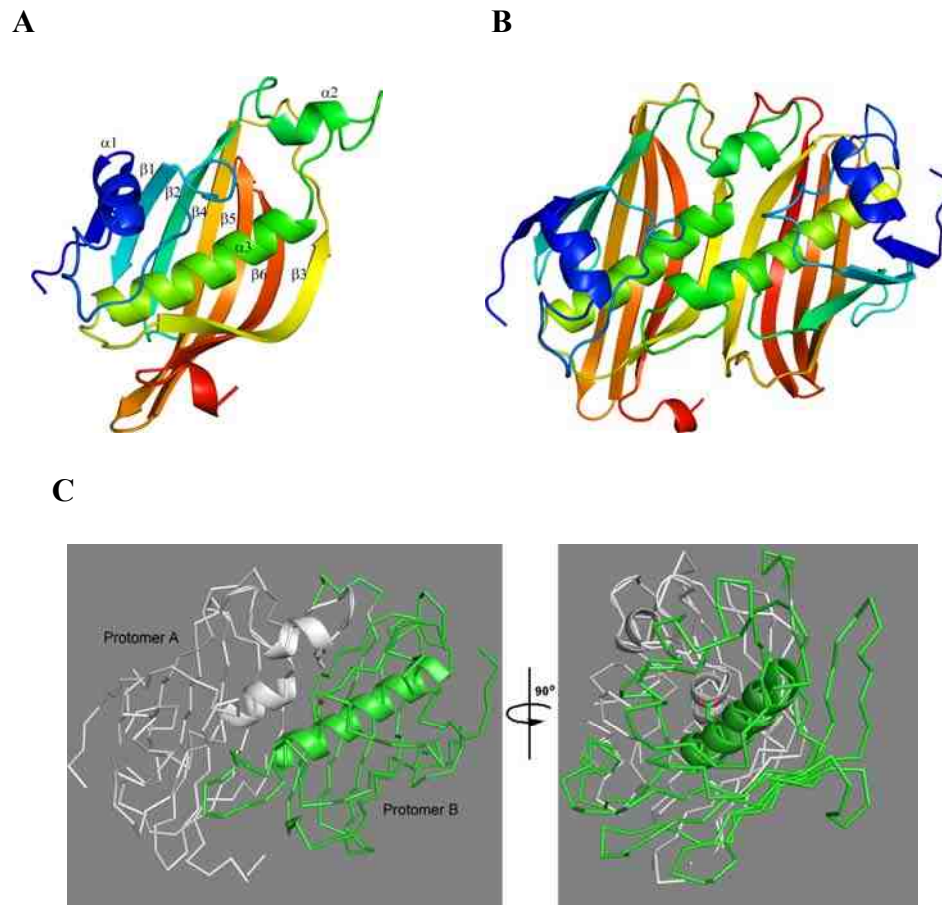


Figure 1.1 X-ray crystal structure of apo-FabA. **(A)** monomeric structure of FabA and **(B)** dimeric structures of FabA. Cartoon representation is colored in rainbow ranging from blue (N-terminus) to red (C-terminus). **(C)** catalytic site. Structure elements within a single catalytic site carrying catalytically relevant residues are represented in cartoon, remaining structure in ribbon. Protomer A is highlighted in white with His70 shown in sticks and protomer B in green with Asp84 shown in sticks. Atomic coloring is as

follows, carbon in white or green, oxygen in red and nitrogen in blue.

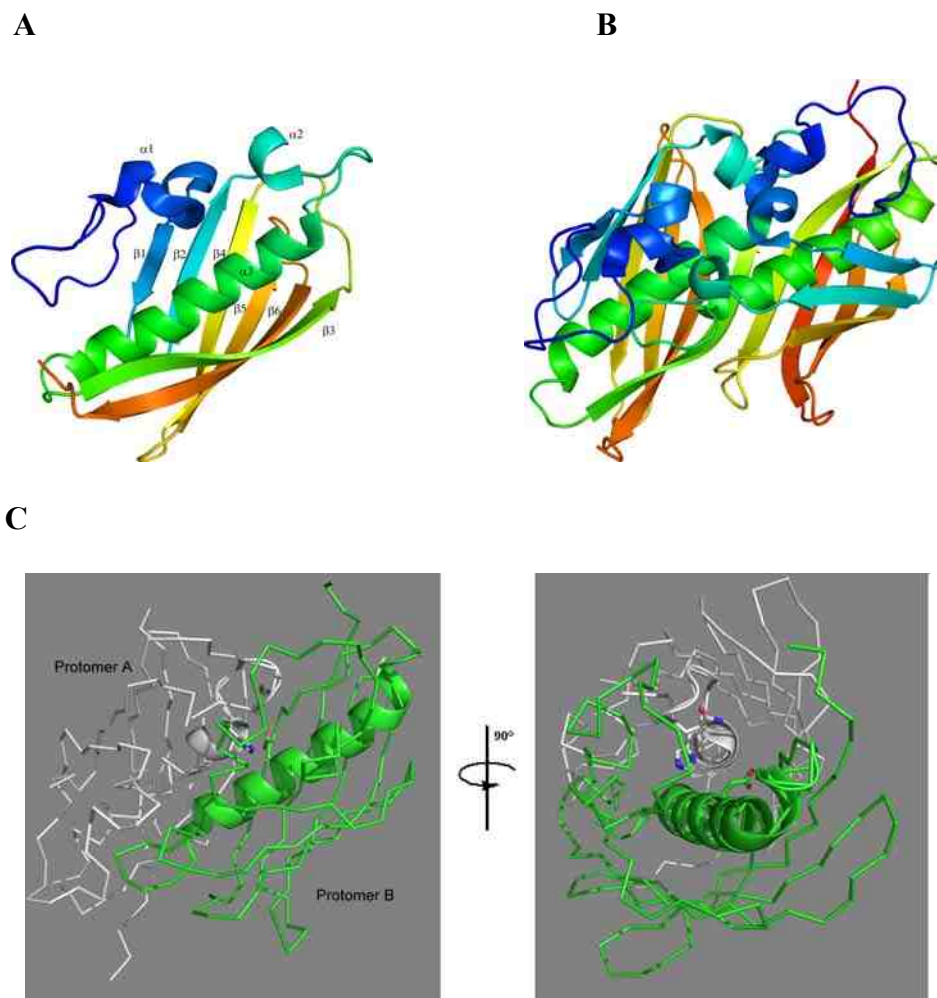


Figure 1.2 X-ray crystal structure of apo-4-HBA-CoA thioesterase from *Arthrobacter. sp* strain SU. **(A)** monomeric and **(B)** dimeric structures. Cartoon representation is colored in rainbow ranging from blue (N-terminus) to red (C-terminus). **(C)** Catalytic site structure elements within a single catalytic site carrying catalytically relevant residues are represented in cartoon, remaining structure in ribbon. Protomer A is highlighted in white with His64 and Gln58 shown in sticks and protomer B in green with Glu73 shown in

sticks. Atomic coloring is as follows, carbon in white or green, oxygen in red and nitrogen in blue.

The overall structure of the *Arthrobacter. sp* strain SU 4-HBT monomer is quite similar to that of FabA (Figure 1.2 A). Firstly, the secondary structure is the same as FabA: $\alpha 1$ - $\beta 1$ - $\beta 2$ - $\alpha 2$ - $\alpha 3$ - $\beta 3$ - $\beta 4$ - $\beta 5$ - $\beta 6$. Secondly, the anti-parallel β -sheet has an identical sequential order of 1-2-4-5-6-3. Thirdly, the functional dimer of 4-HBT is associated by extensive interactions between $\beta 3$ strands and $\alpha 3$ helices of the two protomers (Figure 1.2 B). In addition, two short helices, $\alpha 1$ and $\alpha 2$, are involved in dimerization. Fourthly, the catalytic sites of 4-HBT are also formed at the dimer interface with key residues located on $\alpha 3$, $\alpha 2$ and the connecting loop of one subunit, and the $\alpha 3$ helices of the other subunit (Figure 1.2 C).

Although sharing the same hotdog fold (albeit with low sequence identity), FabA and 4-HBT carry out distinct chemical reactions. 4-HBT catalyzes the hydrolytic cleavage of 4-hydroxybenzoyl-CoA thioester bond, releasing 4-hydroxybenzoic acid (4-HBA) and CoA, whereas FabA dehydrates (*R*)-3-hydroxydecanoyl-ACP to (*E*)-2-decenoyl-ACP. Through mutagenesis and X-ray structural structure studies, key residues of the two hotdog fold proteins are identified: His70 and Asp84 in the case of FabA, and His64, Gln58 and Glu73 in the case of *Arthrobacter* 4-HBT.

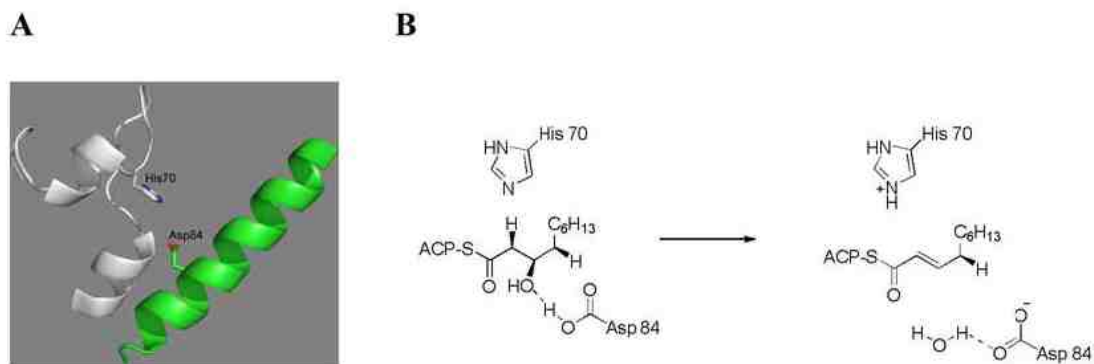


Figure 1.3 Illustration of the active site (**A**) and catalytic roles of Asp84 and His70 of FabA (**B**).

Interestingly, oriented at similar position in catalytic sites, His70/Asp84 pair in FabA and His64/Glu73 in *Arthro* 4-HBT assume different catalytic roles (Figure 1.3 A and 1.4 A), which support the different chemistry of the two enzymes. His70 and Asp84, located on different FabA subunits, serve as acid/base catalysts by shuttling protons during the dehydration reaction, while Glu73 of 4-HBT functions as a base or nucleophile in hydrolytic activities towards 4-HBA-CoA (Figure 1.3 B and 1.4 B). This active site carboxylate residue is well conserved in hotdog thioesterase members and can be either a Glu or Asp. Depending on the particular enzyme it acts as a base to activate the water nucleophile or as a nucleophile in attacking carbonyl carbon and mediating the transfer of the acyl group to water in a two-step reaction pathway.

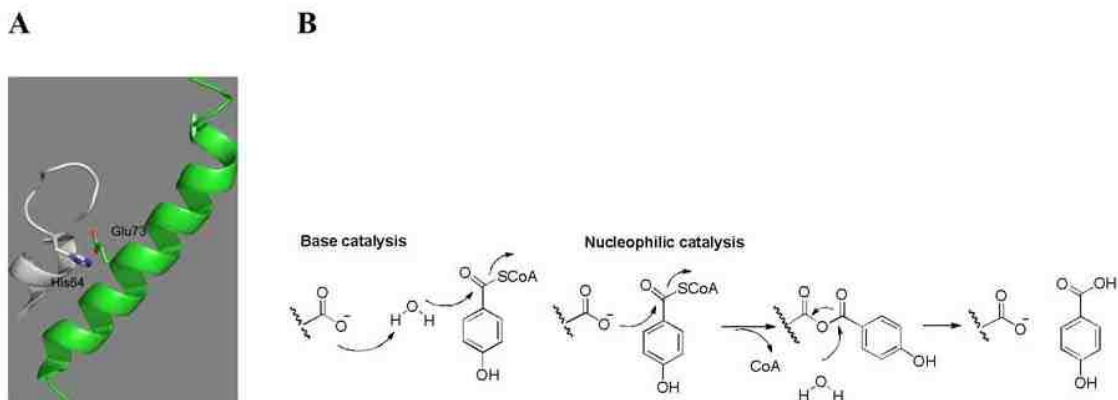


Figure 1.4 Snapshot of active site (A) of *Arthro* 4-HBT and proposed catalytic role of a conserved Asp/Glu residue (B) in Hotdog thioesterase as base or nucleophile.

Although catalyzing different reactions, the two typical hotdog fold enzymes target a thioester substrate. Unlike FabA whose substrate is ACP thioester, 4-HBT acts on acyl-CoAs. Although the physiological substrate is 4-HBA-CoA, 4-HBT shows hydrolytic activity with other aryl-CoA thioesters (5). Indeed, substrate promiscuity is a common feature of hotdog fold thioesterases (9,11-12). Recent reports on the interaction between the hotdog thioesterase YbgC and the acyl carrier protein ACP, and on the identification of putative hotdog fold ACP thioesterases from plant and algae suggest that the substrate range of some hotdog thioesterase is not limited to acyl-CoAs but includes acyl-ACPs. X-ray crystal structures of various hotdog fold proteins, show that the phosphopantetheinyl arm, which is common to both acyl-CoAs and acyl-ACPs, binds the the enzyme along a deep channel that leads to the acyl-group binding pocket..

Recent efforts to elucidate the biological functions of hotdog fold thioesterases, such as hTHEM4 (13), fluoroacetyl-CoA thioesterase (FIK) (14), Acot8 (15) and CalE7

(16), have unveiled intriguing involvements of these proteins with a variety of cellular processes, suggesting broad functional divergence of this subfamily. To appreciate the divergence the hotdog fold thioesterase subfamily, detail discussions of representative thioesterases, 4-HBA-CoA thioesterase, PaaI, Acot8, FIK, CalE7 and CACH will be addressed centering on their diverse catalysis mechanisms, distinct cellular functions and structural bases of the functions.

1.2 Divergence of Hotdog Fold Thioesterases

Although functional assignments of most hotdog fold thioesterases have yet to be made, participations of hotdog thioesterases in a broad array of cellular events have been unequivocally evident by detailed catalytic, functional and structural investigations of several members of this protein subfamily. These members include 4-HBA-CoA thioesterases from *Pseudomonas sp.* CBS3 and *Arthrobacter sp.* SU (involved in the dehalogenation pathway of 5-chlorobenzoate), *E. coli* PaaI (involved in the phenylacetic acid degradation pathway), *Homo sapiens* Acot8 (an interacting partner of HIV-Nef protein), *Streptomyces cattleya* FIK (a component of the fluoroacetic acid biosynthetic pathway), CalE7 (involved in polyketide calicheamicin biosynthesis), and the *Homo sapiens* acetyl-CoA hydrolase (CACH). These well-studied hotdog thioesterases represent distinct cellular function, and variations in catalytic mechanism and structure, which together can provide valuable insights into the divergence of the hotdog fold family.

1.2.1 Divergence in Cellular functions

1.2.1.1 4-Hydroxybenzoyl-CoA Thioesterase (4-HBT) in Dehalogenation of 4-Chlorobenzoate

Soil dwelling bacteria, such as *Pseudomonas sp.* CBS3 and *Arthrobacter sp.* SU are capable of surviving on 4-chlorobenzoate (4-CBA) as the only carbon source. Through dehalogenation, these bacterial strains convert 4-chlorobenzoate to 4-hydroxybenzoate (4-HBA), which is further degraded via aromatic catabolic pathways and ultimately converts to acetyl-CoA/succinyl-CoA (17). Three enzymes catalyze the conversion of 4-CBA to 4-HBA, namely 4-chlorobenzoate ligase (CBAL), 4-chlorobenzoyl-CoA dehalogenase and 4-hydroxybenzoyl-CoA thioesterase (4-HBT) (Figure 1.5). 4-HBT catalyzes the final step in the pathway, the hydrolysis of 4-HBA-CoA, releasing 4-HBA and free CoA.

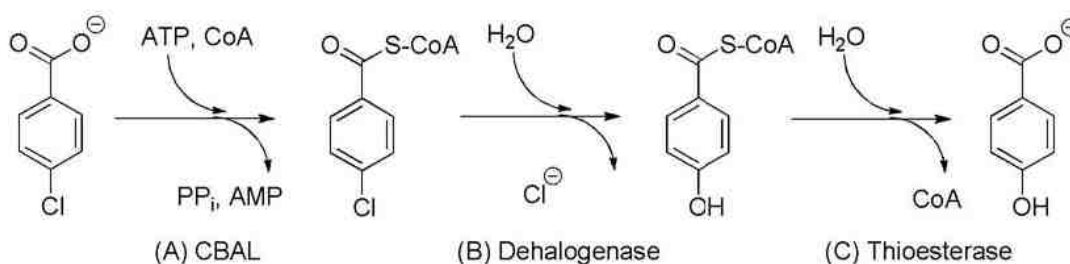


Figure 1.5 4-Chlorobenzoate dehalogenation pathway catalyzed by (A) 4-CBA-CoA ligase (CBAL), (B) 4-CBA-CoA dehalogenase, and (C) 4-HBA-CoA thioesterase.

1.2.1.2 *E. coli* PaaI in Phenylacetic Acid Catabolism Pathway

Bacteria, including *E. coli* and *Pseudomonas* have been previously reported to mineralize a variety of aromatic compounds, including phenylacetic acid and its hydroxylated derivatives (18). A recent effort in elucidating the the phenylacetic acid catabolic pathway revealed a chromosomal gene cluster of the *E. coli* W strain that encodes enzymes involved in the aerobic degradation of phenylacetic acid. This genetic locus is now known as the *paa* cluster (19). The 14 genes are organized in three transcriptional units, *paaZ*, *paaABCDEFGHJK* and *PaaXY* (Figure 1.6 A). Functional annotation predicts that PaaK serves to convert phenylacetate to phenylacetyl-CoA. Analogous to aromatic catabolic pathways from various microorganisms, the *paaABCDE* gene products are predicted to be multi-component oxygenases, which are believed to hydroxylate phenylacetyl-CoA. The *paaFGHIJ* genes encode proteins with high similarities to those involved in fatty acid oxidation. It is likely that the *paaFGHIJ* gene products mediate the downstream degradation steps that connect with central metabolic pathways, such as Krebs cycle.

The gene *paaI*, which encodes a predicted hotdog fold protein, is located within the transcription unit of *paaABCDEFGHJK*. Its identity as a Hotdog thioesterase was established by its demonstrated *in vitro* thioesterase activity and its X-ray structure (8). PaaI efficiently catalyzes the hydrolysis of phenylacetyl-CoA and its ring hydroxylated derivatives, but not the hydrolysis of benzoyl-CoA and its hydroxylated derivatives. The dead end products, 2- and 3-hydroxyphenylacetic acid form from unstable 1,2-dihydrodiol and 2,3-dihydrodiol intermediates of the phenylacetate degradation pathway (Figure 1.6 B) (18,20). The proposed function of PaaI is to release CoA from these dead

end hydroxylated phenylacetyl-CoAs for the purpose of rescuing the CoA.

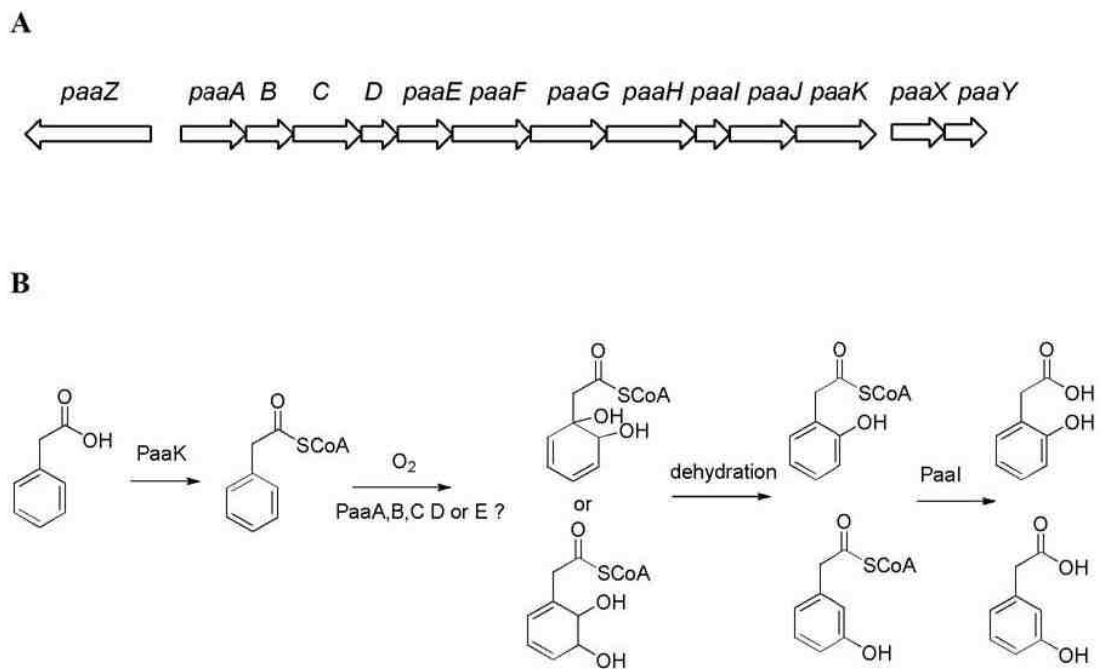


Figure 1.6 Genetic organization of *paa* cluster of catabolism of phenylacetic acid in *E. coli* (A) and putative role of PaaI as an thioesterase in phenylacetic acid pathway (B).

1.2.1.3 CalE7 in Calicheamicin Biosynthesis

Thorson *et al* reported the biosynthetic gene cluster leading to the polyketide calicheamicin, which structure has a 10 membered enediyne core decorated with polysaccharide moieties (Figure 1.7 A) (21). Through BLAST analysis, the genes responsible for the biosynthesis of the enediyne core includes *calE1, E2, E3, E4, E5, E6, E7* and *E8*. The multimodular polyketide synthetase, CalE8 carries out the synthesis of the precursor of calicheamicin enediyne core (Figure 1.7 B) by consuming NADPH and malonyl-CoA in an iterative manner.

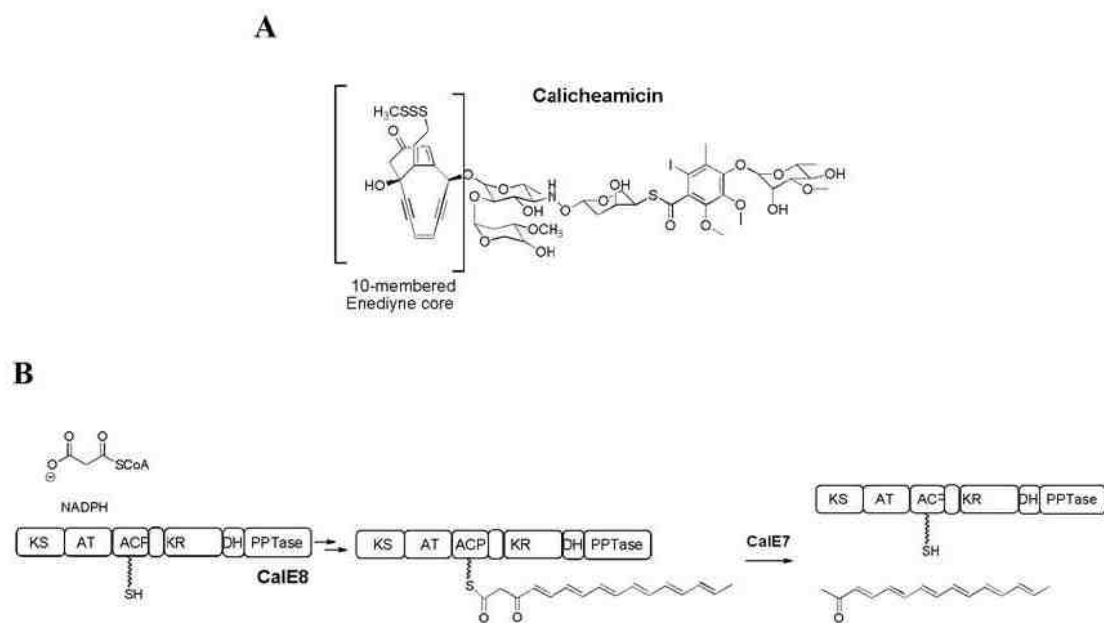


Figure 1.7 The calicheamicin structure (**A**) and the catalytic role of CalE8 and hotdog thioesterase CalE7 in calicheamicin biosynthesis (**B**)

Because CalE8 lacks a terminal *in cis* type I thioesterase domain, it is thought that CalE8 needs an *in trans* thioesterase for product release. Liang *et al* reported that product release from CalE8 is facilitated by the “stand alone” hotdog thioesterase CalE7 (22). Interestingly, NMR and MS based structural determinations of the released product show that major component is a linear carbonyl-conjugated polyene (3,5,7,9,11,13-pentadecen-2-one) (Figure 1.7 B), which is derived from decarboxylation of precursor tethered to the CalE8 pantetheinyl arm. Thus, these results suggest in addition to catalysis of thioester hydrolysis required for product release, CalE7 may also catalyze decarboxylation step. This is the first bi-functional hotdog thioesterase to be reported.

1.2.1.4 FLK in Fluoroacetate Biosynthesis.

The scarcity of fluorometabolites is markedly contrasted with the abundance of fluorine in nature (23). Among the most well known fluorometabolites is fluoroacetate. Huang *et al* reported the discovery of a gene cluster for fluoroacetate biosynthesis from *Streptomyces cattleya*, composed of 12 open reading frame, *flA, B, C, D, E, F, G, H, I, J, K* and *L* (Figure 1.8 A) (24).

Fluoroacetate is assembled by a multistep pathway. The initial reaction is catalyzed by fluorinase FIA, which converts S-adenosylmethionine (SAM) to 5'-fluoro-5'-deoxyadenosine (5'-FDA) by incorporating fluoride and generating methionine (Figure 1.8 B). The 5'-FDA is then converted to the next biosynthetic intermediate, 5-fluoro-5-deoxy-D-ribose-1-phosphate (5-FDRP) by nucleoside phosphorylase FIB (25). Although it is believed that 5-fluoroacetaldehyde is the common precursor to fluoroacetate and fluorothreonine, details of the conversion steps from 4-FDRP to 5-fluoroacetaldehyde remain to be elucidated.

However, *fl* gene cluster does not appear to contain the genes encoding the enzymes for these conversions leading to 5-fluoroacetaldehyde. Thus, a second locus containing the missing genes may be present and remained to be identified.

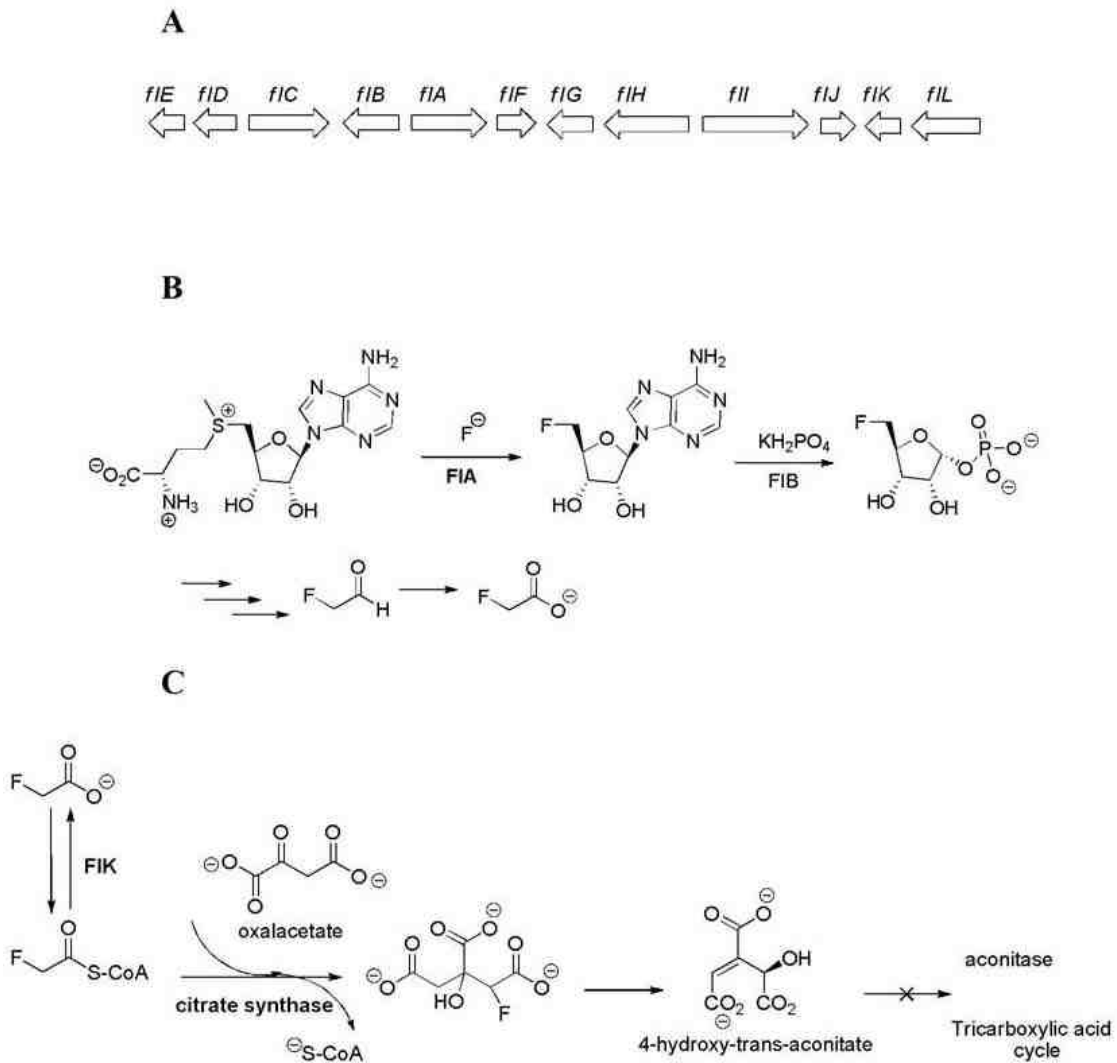


Figure 1.8 Gene cluster (A) and biosynthesis of fluoroacetate (B), and proposed function of FIK in the biosynthesis of fluoroacetate (C).

A hotdog fold protein, FIK, is encoded in the *fl* gene cluster. In vitro studies have shown that FIK catalyzes the hydrolysis of fluoroacetyl-CoA, but no observable activity towards acetyl-CoA was found (14). Fluoroacetyl-CoA does not appear to be a biosynthetic intermediate leading to fluoroacetate. However, it was shown that

fluoroacetate can be converted to fluoroacetyl-CoA, which reacts with oxaloacetate in the presence of the catalyst citrate synthase to generate 2-fluorocitrate (Figure 1.8 C). Further conversion of 2-fluorocitrate would produce aconitase inhibitor, 4-hydroxy-trans-aconitate, which is known to block the operation of tricarboxylic acid cycle. Thus, it is proposed that FIK functions as a thioesterase specifically targeting fluoroacetyl-CoA, thereby conferring fluoroacetate resistance to the fluoroacetate producer without interfering the cellular content of acetyl-CoA, an essential metabolite for cell survival. The remarkable feature of FIK is its ability to discriminate between the essential cellular metabolite acetyl-CoA and the targeted substrate fluoroacetyl-CoA. Whereas most hotdog thioesterases are promiscuous, this one can distinguish a fluorine atom from a hydrogen atom.

1.2.1.5 Acot8, A HIV-1 Nef Binding Protein

hTE, renamed as Acot8 according to the revised nomenclature for mammalian thioesterases, was identified as a potential binding partner to Nef of the human immunodeficiency virus (HIV) by yeast two-hybrid screening and immunoprecipitation (15,26). Acot8 has a predicted hotdog fold, and it shows efficient thioesterase activity to medium to long chain (C₆-C₁₆) fatty acyl-CoAs. Acot8 is homologous to the *E. coli* TEII, which has a known structure consisting of two subunits of tandem hotdog thioesterase domains. Acot8 has a C-terminal Serine-Lysine-Leucine (SKL) peroxisomal importing sequence, thus it might be imported to peroxisome to perform its biological function.

The interacting partner of Acot8, Nef is a myristoylated protein involved in HIV

virus load and progression (27). It was reported to trigger CD4 endocytosis (28-29). The interaction of Nef to Acot8 does not affect the thioesterase activities of Acot8 to an observable extent. This interaction is independent of myristoylation of Nef, whereas 17 C-terminal residues of Nef are indispensable for the interaction as was shown in mutagenesis studies. Although the physiological characterization of Nef-Acot8 interaction awaits comprehensive *in vivo* studies, there are several lines of evidence for the requirement of Nef interaction with Acot8 to down-regulate CD4 expression. This in turn suggests that Acot8 may be a cellular mediator of Nef function in HIV infection.

1.2.1.6 Cytosolic Acetyl-CoA Hydrolase (CACH)

Prass *et al* first reported the cytoplasmic mammalian acetyl-CoA hydrolase known as CACH (and later as Acot12) (30-31). In solution CACH exists as a homodimer or homotetramer, while in the crystal form it is a homotetramer (32). CACH has a modular structure consisting of tandem units of two fused hotdog thioesterase domains and a C-terminal steroidogenic acute regulatory protein domain. The steroidogenic acute regulatory protein domain (START domain) adds to its acetyl-CoA thioesterase activity, the capacity for the binding and transport of a yet to be identified lipid (33). Its close homolog in amino acid sequence and in domain modular design is the human brown fat inducible thioesterase (BFIT). The BFIT synthesis is induced during the physiological state of the thermogenesis transition (34).

Thioesterase activities of CACH are limited to acetyl-CoA and propionyl-CoA, with no observed activities towards malonyl-CoA. Interestingly, the activities of CACH

are regulated by ATP (activator), ADP (inhibitor) and free CoA (inhibitor) *in vitro*. The *in vivo* assay of CACH thioesterase activities show that CACH is up-regulated during enhanced β -oxidation and heightened fatty acid synthesis (35). The biological role of CACH is to release CoA from acetyl-CoA and as CoA needed for fatty acid synthesis and β -oxidation.

In summary, although cellular functions of most of the hotdog fold thioesterases remain to be addressed, sufficient lines of evidence gleaned from well studied thioesterase members support their versatile and divergent functions in a variety of cellular events, including fatty acid synthesis and oxidation, biosynthesis of secondary metabolites, and development and progression of obesity, diabetes, cancers and other diseases.

1.2.2 Divergence in Protein Folding of Hotdog Thioesterase

1.2.2.1 Tertiary Structure

As described previously, studies of hotdog fold thioesterases unveil their distinct features in substrate preferences, specificities, and cellular functions. Recent progress in X-ray crystal structural determination of minimal hotdog fold thioesterase, *Arthrobacter* sp 4-HBT, PaaI, CalE7 and FIK, and those containing internal repeats of Hotdog fold domains, in the case of *E coli* TEII and CACH, enables the pursuit of structural explanations for the marked distinctions among these well studied thioesterases. It is notable that the divergence of thioesterases is also evident in folding.

A

	10	20	30	40	50	60
4-HBT	--MHRTSNGSHATGGNLPDVASHYPVAYEQTLDGTVGFVIDEMTPERATASVEVTDTLRQ					
PaaI	-MSHKAWQNAHAM---YEN-----DACAKALGIDIISMDEGFAVVTMTVTAQMLN					
CalE7	MSMPRYEYRHVVG--FEETN----LVGNVYYVNYLRWQGRCREMFLYEHAP EILDELRA					
Flk	-----MKDGMRVGERFTH-D-----FVVPPHKTVRHLYPESPEFAEFPEVFATGFMVG					

	70	80	90	100	110	120
4-HBT	RWGLVHGGAYCALAEMLA-TEATVAVVHEKGMMAVGQSNHTSFFRP-VKEGHVRAEAVRI					
PaaI	GHQSCHGGQLFSLADT-----AFAYACNSQGLAAVASACTIDFLRPGFAGDTLTATAQVR					
CalE7	DLKLF TLKAECEFFAELAPFDRLAVRMRLVELTQTQME LGFDYLR LGDDLLVARGRQRI					
Flk	LMEWACVRAMAPYLEPGEGLGTAICVTHTAATPPGLTVTVTAELR-----SVEGRRLSW					

	130	140	150
4-HBT	HAGSTTWFDVSLRDDAGRLCAVS---SMSIAVRPRRD-		
PaaI	HQKQTGVYDIEIVNQQQKTVALFRGKSHRIGGTITGEA		
CalE7	ACMRGPNGRTEPVRVPAGLVRAFAPFRSATVGQ-----		
Flk	RVS AHDGVDEIGSGTHERAVIHLEKFNKVRQKTPAG-		

B

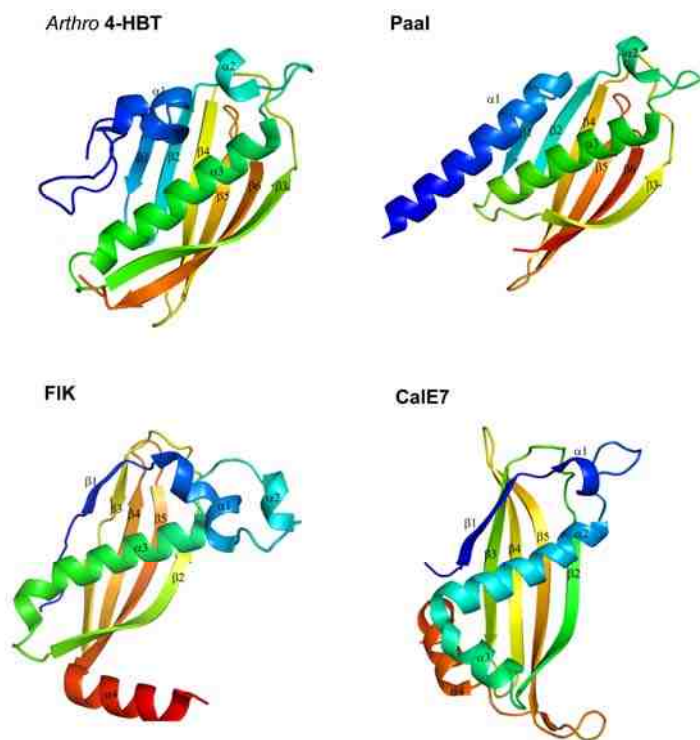


Figure 1.9 Sequence alignment of *Arthrobacter sp* 4-HBT, Paal, CalE7 and FIK (A) and tertiary structure of monomers of *Arthrobacter sp* 4-HBT, Paal, CalE7 and FIK. Secondary structures are color ramped starting with the blue N-terminus and extending to the red C-terminus (B).

Whereas hotdog thioesterases having different biological functions share no significant sequence identity (Figure 1.9 A) their tertiary structures are conserved (viz. central α - helix wrapping around by an extended anti-parallel β -strand).

Arthrobacter sp 4-HBT has an order of secondary structure of $\alpha 1$ - $\beta 1$ - $\beta 2$ - $\alpha 2$ - $\alpha 3$ - $\beta 3$ - $\beta 4$ - $\beta 5$ - $\beta 6$ from N- to C- terminus, same as with the hotdog desaturase, FabZ (Figure

1.9 B). PaaI, on the other hand, has high similarity in tertiary fold to 4-HBT. It has not only the same number and order of secondary structural elements, but also a comparable orientation of these elements to form a tertiary structure. For both of the two hotdog thioesterases, the central α -helix (α 3 helix) is wrapped by a β -sheet formed by 6 anti-parallel β -strands with an order 1-2-4-5-6-3, oriented from left to right. One of the obvious differences between structures of two hotdog members is the length of their respective α 1 helix. In the case of 4-HBT, this helix, following the N-terminal irregular loop, has total of 2 turns with a twist in between turns. PaaI, however, has α 1 helix starting from N-terminus and extending for 6 turns. The other notable difference is the β 3 strand, which in PaaI is significantly shorter.

If the similarities in the tertiary structures of 4-HBT and PaaI can be explained by their similar substrate targets (hydroxylated benzoyl-CoA for 4-HBT and hydroxylated phenylacetyl-CoA), the similarities between FIK and CalE7 in tertiary structures are beyond a simple explanation. Here the order of secondary structural elements is identical: β 1- α 1- α 2- α 3- β 2- β 4- β 4- α 4. However, close examination reveals the notable differences of the orientation of these structural elements. The central helix of FIK, α 3-helix, is wrapped by β -strand 5, instead of β -strand 6, whereas the central helix of CalE7 is α 2. The α 3-helix of CalE7 originates from the 90° twist at the C-terminal end of the α 2. The source of a second significant difference between the two tertiary structures is the different orientation of their C-terminal α 4 helices.

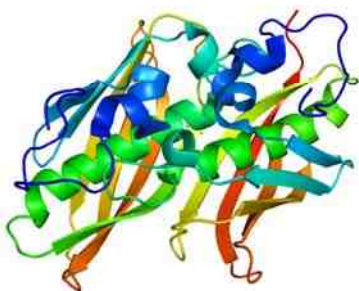
At this point, a brief summary of the versatility of the hotdog thioesterases' tertiary structure is appropriate. Firstly, the tertiary structures from the 4 well-studied

minimal hotdog thioesterases can be divided into two clades, the 4-HBT like clade, which includes 4-HBT and PaaI, and the FIK like clade which includes FIK and CalE7. Secondly, notable variations exist between the two clades. The 4-HBT-like clade has 6 β -strands with an orientation order of 1-2-4-5-6-3. The FIK-like clade has 5 strands with an order of 1-3-4-5-2. In addition, the 4-HBT like clade fold starts with a short helix following a loop or a more extensive helix, whereas the FIK like clade structure forms a β strand at its N-terminus. Furthermore, the FIK like clade has a α 4 helix at the C-terminus, unlike the 4-HBT like clade.

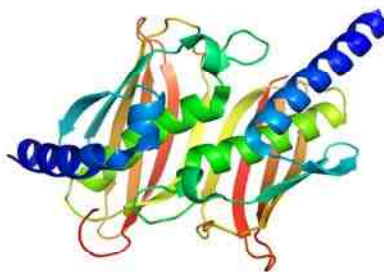
The most minimal hotdog fold shares features with 4-HBT which is the archetype of the 4-HBT-like clade. Members of 4-HBT clade include *Haemophilus influenzae* YciA, *Homo sapiens* THEM2 and THEM4, to name a few. The FIK-like clade includes *Campylobacter jejuni* thioesterase Cj0915 (36) in addition to FIK and CalE7. As will be addressed below, FIK clade members may have diverse ligand binding modes, catalytic sites and reaction mechanisms, compared to the 4-HBT clade.

1.2.2.2 Quaternary Structure

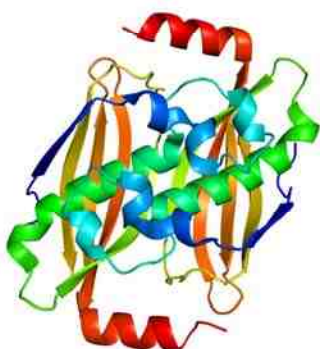
Arthro sp 4-HBT



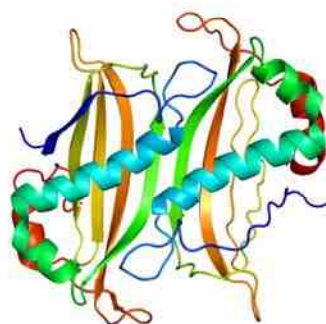
Paal



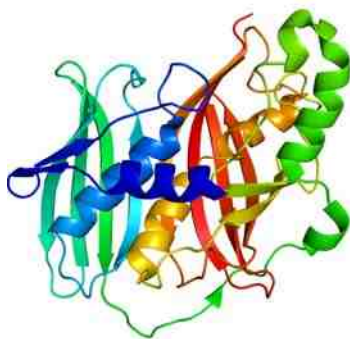
FIK



CalE7



E coli TEII



CACH

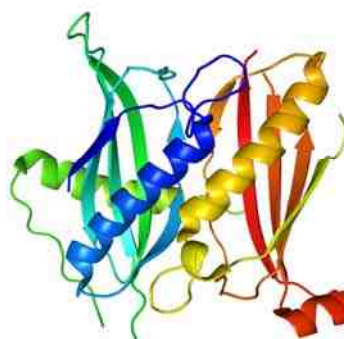


Figure 1.10 Tertiary structures of functional units of *Arthro sp* 4-HBT, Paal, FIK, CalE7, *E coli* TEII and CACH colored in rainbow, ramped from blue (N-terminus) to red (C-terminus).

The quaternary structure of the typical hotdog fold thioesterase consists of a dimer of dimers, with the dimer being the minimal unit that can mediate catalysis. Trimers of dimer (YciA) have been reported, as have simple dimers (hTHEM4). The variations of tertiary structures between the two clades do not significantly alter their shared features in dimerization strategy. 4-HBT, PaaI, FIK and CalE7 dimerizes mainly through interaction between a β -strand from each of the two monomers, and stacking interactions of N-terminal first two turns of central α helix from each monomer (Figure 1.10). *Arthro* 4-HBT monomers appear to pack tighter than those of PaaI, due to the additional interaction between short helices ($\alpha 1$ and $\alpha 2$) from two monomers. A similar fashion of interaction is also observed with the members of the FIK clade.

Some bacterial (*E. coli* TEII for example) and several human hotdog thioesterases contain fused hotdog fold domains which are known as double hotdog (DH) units. In the *E. coli* TEII, two DH monomers associated to form a dimer which closely resembles the homotetramer of the 4HBA clade. The modular human hotdog thioesterases (Acot7 and CACH) possess two DH units within a subunit, which associates to form a dimer. Intrinsically heterogeneous, monomers of CACH and *E. coli* TEII dimers lack the symmetry, although dimerization interactions are reminiscent to those of homodimers (Figure 1.10). In the case of *E. coli* TEII heterodimer, the central α -helix of the N terminal hotdog domain has 4 turns, while that of the C-terminus domain has 7 turns. The dimeric interaction interface involving a 4 helical turn of the two helices is much more extensive, when compared to the typical interface of 2 turns for minimal hotdog thioesterases. Owing to the orientation of the two fused hotdog domains, a long loop

must connect the domains. The helix following this loop of C-terminal hotdog domain is α_1 , which does not orient in the same fashion as that of the N-terminal hotdog domain, thus contributing to the asymmetry.

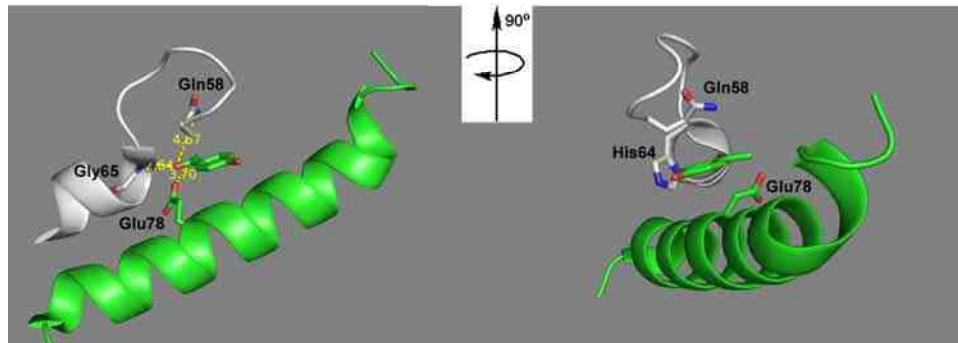
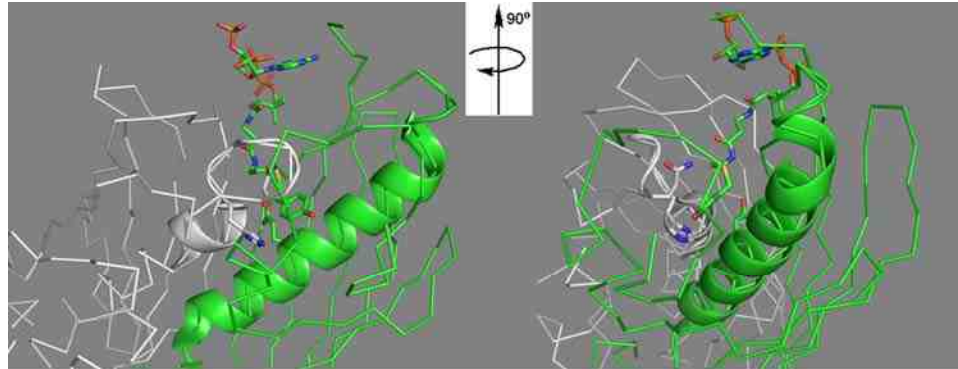
1.2.2.3 Catalytic Mechanisms

The crystal structure of *Arthrobacter sp* 4-HBT, together with detailed kinetic studies unveils its catalytic mechanism. The structure of the enzyme bound with an inert substrate analog reveals that the nucleotide moiety is docked on the top of the entrance to the active site (Figure 1.11 A). As mentioned earlier in this chapter, the phosphopantetheine carrying thioester moiety penetrates deep into enzyme to reach the reaction center, where catalytic residues are contributed mainly by the loop connecting to the N-terminus of the central α -helix of one monomer and the middle of the helix of the other monomer (Figure 1.11 B). The distance between the carboxylate of Glu78 and the carbonyl carbon of the thioester is within the range for nucleophilic attack. On the opposite side of the substrate carbonyl group is Gln58, which may assume a role in stabilizing the tetrahedral intermediate formed during the attack of the Glu78 carboxylate. The amide backbone of Glu65 (located at the N-terminus of the central helix) engages in hydrogen bond interaction with the substrate carbonyl group, rendering it more electrophilic.

The proposed stepwise mechanism of 4-HBT catalyzed hydrolysis of 4-HBA-CoA is as follows: Residues Gln58 and Gly65 polarize carbonyl carbon of the 4-HBA-CoA substrate, which facilitates nucleophilic attack by the carboxylate of Glu78.

Following the release of CoA thiolate from active site, the resulting 4-HBA-enzyme mixed anhydride intermediate putatively undergoes hydrolysis, and the 4-HBA product dissociates freeing the active site for the next turnover reaction.

A



B

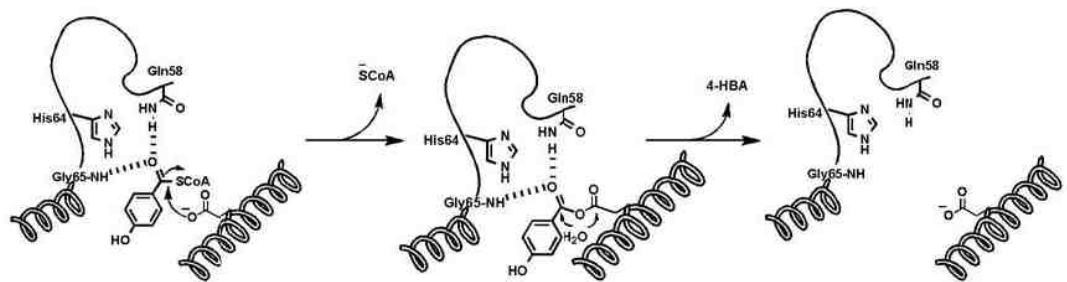
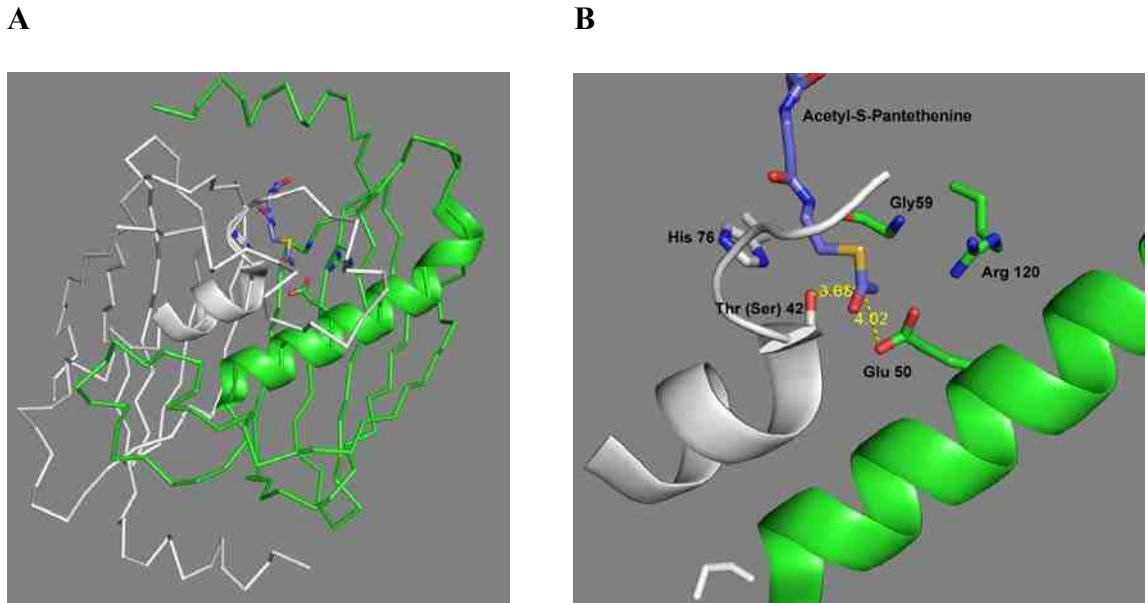


Figure 1.11 Snapshots of the ligand binding site and the catalytic site (A) and the proposed catalytic mechanism (B) of *Arthrobacter sp.* 4-HBT.



C

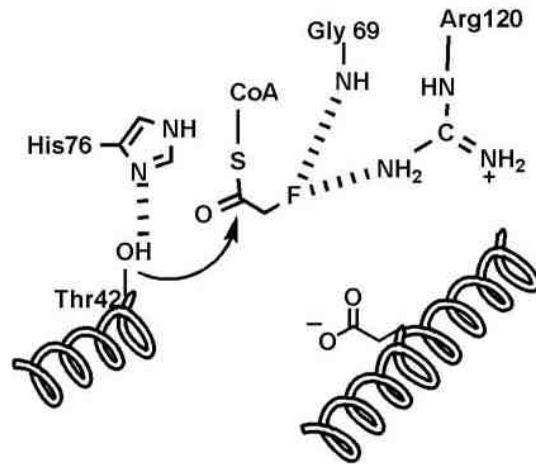


Figure 1.12 Snapshots of the ligand binding site (A), and the catalytic site (B) and the proposed catalytic mechanism (C) of fluoroacetyl-CoA thioesterase.

As described before, based on tertiary and quaternary structures alone, FIK and 4-HBT are grouped in different clades. Notable differences between FIK and other typical hotdog thioesterases revealed by detailed structural and mechanistic investigations support this classification. Based on the reported structure of FIK by Dias *et al*, the dimeric interface of the two FIK monomers forms the substrate binding site bound with acetyl-S-pantetheine. Judging from the orientation of the ligand, the nucleotide moiety is predicted to sit at the entrance of substrate binding site, while the pantetheine arm channels and orients the fluoroacetyl grouping at the reaction center (Figure 1.12 A). Overall, the ligand binding site is the same as those of 4-HBT and its clade-like members.

Residues that fall within a 10 Å range of the substrate thioester moiety are potentially important for catalysis. Unlike 4-HBT, whose catalytic residues are mainly contributed by the loop at the N-terminus of the central helix of one monomer and the middle part of this helix of the other monomer, FIK also utilizes residues from the β -strands of both monomers for catalysis (Figure 1.12 B). The loop at the N-terminus of the central helix contributes Thr42 and His64, whereas the catalytic residues Gly59 and Arg120 are located at the β 2 strand of one monomer and the β 5 of the other monomer. Interestingly, the carboxylate group of the central helix Glu50, is positioned 4.0 Å from the acetyl carbonyl carbon, reminiscent of Glu78 carboxylate nucleophile found in 4-HBT. In contrast to the essential role of the Glu78 in 4-HBT catalysis, mutation Glu50 does not appear to reduce the k_{cat} of FIK, but instead increases the K_m ~10 fold, suggestive of its role in substrate recognition. In contrast, the mutation of Thr42 to alanine inactivates FIK, suggesting a direct involvement of catalysis. The two residues

located at β -sheet, Gly59 and Arg120, are proposed H-bonding with fluoride atom of the substrate, thereby orienting the carbonyl group at close proximity of $-\text{OH}$ of Thr42.

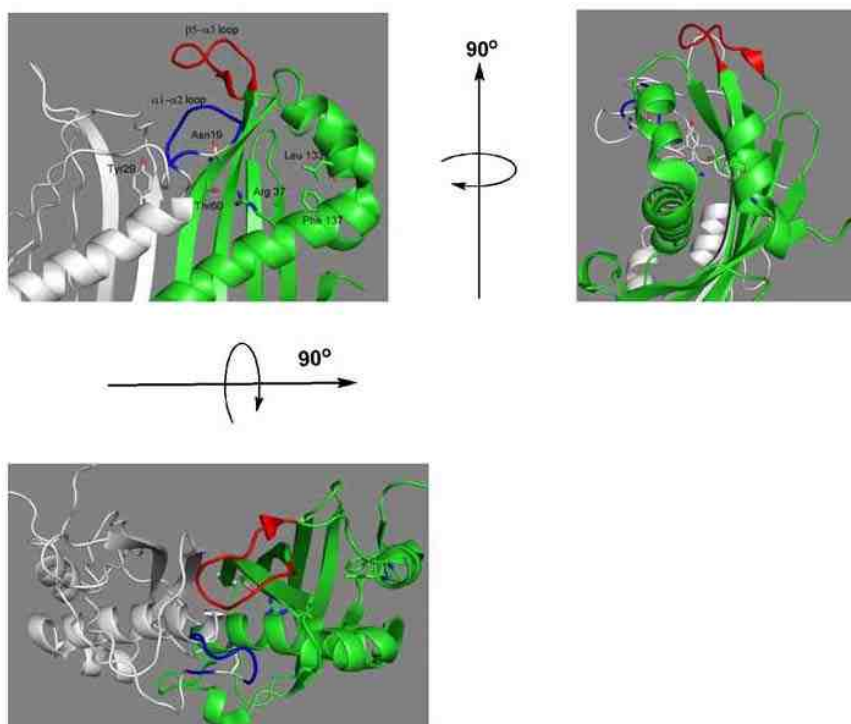
Lines of evidence supports that FIK hydrolyzed fluoroacetyl-CoA in a distinctive mechanism compared to 4-HBT (Figure 1.12 C). The substrate oriented partially by Gly59 and Arg120, and putatively by Glu50, is attacked by the activated $-\text{OH}$ of Thr42 by His 69, and subsequently hydrolyzed.

Although grouped as clade-like member of FIK, CalE7 has unique structural and catalytic features. The apo- structure of CalE7 reported by Liang *et al* possesses two notable features in substrate binding, compared to both 4-HBT and FIK (Figure 1.13 A). First, the entrances to the active sites in both 4-HBT and FIK are located at the top of the dimeric interface and allow the docking of nucleotide moiety of CoA substrate and snaking of pantetheine arm through a channel formed by dimerization of the interfaced subunits. This substrate binding mode is seen in numerous reported crystal structures of hotdog thioesterases. In the case of CalE7, the $\beta 5$ - $\alpha 3$ and $\alpha 1$ - $\alpha 2$ loops appear to prevent this orientation. Lacking a co-crystal structure of CalE7 and its ACP thioester substrate, the authors proposed that ACP orients on CalE7 such that its attached phosphopantetheine arm can snake into active site through the narrow cleft which is formed by the two $\beta 2$ strands and roofed by the $\beta 5$ - $\alpha 3$ and $\alpha 1$ - $\alpha 2$ loops. Secondly, due to this unprecedented orientation of the pantetheine arm in CalE7, the construct of the catalytic site of CalE7, to which polyene carbonyl is directed, is markedly different from the other known hotdog thioesterases. The long chain polyene is mainly accommodated in the narrow hydrophobic space between the central α -helix and the β -strands from one

of the monomers.

Through site-directed mutagenesis, the catalytic relevant residues are identified. CalE7 does not appear to possess a carboxylate residue this is positioned so that it might function as a nucleophile or general base. The proposed catalytic mechanism of CalE7 is as follows: oriented and activated by H-bonding between the Asn19 side chain and the β -carbonyl, and between the Arg37 and the carboxylate carbon, the substrate is subjected to attack by a putative hydroxide anion stabilized through hydrogen bonding to the Tyr29 ring hydroxyl group. The resulting β -keto-polyene carboxylate undergoes spontaneous decarboxylation, generating the precursor of the enediyne product.

A



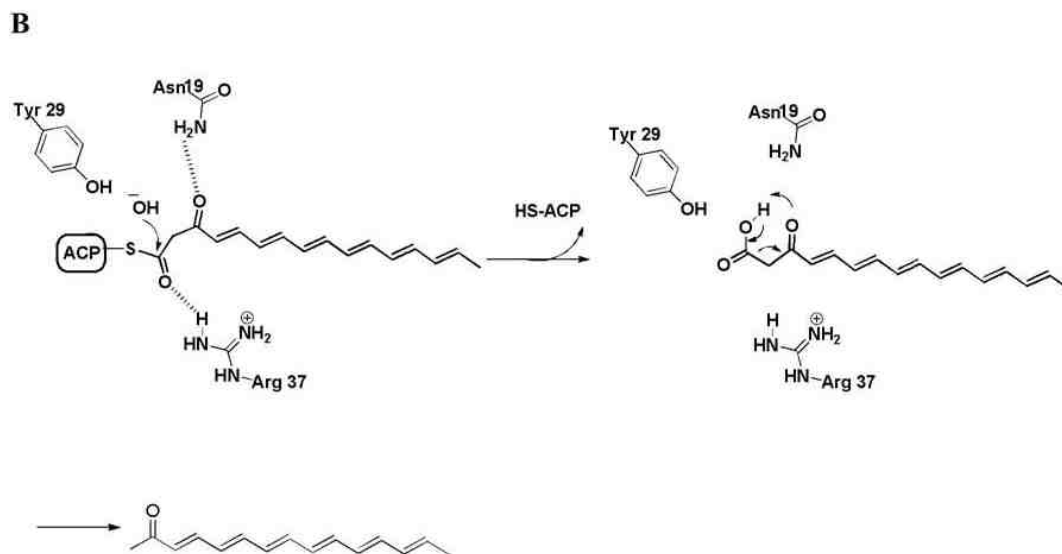


Figure 1.13 Snapshots of the ligand binding site and the catalytic site (A) and the proposed catalytic mechanism (B) of the hotdog thioesterase CalE7.

1.3 Conclusions.

In this chapter, various aspects of the divergence of the structure, function and mechanism among hotdog fold thioesterases were presented. Not only have hotdog thioesterases diverged to participate in a variety of cellular processes, but they utilize different structural elements and mechanisms to carry out the thioester hydrolysis reaction. The large size of the hotdog thioesterase subfamily is indicative of its success in evolution. In this thesis I will provide further insight into the structural determinants of function and mechanism through my study of the *E. coli* hotdog thioesterases EntH (YbdB) and YdiI and the modular human hotdog thioesterase BFIT.

1.4 References

1. Leesong, M., Henderson, B. S., Gillig, J. R., Schwab, J. M., and Smith, J. L. (1996) *Structure* **4**, 253-264
2. Dillon, S. C., and Bateman, A. (2004) *BMC Bioinformatics* **5**, 109
3. Benning, M. M., Wesenberg, G., Liu, R., Taylor, K. L., Dunaway-Mariano, D., and Holden, H. M. (1998) *J Biol Chem* **273**, 33572-33579
4. Thoden, J. B., Holden, H. M., Zhuang, Z., and Dunaway-Mariano, D. (2002) *J Biol Chem* **277**, 27468-27476
5. Zhuang, Z., Song, F., Zhang, W., Taylor, K., Archambault, A., Dunaway-Mariano, D., Dong, J., and Carey, P. R. (2002) *Biochemistry* **41**, 11152-11160
6. Thoden, J. B., Zhuang, Z., Dunaway-Mariano, D., and Holden, H. M. (2003) *J Biol Chem* **278**, 43709-43716
7. Kunishima, N., Asada, Y., Sugahara, M., Ishijima, J., Nodake, Y., Miyano, M., Kuramitsu, S., and Yokoyama, S. (2005) *J Mol Biol* **352**, 212-228
8. Song, F., Zhuang, Z., Finci, L., Dunaway-Mariano, D., Kniewel, R., Buglino, J. A., Solorzano, V., Wu, J., and Lima, C. D. (2006) *J Biol Chem* **281**, 11028-11038
9. Cao, J., Xu, H., Zhao, H., Gong, W., and Dunaway-Mariano, D. (2009) *Biochemistry* **48**, 1293-1304
10. Cheng, Z., Bao, S., Shan, X., Xu, H., and Gong, W. (2006) *Biochem Biophys Res Commun* **350**, 850-853
11. Willis, M. A., Zhuang, Z., Song, F., Howard, A., Dunaway-Mariano, D., and Herzberg, O. (2008) *Biochemistry* **47**, 2797-2805

12. Zhuang, Z., Song, F., Zhao, H., Li, L., Cao, J., Eisenstein, E., Herzberg, O., and Dunaway-Mariano, D. (2008) *Biochemistry* **47**, 2789-2796
13. Zhao, H., Martin, B. M., Bisoffi, M., and Dunaway-Mariano, D. (2009) *Biochemistry* **48**, 5507-5509
14. Dias, M. V., Huang, F., Chirgadze, D. Y., Tosin, M., Spitteller, D., Dry, E. F., Leadlay, P. F., Spencer, J. B., and Blundell, T. L. (2010) *J Biol Chem*
15. Liu, L. X., Margottin, F., Le Gall, S., Schwartz, O., Selig, L., Benarous, R., and Benichou, S. (1997) *J Biol Chem* **272**, 13779-13785
16. Kotaka, M., Kong, R., Qureshi, I., Ho, Q. S., Sun, H., Liew, C. W., Goh, L. P., Cheung, P., Mu, Y., Lescar, J., and Liang, Z. X. (2009) *J Biol Chem* **284**, 15739-15749
17. Dunaway-Mariano, D., and Babbitt, P. C. (1994) *Biodegradation* **5**, 259-276
18. Burlingame, R. P., Wyman, L., and Chapman, P. J. (1986) *J Bacteriol* **168**, 55-64
19. Ferrandez, A., Minambres, B., Garcia, B., Olivera, E. R., Luengo, J. M., Garcia, J. L., and Diaz, E. (1998) *J Biol Chem* **273**, 25974-25986
20. Ismail, W., El-Said Mohamed, M., Wanner, B. L., Datsenko, K. A., Eisenreich, W., Rohdich, F., Bacher, A., and Fuchs, G. (2003) *Eur J Biochem* **270**, 3047-3054
21. Ahlert, J., Shepard, E., Lomovskaya, N., Zazopoulos, E., Staffa, A., Bachmann, B. O., Huang, K., Fonstein, L., Czisny, A., Whitwam, R. E., Farnet, C. M., and Thorson, J. S. (2002) *Science* **297**, 1173-1176
22. Kong, R., Goh, L. P., Liew, C. W., Ho, Q. S., Murugan, E., Li, B., Tang, K., and Liang, Z. X. (2008) *J Am Chem Soc* **130**, 8142-8143

23. Harper, D. B., and O'Hagan, D. (1994) *Nat Prod Rep* **11**, 123-133
24. Huang, F., Haydock, S. F., Spitteller, D., Mironenko, T., Li, T. L., O'Hagan, D., Leadlay, P. F., and Spencer, J. B. (2006) *Chem Biol* **13**, 475-484
25. Cobb, S. L., Deng, H., Hamilton, J. T., McGlinchey, R. P., and O'Hagan, D. (2004) *Chem Commun (Camb)*, 592-593
26. Aiken, C., Krause, L., Chen, Y. L., and Trono, D. (1996) *Virology* **217**, 293-300
27. Kestler, H. W., 3rd, Ringler, D. J., Mori, K., Panicali, D. L., Sehgal, P. K., Daniel, M. D., and Desrosiers, R. C. (1991) *Cell* **65**, 651-662
28. Garcia, J. V., Alfano, J., and Miller, A. D. (1993) *J Virol* **67**, 1511-1516
29. Rhee, S. S., and Marsh, J. W. (1994) *J Virol* **68**, 5156-5163
30. Prass, R. L., Isohashi, F., and Utter, M. F. (1980) *J Biol Chem* **255**, 5215-5223
31. Suematsu, N., Okamoto, K., Shibata, K., Nakanishi, Y., and Isohashi, F. (2001) *Eur J Biochem* **268**, 2700-2709
32. Suematsu, N., Okamoto, K., and Isohashi, F. (2003) *J Chromatogr B Analyt Technol Biomed Life Sci* **790**, 239-244
33. Alpy, F., and Tomasetto, C. (2005) *J Cell Sci* **118**, 2791-2801
34. Adams, S. H., Chui, C., Schilbach, S. L., Yu, X. X., Goddard, A. D., Grimaldi, J. C., Lee, J., Dowd, P., Colman, S., and Lewin, D. A. (2001) *Biochem J* **360**, 135-142
35. Matsunaga, T., Isohashi, F., Nakanishi, Y., and Sakamoto, Y. (1985) *Eur J Biochem* **152**, 331-336
36. Yokoyama, T., Choi, K. J., Bosch, A. M., and Yeo, H. J. (2009) *Biochim Biophys*

Acta **1794**, 1073-1081

CHAPTER 2

IN VITRO KINETIC ANALYSIS OF SUBSTRATE SPECIFICITY IN ENTEROBACTIN BIOSYNTHETIC LOWER PATHWAY ENZYMES PROVIDES INSIGHT INTO THE BIOCHEMICAL FUNCTION OF ENTH

2.1 Introduction

Iron is essential to nearly all living organisms. During an iron depleting period, microorganisms, including bacteria and fungi, release siderophores which are high affinity iron-chelating compounds, to harvest exogenous iron (1-2). Because pathogenic strains must sequester iron from the human host, mechanistic studies on biosynthesis of siderophores would lead to development of novel antibiotics (3). Enterobactin, synthesized by *Escherichia coli* is among the most thoroughly studied siderophores (4).

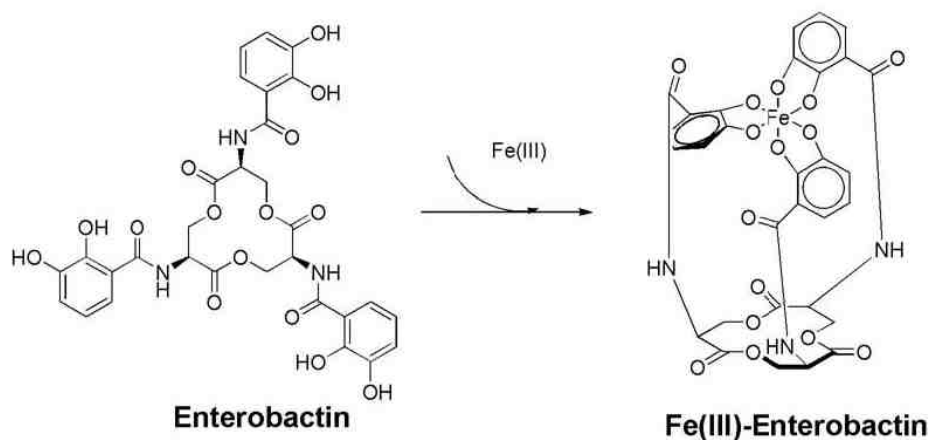


Figure 2.1 Iron-chelating siderophore enterobactin and Fe (III) bound enterobactin

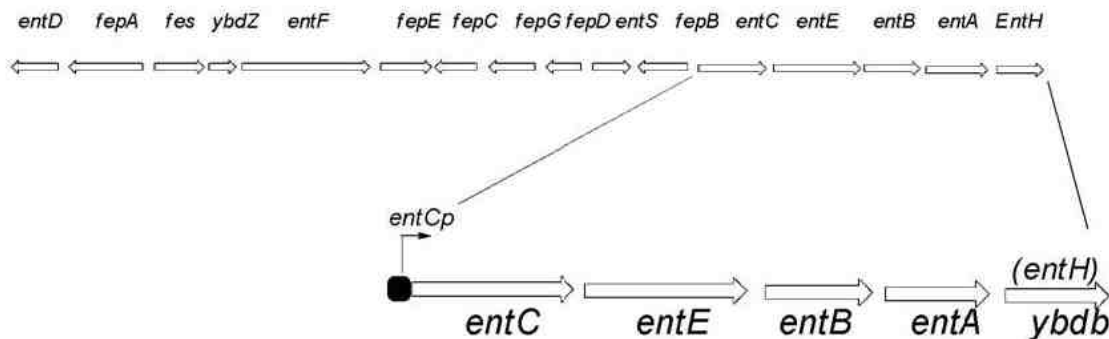


Figure 2.2 Enterobactin biosynthesis gene cluster with transcription unit with *entCEBA* and *ybdb* (*EntH*) genes that are regulated by *entCp* promoter region.

The entire gene collection for biosynthesis and transportation of enterobactin is clustered in *E. coli* and regulated by the ferric uptake regulator (Fur), a transcriptional repressor (Figure 2.1) that represses iron responsive gene transcription in iron non-limiting growth conditions (5). One of the promoter-operator regions regulated by Fur in enterobactin biosynthesis gene cluster is the *entCp* promoter located upstream of gene *entC*. Thus, *entCEBA* genes and *ybdb* (recently renamed as *entH*) are colocalized within a single transcription unit, suggesting the participation of EntH in the biosynthesis of enterobactin.

However, studies on the in vitro reconstitution of enterobactin biosynthesis showed the EntH protein is not required (6). Biosynthetic steps to enterobactin proceed via an upper and lower pathway. The upper chemical pathway forms 2,3-dihydroxybenzoate (2,3-DHB) from chorismate, catalyzed cooperatively by EntC, EntB and EntA. The lower pathway (Figure 2.2) employs the two nonribosomal peptide synthetases (NRPs) EntB and EntF (7-8). The phosphopantetheinyl transferase (PPTase)

EntD catalyzes the transfer of pantetheinephosphate from CoA to the Ser residue of the 2,3-DHB carrier domain (ArCP) of EntB and of the peptidyl carrier domain (PCP) of EntF to generate holo-EntB and holo-EntF. EntE catalyzes the adenylation of 2,3-DHB and the subsequent acyl transfer to the pantetheine thiol of holo-EntB. The adenylation domain (A) of holo-EntF catalyzes the adenylation of L-Ser and the subsequent transfer of an acyl group to the pantetheine thiol carrier domain of EntF. The EntF condensation domain then catalyzes the transfer of the 2,3-DHB unit to the amine group of the EntF-tethered L-Ser. Next, the 2,3-DHB-Ser unit is transferred to the active site Ser of the EntF cyclizing thioesterase domain (TE), thereby freeing the pantetheine thiol for the acquisition of a second 2,3-DHB-Ser unit. It is in turn transferred to the Ser moiety side chain of 2,3-DHB-Ser tethered to the TE domain. Following the addition of a third 2,3-DHB-Ser unit, the chain is cyclized by the TE domain to form the tri-N-2,3-DHB-Ser-lactone, Enterobactin.

Through sequence analysis of proteins in Pfam database and X-ray crystallization, EntH is identified as one member of the “hot dog” fold thioesterase family (9-10). It is not unusual for a biosynthetic pathway gene cluster leading to nonribosomal peptide (NRP) to possess a stand-alone thioesterase. Previous studies have provided evidence of these thioesterases function as “housekeepers” by catalyzing the removal of non-native or aberrant acyl extending units from the pantetheine thiols of misacylated NRPs (11-13). We suspected that EntH might play such a role in enterobactin biosynthesis even though the NRP housekeepers characterized to date belong to the α/β hydrolase fold family. Because of its “hot dog” fold, EntH is unique and with the exception of the double hotdog fold thioesterase TESII, hot dog thioesterases characterized to date target acyl-

CoA substrates. On the other hand, the EntH paralog YbgC was discovered by Bouveret and co-workers to bind with the acyl carrier protein of the fatty acid synthesis pathway (14). Although thioesterase activity towards the fatty acyl-ACP has yet to test to confirm this partnership, we were nevertheless encouraged about the possible biochemical function of EntH as an acyl-ACP thioesterase which targets misacylated EntB and/or EntF.

During the course of my investigation of EntH substrate specificity, Bouveret and co-workers reported the results of an *in vivo* two-hybrid experiment and an *in vitro* pull down experiment, which suggested binding interaction between EntH and holo-EntB (15). These investigators proposed a proofreading role for EntH in which holo-EntB acylated with salicylic acid produced by competing bacteria, is hydrolyzed. While, this hypothesis was similar to our own in so far as misacylated EntB is the suggested physiological substrate, the experimental findings would in the end serve to complement our *in vitro* activity analyses and strengthen the case for the house keeping function of EntH.

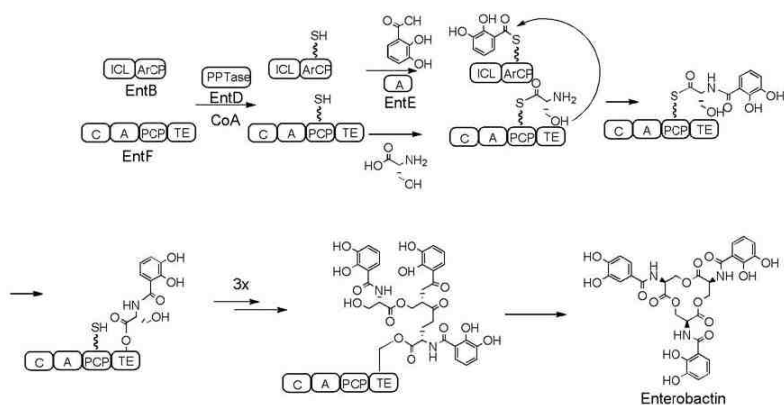


Figure 2.3 Lower chemical pathway for biosynthesis of enterobactin.

As part of a larger effort to define structural determinants of substrate recognition within the hotdog-fold thioesterase family, we have targeted EntH for structure-function analysis. Little insight into the structure the physiological substrate can be gleaned from the structure of unliganded *E. coli* EntH (PDB ID 1vh9; Structural Genomic 2003).

In this chapter, through *in vitro* kinetic analysis we demonstrate two potential routes to EntB misacylation and provide evidence for two mechanisms by which the hotdog-fold thioesterase EntH can potentially prevent or reverse EntB misacylation.

2.2 Experimental

2.2.1 Materials.

2,3-Dihydroxyl- (2,3-HB), 2,4-dihydroxyl- (2,4-HB), 2-hydroxyl-(2-HB), 3-hydroxyl- (3-HB), 4-hydroxylbenzoate (4-HB) and pyridyl were purchased from Sigma-Aldrich as were the acyl-CoAs and CoA. Plasmids for expression of EntE and EntF were kindly provided by Dr. Andrew Gulick and Dr. Chris Walsh, respectively. 3-HB-CoA, 4-HB-CoA, benzoyl-CoA, 3-hydroxyl-phenylacetyl-CoA (3-HP-CoA) were synthesized according to published procedure (16-17). Wild type BW2513, $\Delta entH$ /BW2513 and $\Delta entB$ /BW2513 strains were obtained from Keio knockout collection.

2.2.2 Growth monitoring of type BW2513, $\Delta entH$ /BW2513 and $\Delta entB$ /BW2513 strains under iron-rich or iron-depleting conditions.

Wild type BW2513, $\Delta entH$ /BW2513 and $\Delta entB$ /BW2513 strains were cultured in dish culture at 37°C for overnight before culturing in 10 mL LB medium in the absence of antibiotic. Larger scale cultures in 2 L LB medium with (iron depleting) or without

(iron rich) 400 μ M dipyriddy were initiated by introducing equal amount of cells from the 10 mL small scale cultures of the above-mentioned strains. Cell density was monitored at 600 nm for up to 12 hrs starting from the initiation of the large scale cultures.

2.2.3 Electrospray Ionization Mass Spectroscopy.

ESI-MS determination of molecular weight was carried out by the mass spectrometry facility of the University of New Mexico. Protein samples were diluted in 10% acetonitrile, 90% water and 0.1% of formic acid and loaded into a NewObjective Pico tip. The solution was infused into an ion trap mass spectrometer (Finnigan LCQ DUO). The distance between the spray needle and counter electrode was 3 mm. A voltage of 7.0 kV was applied at the spray needle.

2.2.4 *EntH* (*ybdB*) Cloning and Expression and EntH Purification.

The gene encoding the thioesterase EntH was cloned by using a PCR based strategy in which genomic DNA prepared from *E. coli* strain K12 (substrain W3110) was used as template, commercial oligonucleotides as primers, *Pfu* as the DNA polymerase. The *NdeI* and *XhoI* treated PCR product of *entH* was ligated by T4 DNA ligase to *NdeI/XhoI* sites of pET-23a (Novabiochem) to give the plasmid, *entH/pET-23a*. The resulting clone was use to transform competent *E. coli* BL21 (DE3) Star cells (Novagen). The cells were grown at 37°C in LB medium containing 50 μ g/mL ampicillin. The cell culture was induced with 0.4 mM isopropyl- β -D-galactopyranoside (IPTG) when the cell optical density had reached 0.6 (OD at 600nm). Following an overnight induction, the cells were harvested by centrifugation at 5,000 g for 15 min and then resuspended in 150

mL of lysis buffer (20 mM Tris•HCl, 10 mM imidazole, 1 mM DTT and 300 mM NaCl, pH 7.5). The cell suspension was lysed by passing through a French pressure cell at 1,200 psi before centrifugation at 48,000 G for 1 h at 4°C. The supernatant was loaded onto a Ni-NTA column (2 cm X 11cm) at a flow rate of 0.5 mL/min. The column was washed with 150 mL of lysis buffer followed by 100 mL of wash buffer (20 mM Tris•HCl, 300 mM NaCl, 50 mM imidazole and 1mM DTT at pH 7.5). EntH was eluted by applying 90 mL of elution buffer (20 mM Tris •HCl, 300 mM NaCl, 250 mM imidazole and 1mM DTT at pH 7.5). The thioesterase containing fractions were analyzed by SDS-PAGE before concentrating for storage in at -80°C.

2.2.5 Preparation of Recombinant Apo-EntB, EntD, EntE and Apo-EntF.

These proteins were produced and purified according to published methods (6,18). Both EntB and EntD possess a C-terminal His₆ tag, whereas EntE and EntF possess a N-terminal His₆ tag.

2.2.6 EntD Catalyzed Formation of Misacylated-holo-EntB with Aroyl- and Acyl-CoAs.

The acyl-CoAs used as substrates included lauroyl-, palmitoyl-, hexanoyl-, isobutyryl- and acetyl-CoA. The misacylated EntBs were generated by incubation of the following components in a 500 µL volume at 37°C for 30 min: 5 mM DTT, 50mM HEPES (pH 7.5) with 50 mM NaCl, 100 µM of one of the above-mentioned acyl-CoA, 25 µM apo-EntB-His₆ and 2 µM EntD. The resulting reaction mixtures were subject to ESI-MS analysis.

2.2.7 Formation of Misacylated-holo-EntB by EntE Catalyzed Loading of Benzoic Acid Derivatives onto Holo-EntB.

2,3-DHB, 2,4-DHB, 2-HB, 3-HB, 4-HB and benzoyl- loaded holo-EntB were synthesized from apo-EntB in two steps. In the first step, a 500 μL solution of 125 μM apo-EntB, 400 μM CoA and 2 μM EntD, 50 mM HEPES (pH 7.5), 1 mM MgCl_2 , 5 mM DTT was incubated for 2 h at room temperature. In the second step reaction, an equal volume of 1 mM ATP and 0.5 mM desired benzoate and 1 μM EntE was added to the reaction and following a 1 h incubation period the reaction solution was flash frozen for storage at -80°C until ESI-MS analysis was carried-out.

2.2.8 Steady-state Kinetic Constant Determinations for EntH Thioesterase-catalyzed Hydrolysis of Acyl-CoA and Acyl-holo-EntB.

k_{cat} and K_m determinations were carried-out by using a DTNB [5, 5-dithiobis (2-nitrobenzoic acid)]-based continuous spectrophotometric assay. Hydrolysis reactions of the acyl-thioester substrates were monitored at 25°C by measuring the absorbance of 5-thio-2-nitrobenzoate at 412 nm (ϵ 13.6 $\text{mM}^{-1}\text{cm}^{-1}$), formed by reaction of DTNB with the product, free CoA or holo-EntB. Reaction mixtures (500 μl) contained EntH, acyl-CoA or acyl-holo-EntB substrate (0.5 to 5 K_m), DTNB (1 mM), KCl (0.2 M), and 50 mM HEPES (pH 7.5). The kinetic parameters of V_{max} and K_m were determined from initial velocity data, as measured as a function of substrate concentration, by using eq 1 and computer program KinetAsyst (IntelliKinetics)

$$V=V_{\text{max}}[A]/([A]+K_m) \quad (1)$$

Where $[A]$ is the substrate concentration, V is the initial velocity, V_{\max} is the maximum velocity, and K_m is the Michaelis constant. The reported error was computed for the data fitting. The k_{cat} was calculated from the ratio of V_{\max} to the total enzyme concentration. The enzyme concentration was determined by using the Bradford method.

2.2.9 HPLC Analysis of Formation of 2,4-DHB by Hydrolysis of 2,4-DHB-holo-EntB by EntH.

The 500 μL reaction solution initially contained 50 mM HEPES (pH 7.5), 50 mM NaCl, 250 μM 2, 4-DHB-holo-EntB and 2 μM EntH at 25°C. Aliquots of 100 μL were extracted at 2 min, 8 min and 25 min respectively, and passed through a Amicon YM-10 (Millipore) filter. The control sample did not contain EntH. The filtrates were analyzed by reversed phase HPLC using a C18 reverse phase column (5 μm , 4.6 mm \times 25 cm) and a linear gradient of 0.1% trifluoroacetic acid in deionized water to 0.1% trifluoroacetic acid in 50% acetonitrile. 2,4-DHB has a retention time of 13.9 min.

2.2.10 ESI-MS Analysis of the Hydrolysis of 2,4-DHB-holo-EntB by EntH.

Aliquots (100 μL) were taken from a reaction solution at 25°C (500 μL) initially containing 1mM DTT, 50 mM HEPES buffer (pH 8.5), 50 mM NaCl, 98 μM 2,4-DHB-holo-EntB and 2 μM of EntH at 5 min, 10 min and 25 min incubation, flash frozen and then stored at -80°C until ESI-MS analysis. Alternatively, reaction solutions initially containing 1 mM DTT, 50 mM HEPES buffer (pH 8.5), 50 mM NaCl, 150 μM 2,4-DHB-holo-EntB and 0.5 μM , 2 μM or 8 μM EntH were incubated at 25°C for 15min before freezing and storage at -80°C until ESI-MS analysis.

2.2.11 ESI-MS Analysis of the Processivity of Acylated- or Aroylated-holo-EntB by EntF.

2,3-DHB-, 2,4-DHB- and lauroyl-holo-EntB were prepared as described above and then separately incubated at a concentration of 40 μ M with 20 μ M holo-EntF or apo-EntF at 25°C in 1 mM DTT, 50 mM HEPES (pH 8.5), 50 mM NaCl, 1 mM MgCl₂, 1 mM L-Serine and 2 mM ATP. Aliquots (100 μ L) of the reaction mixture were extracted at 5 min, 15 min and 25 min respectively, flash frozen and stored at -80°C until ESI-MS analysis. The 2,4-DHB-holo-EntB control reaction lacked L-Serine and ATP.

2.3 Results and Discussion

2.3.1 Correlation of YbdB (EntH) Function and Enterobactin Biosynthesis.

Our first approach to establish the direct connection between EntH and Enterobactin biosynthesis is through gene knockout and subsequent phenotype analysis. *E. coli* strains including wild type BW2513, Δ *entH*/BW2513 and Δ *entB*/BW2513 were obtained from Keio gene collection and used to investigate the functional role of *ybdB* (*entH*) gene product. The growth of these strains under iron rich conditions is almost identical (Figure 2.4). Under iron depleting conditions the growth rates of all three strains were inhibited. However, the growth rates of the Δ *entH*/BW2513 and Δ *entB*/BW2513 strains more so than that of the wild type BW2513 strain. The difference in growth curves between Δ *entB*/BW2513 and wild type BW2513 strains is readily explained by the essential role of EntB in enterobactin biosynthesis. Removal of the *entB* gene abolishes the capability of the Δ *entB*/BW2513 mutant to produce enterobactin during the

iron-depleting period and hence, generates a phenotype of further inhibition in growth. Notably, a similar phenotype is also observed for $\Delta entH$ /BW2513. This finding suggested that, EntH also plays an essential role in the biosynthesis of enterobactin.

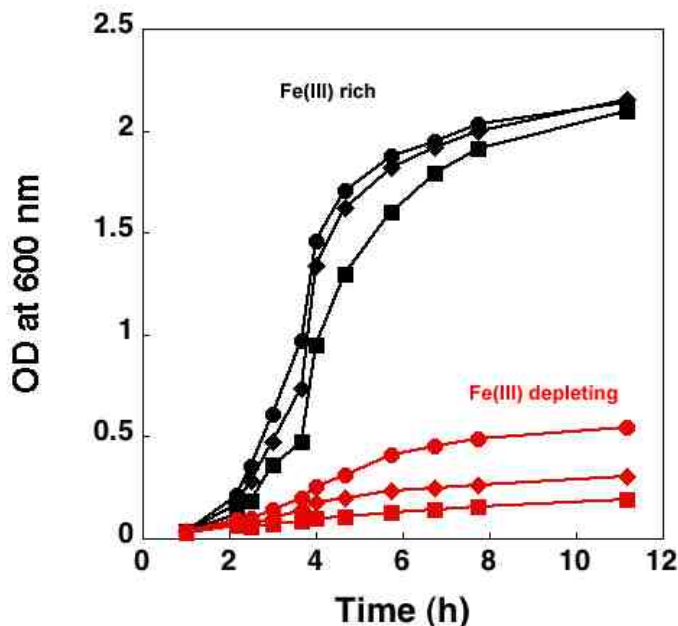


Figure 2.4 Growth curves of wild type BW2513, $\Delta entH$ /BW2513 and $\Delta entB$ /BW2513 in iron-rich or iron-depleted LB medium: (●) wild type BW2513, (◆) $\Delta entB$ /BW2513 and (■) $\Delta entH$ /BW2513 grown in the absence of pyridyl (iron rich medium); (●) wild type BW2513, (◆) $\Delta entB$ /BW2513 and (■) $\Delta entH$ /BW2513 grown in the presence of 400 μ M pyridyl (iron-depleted conditions). Growth was monitored at 600 nm for up to 12 h following the inoculation of the respective culture flasks.

2.3.2 Expression and Purification of EntH and Enzymes of the Lower Pathway of Enterobactin Pathway.

Enzymes of the lower enterobactin pathway include EntB, EntE and EntF. Recombinant EntB features a C-terminal His₆ tag, whereas EntE and EntF are labeled

with N-terminal His₆ tag. Together with the “hot dog” thioesterase EntH, all the recombinant proteins under investigation were purified to homogeneity as judged from SDS-PAGE analysis (data not shown). High resolution ESI-MS analysis confirmed their identities as the target proteins (Figure 2.5 and Table 2.1). Notably, recombinant EntB and EntF from *E. coli* are exclusively in the apo- form. The holo- (pantetheinylated) forms for either protein were not detected by the ESI-MS spectral analysis.

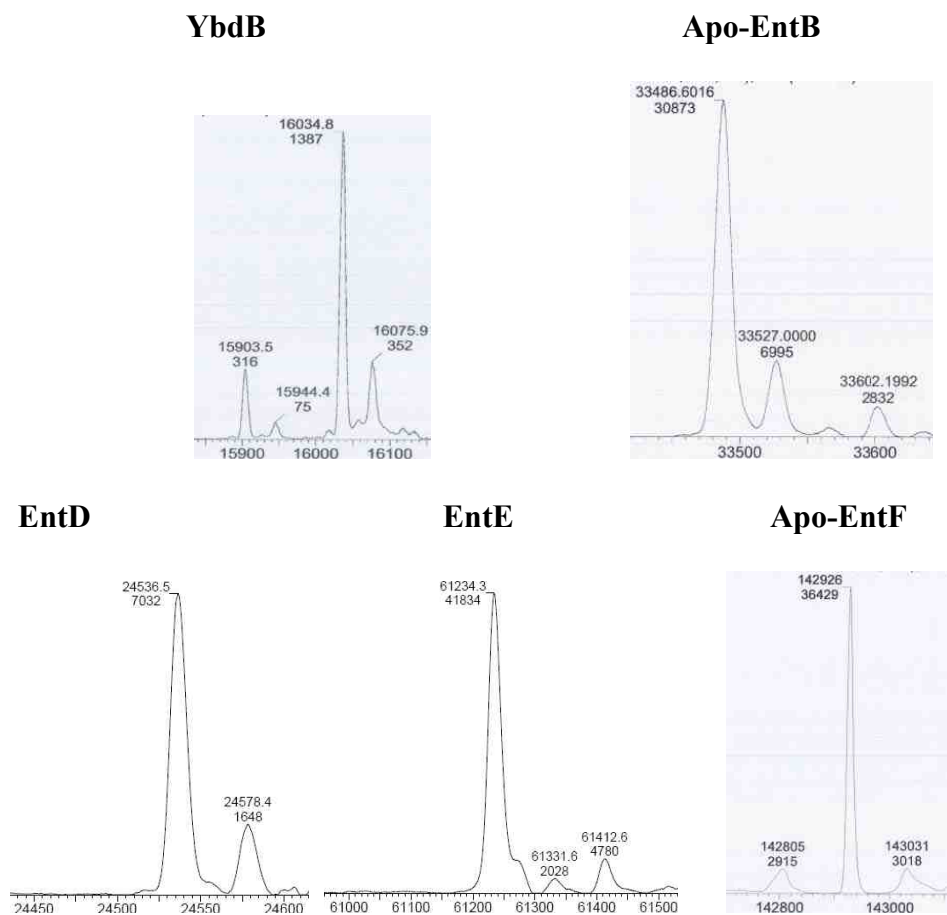


Figure 2.5 Exact molecular weight determination of EntH (YbdB) and enzymes in the lower pathway of enterobactin biosynthesis.

Table 2.1 Exact molecular weight determination of His₆-EntH and His₆-enzymes of the lower Enterobactin pathway.

Enzymes	Predicted M.W. (Da)	Observed M.W (Da)
EntH	16,035.6	16,034.8
EntB	33,619.4	33,486.6 (-Met)
EntE	61,234.3	61,357.6 (-Met)
EntD	24,536.5	24,667.8 (-Met)
EntF	143,057.0	142,926 (-Met)

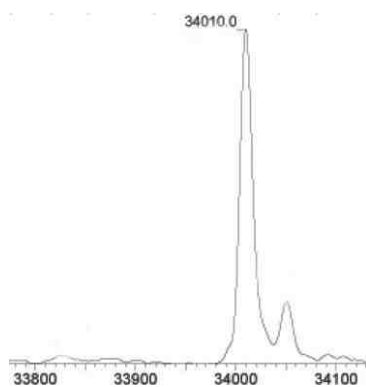
2.3.3 Misacylation of Holo-EntB by EntD and Acyl-CoA.

Building on the hypothesis that EntH functions to release holo-EntB from misacylated holo-EntBs, we first identified the potential sources of misacylated holo-EntBs. One of the potential routes for the formation of misacylated holo-EntB arises from EntD substrate promiscuity. Specifically, PPTases are known to catalyze the transfer of an acyl-pantetheine phosphate unit from an acyl-CoA to the protein acceptor. If CoA is scarce, an endogenous acyl-CoA might substitute as a substrate for EntD, the PPTase in enterobactin biosynthesis.

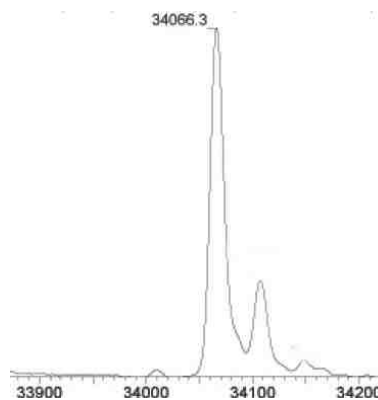
The ability of EntD to utilize CoA, acyl and aroyl-CoA as the pantetheine phosphate donor was tested by incubating reaction solutions initially containing 100 μ M HS-CoA, isobutyryl-CoA, hexanoyl-CoA, lauroyl-CoA, palmitoyl-CoA or 4-hydroxybenzoyl-CoA, 2 μ M EntD, 25 μ M EntB, 50 mM HEPES, 5 mM DTT, and 50 mM NaCl (pH 7.5 and 37°C) for 30 min followed by ESI-MS analysis. The percent

conversion of EntB to acylated holo-EntB was determined by calculating the ratio of the corresponding peak areas in the mass spectrum (Figure 2.6) to show 100% conversion to acyl-holo-EntB for reactions of HS-CoA, lauroyl-CoA, palmitoyl-CoA, and 4-hydroxybenzoyl-CoA and 23 and 67% conversion for reactions of isobutyryl-CoA and hexanoyl-CoA respectively (Table 2.2). The results show that EntD is promiscuous and suggest a potential source of misacylated holo-EntB *in vivo*.

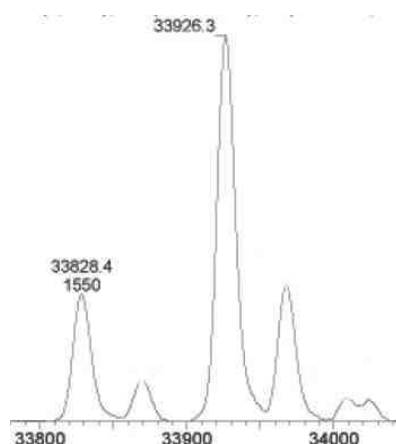
A Lauroyl-holo-EntB



B Palmitoyl-holo-EntB



C Hexanoyl-holo-EntB



D Isobutyl-holo-EntB

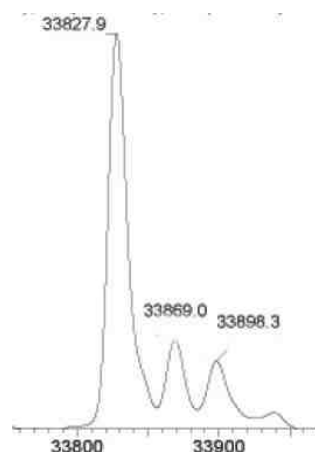


Figure 2.6 ESI-MS spectra of the reaction mixtures formed in the reaction of EntD with **A)** lauroyl-holo-EntB; **B)** palmitoyl-holo-EntB; **C)** hexanoyl-holo-EntB; **D)** isobutyryl-holo-EntB.

Table 2.2 Summary of ESI mass spectral data obtained for the formation of misacylated holo-EntB by EntD catalyzed reactions between apo-EntB various acyl-CoAs.

Target Protein	Predicted M.W. (Da)	Observed M.W. (Da)	Yield of acylated-S-holo-EntB ^a
lauroyl-holo-EntB	34,011.0	34,010.0	100%
palmitoyl-holo-EntB	34,066.9	34,066.3	100%
hexanoyl-holo-EntB	33,927.0	33,926.3	67.7%
isobutyl-holo-EntB	33,898.9	33,898.3	23.3%
4-HBA-holo-EntB	33,948.9	33,948.4	100%

^a Calculated from: (peak area of acylated EntB/ peak area of apo-EntB + peak area of acylated holo-EntB) × 100%.

2.3.4 EntE Misacylation of Holo-EntB with Aromatic Acids Plus ATP.

EntE catalyzes the formation of 2,3-DHB-holo-EntB by first activating 2,3-DHB ATP adenylation and then transferring the 2,3-DHB unit to pantetheinyl thiol of holo-EntB. If 2,3-DHB is scarce, an endogenous carboxylate metabolite might serve as a substitute substrate for EntE, resulting in the formation of misacylated holo-EntB. Thus, the specificity of EntE toward aromatic acid metabolites was next tested by incubating reaction solutions initially containing 125 μM holo-EntB (generated *in situ* using 400 μM CoA and 2 μM EntD), 1 μM EntE, 500 μM 2,3-DHB, 2,4-DHB, 2-hydroxybenzoate (2-

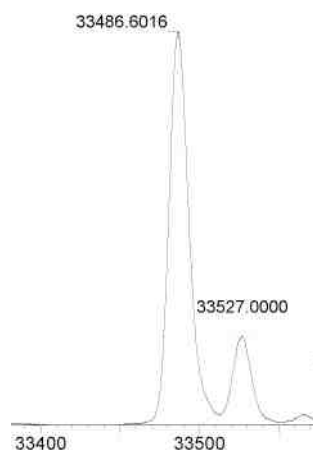
HB), 3-hydroxybenzoate (3-HB), 4-hydroxybenzoate (4-HB), benzoate, phenylacetate, 4-hydroxyphenylacetate (4-HPA) or 3-hydroxyphenylacetate (3-HPA) in 50 mM HEPES (pH 7.5), 1 mM ATP, 10 mM MgCl₂ and 5 mM DTT for 2 h at 25 °C. The percent conversion of holo-EntB to acylated holo-EntB was determined by ESI-MS analysis (Table 2.3 and Figure 2.7) to be 100% for reactions of 2,3-DHB, 2,4-DHB and 47, 25, 9, 14, 0% conversions for reactions of 4-hydroxybenzoate, 3-hydroxybenzoate, 2-hydroxybenzoate, benzoate, and phenylacetate/hydroxyphenylacetate, respectively. These results demonstrate that EntE is promiscuous in acylating holo-EntB with benzoic acids and suggest a second possible source of misacylated holo-EntB *in vivo*.

Table 2.3 Summary of the ESI mass spectral data measured for the acyl-holo-EntB products formed from the reaction of holo-EntB, ATP and the corresponding aromatic acid in the presence of EntE and the formation of the misacylated holo-EntBs via thioesterification of benzoic acid derivatives to ArCP domain catalyzed by EntE.

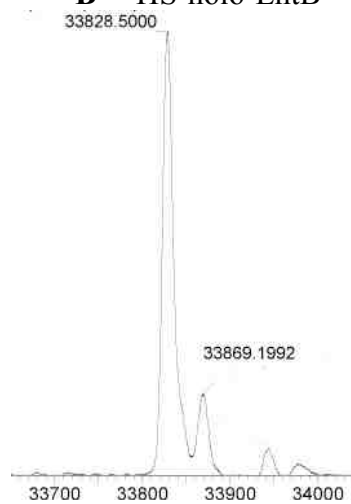
Target Protein	Predicted M.W.(Da)	Observed M.W. (Da)	Yield of acyl- holo-EntB ^a
apo-EntB	33,486.6 (-Met)	33,486.6	---
holo-EntB	33,827.9	33,828.5	0
2,3-DHB-holo-EntB	33,964.9	33,964.0	100%
2,4-DHB-holo-EntB	33,964.9	33,963.8	100%
4-HBA-holo-EntB	33,948.9	33,947.9	46.9%
2-HBA-holo-EntB	33,948.9	33,947.9	9.3%
3-HBA-holo-EntB	33,948.9	33,948.3	25.4%
Benzoyl-holo-EntB	33,932.9	33,931.9	13.7%

^a Calculated from: (peak area of acylated EntB/ peak area of apo-EntB + peak area of acylated holo-EntB) × 100%.

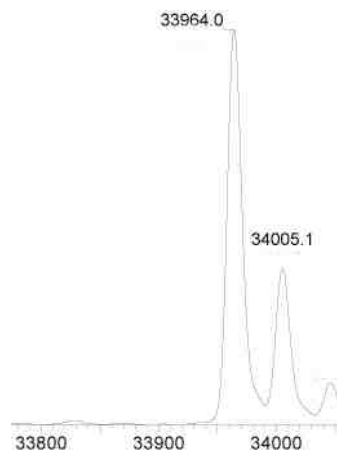
A Apo-EntB



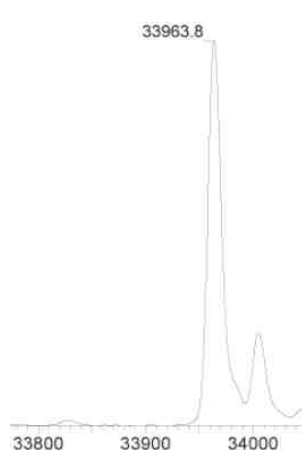
B HS-holo-EntB



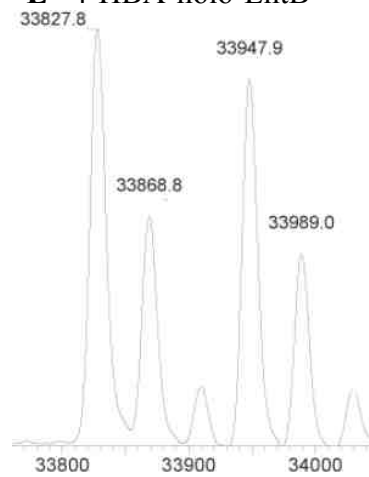
C 2,3-DHB-holo-EntB



D 2,4-DHB-holo-EntB



E 4-HBA-holo-EntB



F 2-HBA-holo-EntB

G 3-HBA-holo-EntB

H Benzoyl-holo-EntB

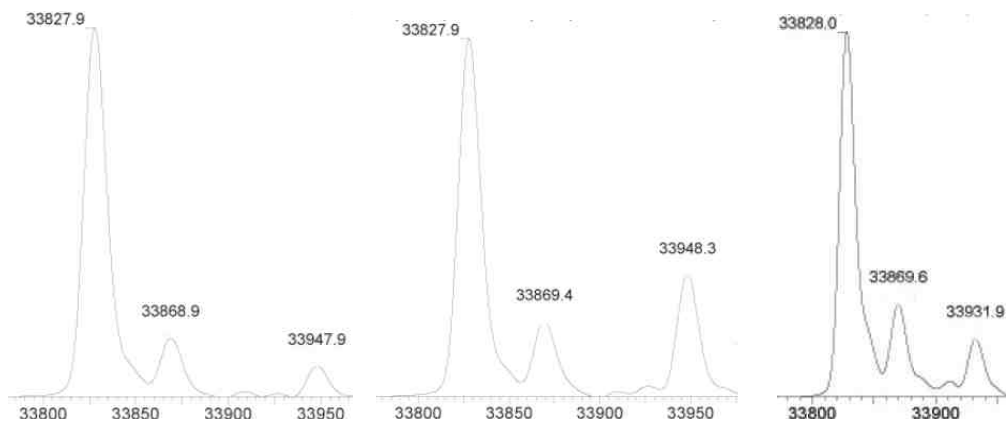


Figure 2.7 ESI mass spectra of **A)** apo-EntB, **B)** holo-EntB and of the reaction products **C)** 2,3-DHB-holo-EntB; **D)** 2,4-DHB-holo-EntB; **E)** 4-HBA-holo-EntB; **F)** 2-HBA-holo-EntB; **G)** 3-HBA-holo-EntB; **H)** benzyol-holo-EntB formed from the reaction of holo-EntB, ATP and the corresponding aromatic acid in the presence of EntE.

2.3.5 Steady-state Kinetic Constant Determinations for EntH Catalyzed Thioester Hydrolysis of Acyl-CoAs and Acyl-holo-EntBs.

Having provided support for the hypothesis that misacylated holo-EntBs might be formed *in vivo*, we focused our attention on the possible roles that EntH might play in “prevention and rescue”.

Firstly, EntH thioesterase activity toward acyl- and aroyl-CoAs was tested by using a DTNB-based continuous spectrometric assay. The steady-state kinetic constants obtained and reported in Table 2.4, show that EntH is promiscuous. All thioesters tested were hydrolyzed, however on the basis of the substrate specificity constants (k_{cat}/K_m) there is a noticeable preference for hydroxylated benzoyl-CoAs and hydroxylated phenylacetyl-CoAs ($k_{cat}/K_m > 1 \times 10^4 \text{ M}^{-1}\text{s}^{-1}$). Thus, it is conceivable that EntH is used to

harvest CoA from endogenous acyl and aroyl-CoAs, thereby facilitating EntD-catalyzed formation of holo-EntB and avoiding formation of (mis)acylated-holo-EntB.

Next, EntH-catalyzed hydrolysis of acyl- and aroyl-holo-EntBs was examined. First, the reaction of 98 μM 2,4-DHB-holo-EntB and 2 μM EntH in 50 mM HEPES buffer (pH 7.5 and 25°C) containing 50 mM NaCl and 1 mM DTT was monitored by ESI-MS over a 25 min period to demonstrate the time-dependent consumption of 2,4-DHB-holo-EntB and formation of holo-EntB (Figure 2.8).

Table 2.4 Steady-state kinetic constants for EntH catalyzed hydrolysis of acyl-CoA and acyl-holo-EntB thioesters pH 7.5 and 25 °C measured using the DTNB spectrophotometric assay.

Substrate	k_{cat} (s^{-1})	K_{m} (μM)	$k_{\text{cat}}/K_{\text{m}}$ ($\text{M}^{-1}\text{s}^{-1}$)
Acetyl-CoA	0.0044±0.0002	800±90	5.5
Propionyl-CoA	0.0125±0.0006	400±40	31.2
Decanoyl-CoA	0.0267±0.0004	48.6±3.1	5.4×10^2
Lauroyl-CoA	0.0281±0.0003	44.6±2.3	6.2×10^2
Pamitoyl-CoA	0.0854±0.0044	55.3±9.1	1.5×10^3
3-HPA-CoA ^b	2.12±0.48	36.9±0.2	5.7×10^4
3-HBA-CoA	1.21±0.01	34.7±1.2	3.4×10^4
4-HBA-CoA	1.55±0.05	21.1±1.7	7.3×10^4

2,4-DHB-holo-EntB	3.7±0.1	25.0±1.0	1.4×10 ⁵
2,3-DHB-holo-EntB	2.8±0.1	15.4±1.0	1.8×10 ⁵
lauroyl-holo-EntB	0.01±0.0	31.5±2.1	6.3×10 ²

^a no detection

^b 3-hydroxyphenylacetyl-CoA

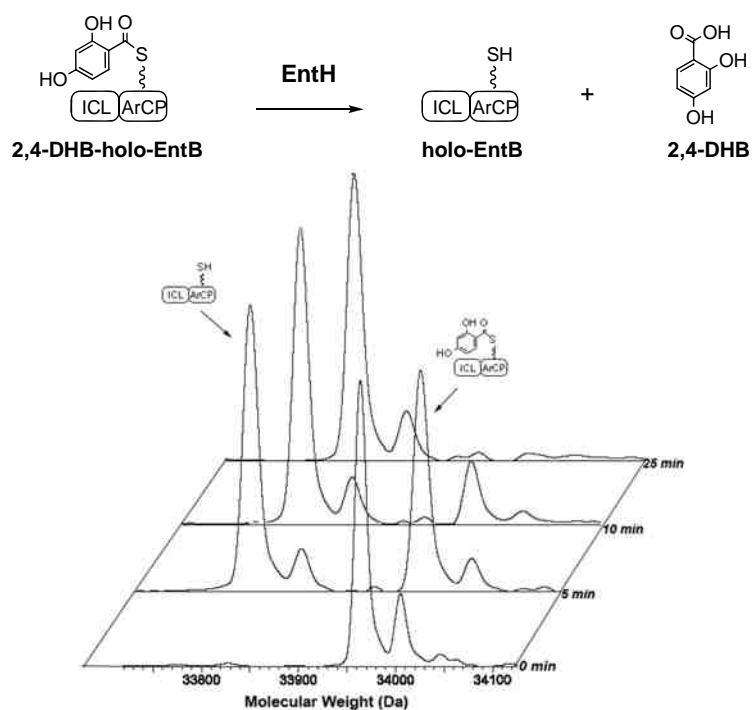


Figure 2.8 ESI-MS analysis of EntH catalyzed 2,4-DHB-holo-EntB hydrolysis at pH 7.5 and 25°C.

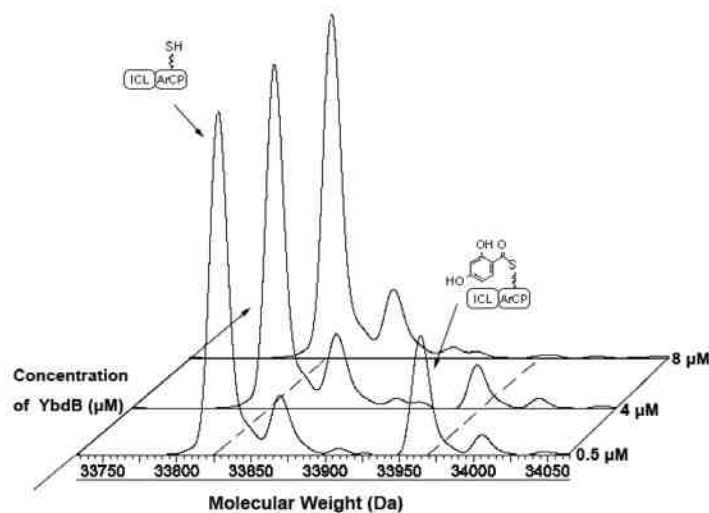


Figure 2.9 ESI-MS analysis of the hydrolysis of 2,4-DHB-S-holo-EntB by incubation with 0.5 μM , 2 μM and 8 μM of EntH for 15 min at pH 7.5 and 25 $^{\circ}\text{C}$.

As judged from ESI-MS spectrum, the control reaction (which did not include EntH) showed <5% spontaneous hydrolysis of 2,4-DHB-holo-EntB 25 min incubation at 25 $^{\circ}\text{C}$. The rate of 2,4-DHB-holo-EntB thioester hydrolysis was measured by ESI-MS as a function of EntH concentration to demonstrate linearity in the time course (Figure 2.9). In addition, a reaction of 250 μM 2,4-DHB-holo-EntB and 2 μM EntH in 50 mM HEPES buffer (pH 7.5) containing 50 mM NaCl was monitored by HPLC to demonstrate the time-dependent formation of 2,4-DHB (Figure 2.10).

Having demonstrated EntH-catalyzed conversion of 2,4-DHB-holo-EntB to holo-EntB and 2,4-DHB, a DTNB-based spectrometric assay was employed to determine the steady-state kinetic constants for the hydrolysis of lauroyl-holo-EntB, 2,4-DHB-holo-EntB and 2,3-DHB-holo-EntB.

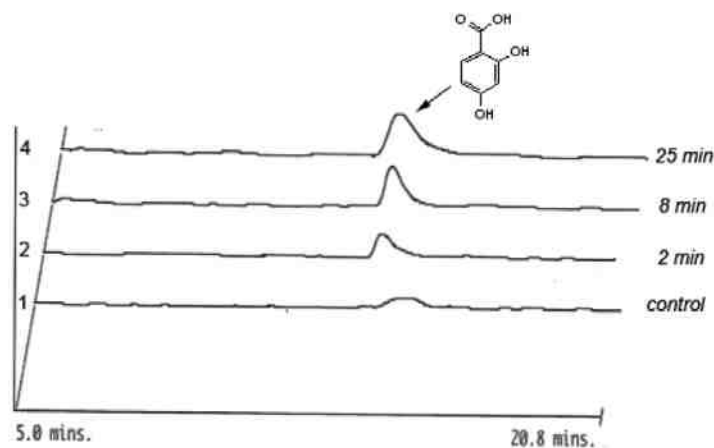


Figure 2.10 HPLC analysis of formation of 2,4-DHB during the hydrolysis of 2,4-DHB-holo-EntB by EntH at 314 nm at 25°C. Lane 1(control): 2,4-DHB-holo-EntB incubated for 25 min in the absence of EntH; Lane 2 to 4: 2,4-DHB-holo-EntB incubated with EntH for 2 min, 8 min and 25 min, respectively.

The k_{cat}/K_m value measured for lauroyl-holo-EntB thioester hydrolysis catalyzed by EntH is equivalent to that measured for lauroyl-CoA hydrolysis. Owing to the repeated failed attempts to chemically synthesize 2,3-DHB-CoA and 2,4-DHB-CoA (the ring C2OH interferes) a direct comparison of their EntH activities with 2,3-DHB-holo-EntB and 2,4-DHB-holo-EntB could not be made. On the other hand, 2,3-DHB-holo-EntB and 2,4-DHB-holo-EntB displayed 2-4 fold higher substrate specificity constants than 3-HB-CoA and 4-HB-CoA. The results suggest that EntH binds the phosphopantetheine arm of the thiol unit and does not bind the nucleotide of CoA or EntB moieties. This kinetics-based finding is consistent with the report that EntH binds holo-EntB but does not bind EntB.

We conclude that EntH recognizes the acyl-pantetheinephosphate and aroyl-pantetheinephosphate of the corresponding CoA or EntB thioester substrate and that it discriminates potential substrates on the basis of the structure of the acyl/aroyl unit, as indicated by the distinct preference observed for aromatic substrates versus aliphatic substrate.

2.3.6 ESI-MS Analysis of the Processivity of Acylated- or Aroylated-holo-EntB by EntF.

The observation that the lower pathway intermediate 2,3-DHB-holo-EntB is efficiently hydrolyzed by EntH raises the important question of how EntH might selectively remove misacylated holo-EntB. One possible scenario is that EntF removes the 2,3-DHB unit from 2,3-DHB-holo-EntB more effectively than does EntH, whereas the reverse might be true for misacylated holo-EntB substrates.

To explore this idea, the comparative rates at which L-Ser-holo-EntF abstracts the 2,3-DHB, 2,4-DHB, or lauroyl group from the corresponding charged holo-EntB were measured by using ESI-MS to monitor the reactions. The results showed that 2,3-DHB-holo-EntB (Figure 2.11 A), the correct intermediate in lower pathway of enterobactin biosynthesis, is processed significantly faster than 2,4-DHB-holo-EntB (Figure 2.11 B, control reaction in Figure 2.11 C)) and the rate of formation of holo-EntB from lauroyl-holo-EntB (Figure 2.11D) does not exceed the rate of spontaneous hydrolysis.

The slower reaction rate with misacylated holo-EntB might provide EntH with the opportunity to hydrolyze the misacylated holo-EntB, thus avoiding precursor misincorporation and/or premature termination of the assembly process. Overproduction

of EntH, however, is known to impair enterobactin biosynthesis (15) and this can be rationalized by having significantly more EntH present than the endogenously expressed EntF with the consequence of 2,3-DHB-holo-EntB being hydrolyzed.

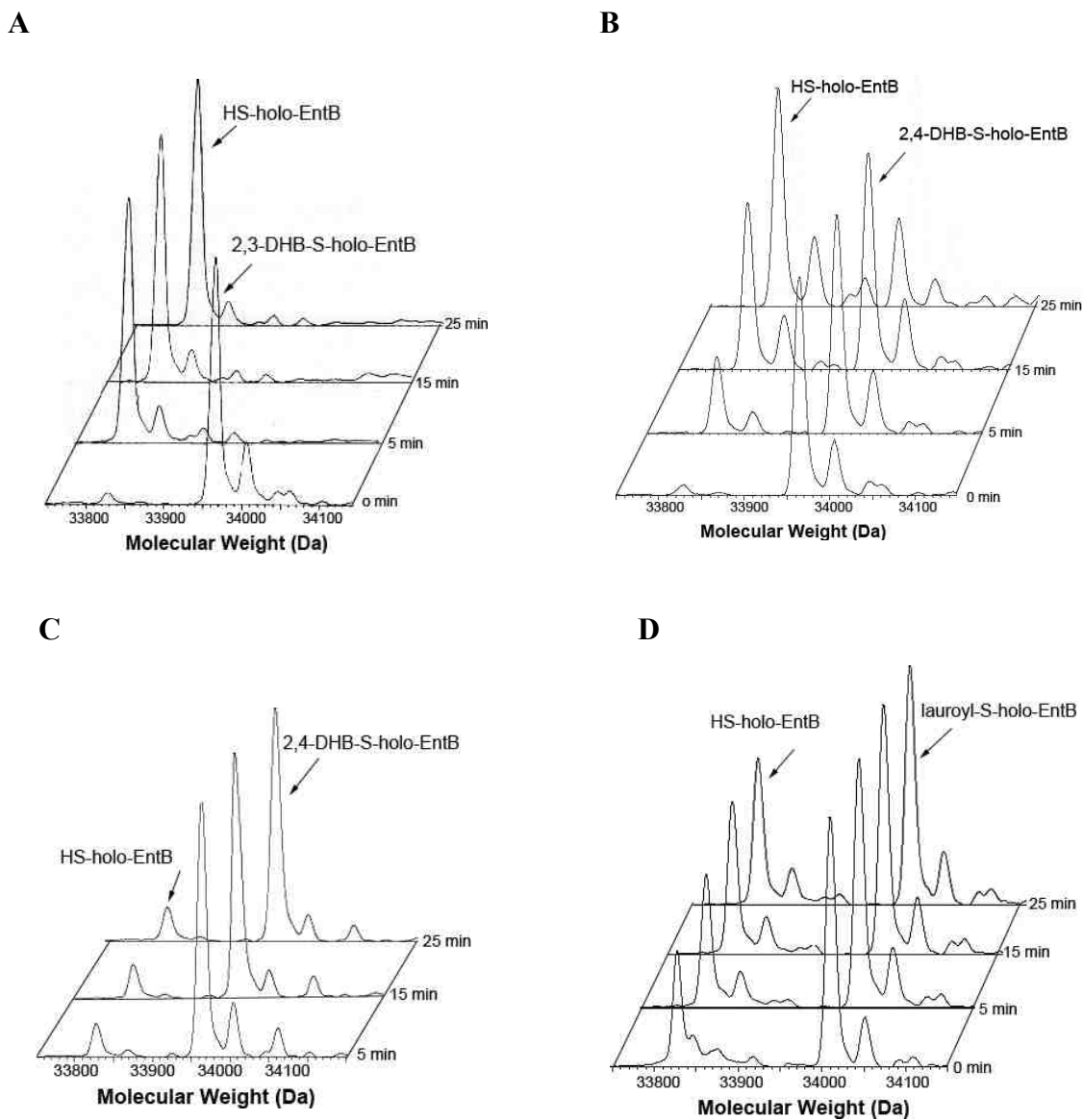


Figure 2.11 ESI-MS spectra measured as a function of incubation time for reactions of holo-EntF with A) 2,3-DHB-holo-EntB, B) 2,4-DHB-holo-EntB, C) 2,4-DHB-holo-EntB without ATP and L-ser and D) lauroyl-holo-EntB.

2.4 Conclusion

In summary, two possible routes for formation of misacylated holo-EntB that are made by the low substrate specificity of EntD and EntE were identified. In addition, kinetic analysis showed that EntH can potentially prevent EntD-catalyzed mischarging of EntB through hydrolysis of endogenous acyl- or aroyl-CoAs and that EntH can release holo-EntB from misacylated holo-EntB formed by EntE. The EntB does not successfully compete with EntF for 2,3-DHB-holo-EntB but does successfully compete with EntF for misacylated holo-EntB and in this manner rescues holo-EntB without interrupting the flux through the lower Enterobactin pathway.

2.5 References

1. Barry, S. M., and Challis, G. L. (2009) *Curr Opin Chem Biol* **13**, 205-215
2. Challis, G. L., and Naismith, J. H. (2004) *Curr Opin Struct Biol* **14**, 748-756
3. Raymond, K. N., Dertz, E. A., and Kim, S. S. (2003) *Proc Natl Acad Sci U S A* **100**, 3584-3588
4. Crosa, J. H., and Walsh, C. T. (2002) *Microbiol Mol Biol Rev* **66**, 223-249
5. Brickman, T. J., Ozenberger, B. A., and McIntosh, M. A. (1990) *J Mol Biol* **212**, 669-682
6. Gehring, A. M., Mori, I., and Walsh, C. T. (1998) *Biochemistry* **37**, 2648-2659
7. Lai, J. R., Fischbach, M. A., Liu, D. R., and Walsh, C. T. (2006) *J Am Chem Soc* **128**, 11002-11003
8. Zhou, Z., Lai, J. R., and Walsh, C. T. (2006) *Chem Biol* **13**, 869-879
9. Dillon, S. C., and Bateman, A. (2004) *BMC Bioinformatics* **5**, 109
10. Cao, J., Xu, H., Zhao, H., Gong, W., and Dunaway-Mariano, D. (2009) *Biochemistry* **48**, 1293-1304
11. Schwarzer, D., Mootz, H. D., Linne, U., and Marahiel, M. A. (2002) *Proc Natl Acad Sci U S A* **99**, 14083-14088
12. Reimann, C., Patel, H. M., Walsh, C. T., and Haas, D. (2004) *J Bacteriol* **186**, 6367-6373
13. Yeh, E., Kohli, R. M., Bruner, S. D., and Walsh, C. T. (2004) *ChemBiochem* **5**, 1290-1293
14. Gully, D., and Bouveret, E. (2006) *Proteomics* **6**, 282-293
15. Leduc, D., Battesti, A., and Bouveret, E. (2007) *J Bacteriol* **189**, 7112-7126

16. Thoden, J. B., Zhuang, Z., Dunaway-Mariano, D., and Holden, H. M. (2003) *J Biol Chem* **278**, 43709-43716
17. Song, F., Zhuang, Z., Finci, L., Dunaway-Mariano, D., Kniewel, R., Buglino, J. A., Solorzano, V., Wu, J., and Lima, C. D. (2006) *J Biol Chem* **281**, 11028-11038
18. Gehring, A. M., Bradley, K. A., and Walsh, C. T. (1997) *Biochemistry* **36**, 8495-8503

CHAPTER 3

STRUCTURE-FUNCTION ANALYSIS OF THE *E. COLI*

ENTEROBACTIN PATHWAY THIOESTERASE EntH (YbdB)

3.1 Introduction

EntH (YbdB) is synthesized in enterobactin producing bacteria. In the previous chapter, I provided kinetic evidence, derived from *in vitro* substrate activity assays, that EntH assists in the production of enterobactin by functioning in a prevention-and-rescue capacity. Specifically, EntH was shown to be an active catalyst in the thioester hydrolysis of acyl-CoA metabolites and in this capacity it liberates CoA, the substrate used by the enterobactin pathway PPTase EntD, while consuming the acyl-CoA, which can compete with CoA as substrate for EntD, and hence form a “deadend” acyl-holo-EntB. EntH was also shown to catalyze the hydrolysis of misacylated-holo-EntB, thereby rescuing holo-EntB for EntE catalyzed acylation with 2', 3'-dihydroxybenzoate.

The direct binding of EntH to holo-EntB has been evidenced by *in vivo* yeast 2-hybrid experiments and by *in vitro* pull-down experiments (1). This result suggests that the Michaelis complex formed between EntH and 2', 3'-dihydroxybenzoyl-EntB is stabilized in through electrostatic interactions between EntH and the EntB docked surfaces. On the other hand, the k_{cat}/K_m values measured for EntH catalyzed hydrolysis of acyl-CoA and acyl-holo-EntB substrates are similar (Chapter 2 and (2)), a finding which suggests that EntH is equally effective at forming a Michaelis complex through

electrostatic interactions with the CoA unit. In this chapter the structural basis for the recognition of the thiol unit (*viz.* CoA vs holo-EntB) substrates) is examined.

The EntH substrate specificity profile reported in the Chapter 2 and (2) also provides evidence that the EntH active site can accommodate substrate acyl/aroyl units of various sizes and polarities. However, the comparison of the substrate k_{cat} values indicates that the substrates, which possess an aroyl moiety, bind more (~100-fold) productively than those which possess a fatty acyl moiety. In this chapter the structural basis for catalysis and for productive substrate binding as directed by the aroyl unit are examined.

In the text below I report the X-ray structures of EntH bound with the inert substrate analogs phenacyl-CoA and with 2,4-dihydroxyphenacyl-CoA, determined in collaboration with Professor Karen Allen and Dr. Rui Wu of Boston University and compare these structures to the previously reported X-ray crystal structure of apo-EntH (PDB code: 1vh9) (3). The X-ray structures of the liganded enzyme are used to analyze the ^{31}P -NMR spectral properties of the bound ligand, the kinetic properties of active site mutants (reported here and elsewhere (4)), the conservation of the structure and function of the catalytic scaffold among thioesterases of the same hotdog family clade.

3.2 Experimental

3.2.1 EntH(YbdB) Crystallization and X-ray Diffraction Data Collection. (work carried out by Dr. Rui Wu at Boston University)

Both EntH (2,4-Dihydroxyphenacyl-CoA) and EntH (Phenacyl-CoA) complexes were crystallized by using the hanging-drop vapor diffusion method. Protein at 10 mg/ml

was combined with the substrate analogue 2,4-dihydroxyphenacyl-CoA or phenacyl-CoA to obtain a final concentration of 2 mM.

Crystals of the EntH(2,4-dihydroxyphenacyl-CoA) complex grew in 2 days from 2.0 M sodium malonate (pH 7.0) at 18 °C. The overall crystal dimensions were 0.2 X 0.2 X 0.05 mm. The crystals were cryoprotected by passing them through 100% Paratone-N (Hampton Research) for ~ 5 min and then freezing them in a stream of N₂ gas at 100 °K. Diffraction data were collected on the home source (Boston University) at 100 °K using CuK_α radiation from a Rigaku RUH3 generator equipped with the R axis IV⁺⁺ image plate. The data extending to 2.17 Å resolution were indexed and scaled using DENZO and SCALEPACK (5). The crystals were orthorhombic, belonging to the space group P2₁2₁2 with unit cell dimensions a = 59.5 Å, b = 85.9 Å, and c = 49.9 Å. Data collection and process statistics are summarized in Table 1.

Crystals of the EntH(phenacyl-CoA) complex grew in 2-3 days from 2.0 M sodium malonate (pH 7.0) and 0.1M Cobalt(II) chloride hexahydrate at 18°C. The overall crystal dimensions were 0.1 X 0.2 X 0.05 mm. The crystals were cryoprotected by passing them through 100% Paratone-N (Hampton Research) for ~ 5 min and then freezing them in a stream of N₂ gas at 100 °K. Diffraction data were collected to 1.85 Å resolution at the National Synchrotron Light Source at Brookhaven National Laboratory at beamline X12A. The data set was processed with HKL2000 (5). The crystals were monoclinic, belonging to the space group P2₁ with unit cell dimensions a = 51.4 Å, b = 59.1 Å, c = 85.9 Å, α = 90.0°, β = 90.2°, γ = 90.0°. Data collection and process statistics are summarized in Table 1.

3.2.2 EntH(2,4-Dihydroxyphenacyl-CoA) or EntH(Phenacyl-CoA) Complex Structure Refinement. (work carried out by Dr. Rui Wu at Boston University)

Both structures were solved by molecular replacement using MOLREP of the CCP4 package using the apo EntH structure (Protein Data Bank accession ID 1VH9) as a search model. Successive rounds of manual rebuilding were performed using the molecular graphics program COOT (6) alternated with minimization and simulated annealing in PHENIX. The dihedral angles of all residues are within the allowed region of the Ramachandran plot. The stereochemical quality of the final model was examined by PROCHECK. Refinement statistics are summarized in Table 3.1. The 2,4-dihydroxybenzoyl-CoA ligand and a water molecule were modeled in place of the 2,4-dihydroxyphenacyl-CoA by using COOT. Figures were prepared with PyMol (DeLano Scientific) (7), MOLSCRIPT (8), and POV-Ray (9).

3.2.3 Synthesis of Acyl-CoA analog inhibitors

2,4-Dihydroxyphenacyl-CoA and phenacyl-CoA were synthesized according to the following procedure. Under a N₂ atmosphere, 45 mg of 2-bromo-2',4'-dihydroxyacetophenone or 2-bromo-acetophenone were added dropwise over a 1 hour period to 30 mg of CoA in 1.5 mL of H₂O/THF (1:1). The solution pH was monitored and maintained within a range of 7.2-8.0. After overnight incubation, the reaction mixture was extracted with ethyl acetate twice. The crude product obtained after lyophilization of the aqueous phase was chromatographed on a G-15 Sephadex column. The fractions were monitored by HPLC (dual wavelengths 280 nm/260 nm). The fractions with desirable purity were combined, concentrated and subjected to ¹H-NMR

spectral analysis for structure confirmation (NMR facilities, the Department of Chemistry, UNM).

3.2.4 Enzymatic Assay

Steady-state kinetic methods were used to determine the k_{cat} and K_m for wild-type and mutant of YbdB. The thioesterase activities were monitored at 412 nm ($\Delta\epsilon = 13.6 \text{ mM}^{-1}\text{cm}^{-1}$) by coupling the reaction of 5,5'-dithio-bis(2-nitrobenzoic acid) (DTNB) with the CoA liberated from the hydrolysis of acyl-CoA substrates. All reactions were carried out in 50 mM HEPES (pH 7.5) containing 1 mM DTNB at 25°C.

The kinetic parameters of V_{max} and K_m were determined from initial velocity data, as measured as a function of substrate concentration, by using eq 1 and the computer program KinetAsyst (IntelliKinetics)

$$V=V_{max}\cdot[A]/([A]+K_m) \text{ (eq. 1)}$$

Where [A] is the substrate concentration, V is the initial velocity, V_{max} is the maximum velocity, and K_m is the Michaelis constant. The reported error was computed for the data fitting. The k_{cat} was calculated from the ratio of V_{max} and the total enzyme concentration. The enzyme concentration was determined by using the Bradford method.

The inhibition constant K_i for competitive inhibition was obtained by fitting the initial velocities of reactions of 4-hydroxybenzoyl-CoA (at varied concentration) in the absence or presence of inhibitor (2,4-dihydroxyphenacyl or phenacyl-CoA) to eq 2.

$$V=V_{max}[S]/[K_m(1+[I]/K_i)+[S]] \text{ (eq 2.)}$$

Where [I] is the concentration of the inhibitor, and K_i is the inhibition constant.

3.2.5 Hotdog Thioesterase Sequence and Structure Alignments, and Graphics.

Hotdog thioesterase amino acid sequences and X-ray structure coordinates were obtained from the NCBI Protein data base (<http://www.ncbi.nlm.nih.gov/protein/>) and from the RCSB PDB (<http://www.pdb.org/pdb/home/home.do>), respectively. Pair-wise and multiple sequence alignments were generated using ClustalW (<http://www.ch.embnet.org/software/ClustalW.html>). The structure alignments and graphic figures were produced using PyMOL (<http://www.pymol.org/>).

3.2.6 ³¹P-NMR Spectrum of HS-CoA and Phenacyl-CoA in the Presence and in the Absence of EntH.

Proton-decoupled ³¹P-NMR (202.5 MHz) spectra were recorded at room temperature using 5 mm NMR tubes and a Bruker AC-F 500 NMR spectrometer (calibrated with 85% phosphoric acid). Samples contained 1 mM CoA or phenacyl-CoA in 50 mM HEPES/50 mM NaCl and 0.0, 0.5 μM or 1 mM EntH.

3.3 Results and Discussion

3.3.1 Synthesis and Evaluation of Inert Substrate Analogs for Co-crystallization with EntH.

The inert substrate analogs 2,4-dihydroxyphenacyl-CoA and phenacyl-CoA (Figure 3.1 A and B) were synthesized and characterized as EntH inhibitors and as potential co-crystallization ligands. The presence of methylene (–CH₂) in between the C=O and S atom prevents EntH catalyzed hydrolytic cleavage. The binding affinities of the substrate analogs were determined by measuring their inhibition constants using

steady-state kinetic techniques. Both analogs displayed competitive inhibition (Figure 3.1 B and C). Data fitting to equation 2 defines the $K_i = 65 \pm 6 \mu\text{M}$ for 2,4-dihydroxyphenacyl-CoA, and the $K_i = 5.9 \pm 0.3 \mu\text{M}$ for phenacyl-CoA.

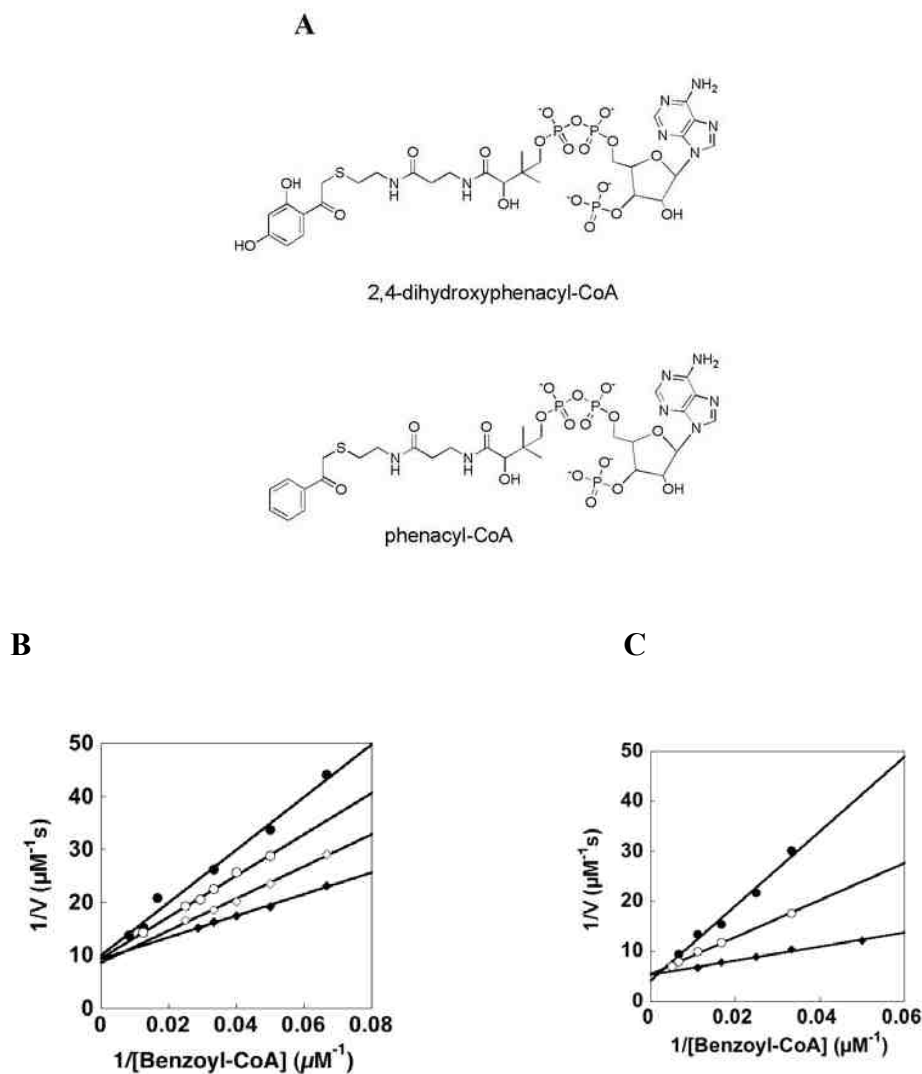


Figure 3.1 Chemical structure of the inert substrate analogs, 2,4-dihydroxyphenacyl-CoA and phenacyl-CoA (A). Initial velocity data measured for EntH (0.0645 μM) catalyzed hydrolysis of benzoyl-CoA (15 to 200 μM) in 50 mM HEPES (pH 7.5, 25 °C) containing

1 mM DTNB and 0 μM (\blacklozenge), 35 μM (\blacklozenge), 70 μM (\circ) and 105 μM (\bullet) 2,4-dihydroxyphenacyl-CoA, (B) 0 μM (\blacklozenge), 5 μM (\circ) and 10 μM (\bullet) phenacyl-CoA (C).

3.3.2 Quaternary and Tertiary Structure of the Apo and Liganded EntH

The three dimensional structure of EntH complexed with the inert substrate analogue, 2,4-dihydroxyphenacyl-CoA or phenacyl-CoA was determined to probe the structural elements that provide substrate-specific interactions. The X-ray structure of EntH complexed with 2,4-dihydroxyphenacyl-CoA was determined at 2.2 Å resolution (Table 3.1; Figure 3.2 A). Within an asymmetric unit, there are two EntH subunits, two 2,4-dihydroxyphenacyl-CoA ligands and one sodium malonate (present in crystallization solution). Based on the native molecular weight determination of the apo EntH at 56 kDa (theoretical molecular weight for the monomer \sim 14 kDa), EntH is a tetramer. The biological structure unit is therefore a dimer of dimer.

The EntH (phenacyl-CoA) complex crystallized in the monoclinic symmetry space group $P2_1$. The structure of this complex refined to 1.9 Å resolution (Table 3.1), with the final model including four EntH subunits and four phenacyl-CoA molecules. Within one asymmetric unit, there are two dimers of EntH (Figure 3.2 B). Four phenacyl-CoA ligands bind to the respective substrate binding sites. The tetramer is formed by back-to-back association of dimer 1 (formed by subunit A and D) and dimer 2 (formed by subunit B and C) (Figure 3.2 C and D).

Table 3.1 Crystallographic and refinement statistics for wild-type EntH bound with 2,4-dihydroxyphenacyl-CoA or phenacyl-CoA.

	EntH + 2,4-Dihydroxyl phenacyl: CoA	EntH + Phenacyl: CoA
Resolution (Å) ^a	50.0 – 2.17 (2.25-2.17)	50.0 – 1.85 (1.92-1.85)
Space Group	P2 ₁ 2 ₁ 2	P2 ₁
Unit Cell (Å)	a) 59.5, b) 85.9, c) 49.9	a) 51.4, b) 59.1, c) 85.9
R _{merge} ^{ab} (%)	12.4 (52.3)	8.5 (4.1)
Completeness ^a	99.8 (100)	99.8 (100)
I/σ ^a	11.1 (2.9)	13.7 (2.7)
no of measured reflections/unique	141246 /14280	164296 /44184
R _{cryst} / R _{free} ^c (%)	20.8/26.7	16.1/20.8
RMS Deviation bond lengths (Å)	0.008	0.006
RMS Deviation bond angles (°)	1.046	1.032
Average B-Factor (Å ²)		
Protein main chain atoms	44.6	25.4
Protein side chain atoms	49.7	31.4
Inhibitor atoms	75.5	57.7
Sodium malonate atoms	62.6	N/A
Solvent	49.8	37.5

^aValues for the highest resolution shell are given in parentheses.

^b $R_{\text{merge}} = \sum |I_i - I_m| / \sum I_i$, where I_i is the intensity of the measured reflection and I_m is the mean intensity of all symmetry-related reflections.

^c $R_{\text{cryst}} = \sum |F_{\text{obs}}| - |F_{\text{calc}}| / \sum |F_{\text{obs}}|$, where F_{obs} and F_{calc} are observed and calculated structure factors. $R_{\text{free}} = \sum T|F_{\text{obs}}| - |F_{\text{calc}}| / \sum T|F_{\text{obs}}|$, where T is a test data set of about 5% of the total reflections randomly chosen and set aside prior to refinement.

The monomer fold consists of six extended β -strands (β 1-6) and the central α -helix (α 3), characteristic of a “hot dog” fold (Figure 3.2 B) (10). The monomers dimerize through interaction between β 3 (residues Val 81 to His 89) from each of subunit, forming an extended anti-parallel β -sheet with 12 strands (6 from each monomer). The dimer is further stabilized by multiple hydrogen bonds between two central α 3 helices.

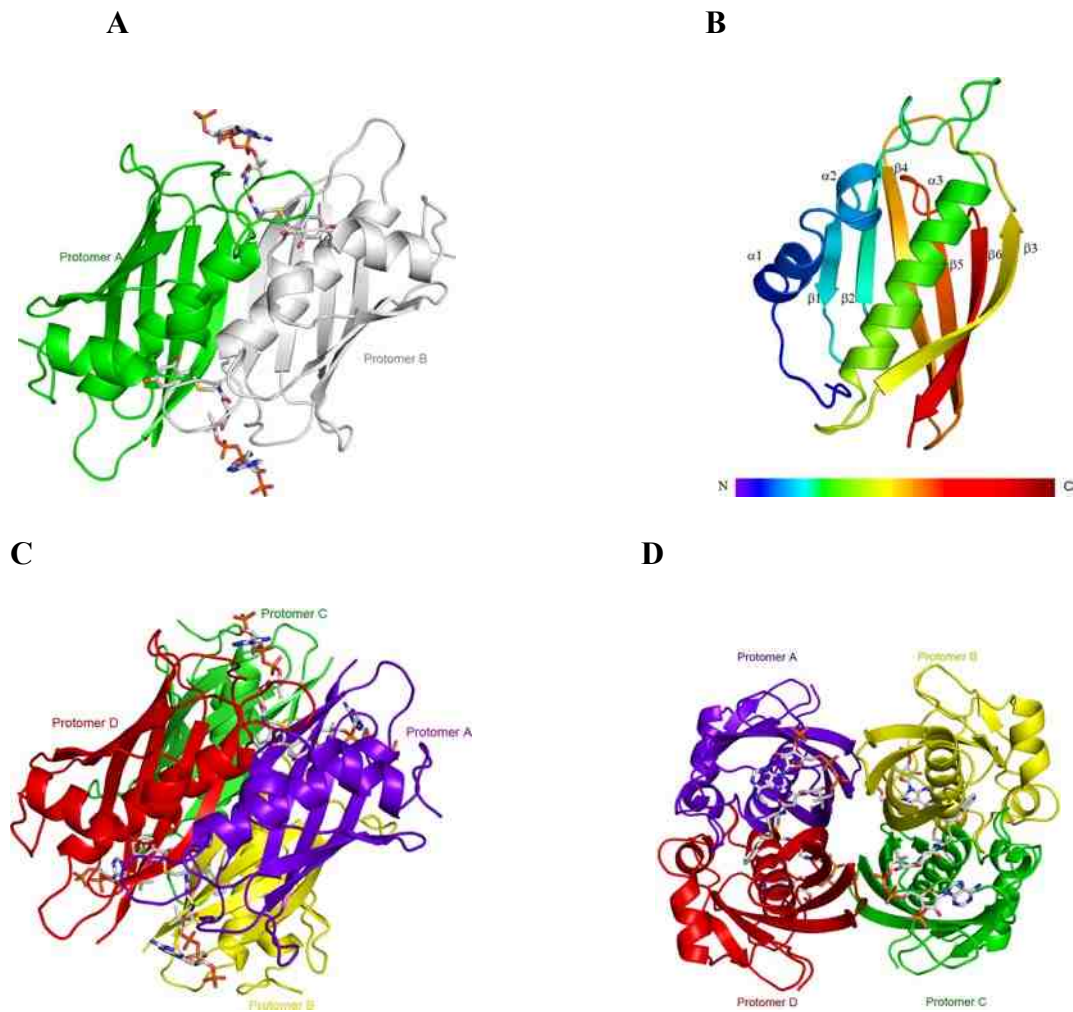
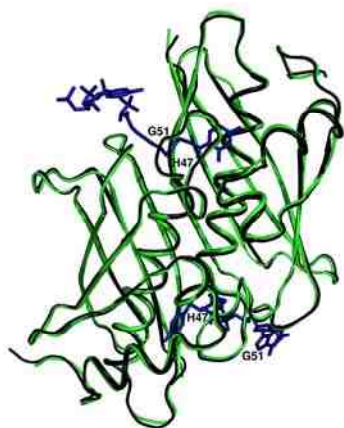


Figure 3.2 (A) The EntH dimer bound with 2,4-dihydroxyphenacyl-CoA (subunit A in green and subunit B in gray). (B) Illustration of the tertiary structure of the EntH monomer. (C) and (D) The EntH tetramer (bound with phenacyl-CoA) shown in two different orientations.

At the protein surface, layered on top of α_3 , are the two short N-terminal α -helices (α_1 and α_2), which likely prohibit the face-to-face dimer-dimer interaction via the central α -helices that is observed for the *Pseudomonas* 4-HBA-CoA thioesterase, the prototype of the clade type AA hotdog fold thioesterase (11). Instead, the two EntH dimers (red plus purple subunits and yellow plus green subunits, Figure 3.2C) associate through the extensive β -sheet, a “back to back” organization (Figure 3.2 D). The observed “back-to-back” dimer association has been reported for the human thioesterase superfamily member 2 (Them2, recently renamed as Acot13) (12), *Arthrobacter sp.* Strain SU 4-hydroxybenzoyl-CoA thioesterase (13-14) and *E. coli* phenylacetyl-CoA thioesterase (Paal) (11,15).

A



B

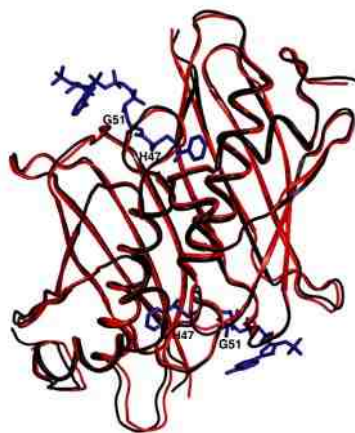


Figure 3.3 (A) The overlay of the *apo*-EntH (black) and EntH (2,4-dihydroxy phenacyl-CoA) complex (green). **(B)** The overlay of the *apo*-EntH (black) and EntH (phenacyl-CoA) complex (red). The ligands are colored blue in both figures. The small loop (His47-Gly51), which leads into the N-terminus of the central α helix and that flanks the ligand binding site, is indicated in both figures.

The structures of the respective dimer units of the EntH (2,4-dihydroxyphenacyl-CoA) and EntH (phenacyl-CoA) complexes superimpose with the structure of the dimer unit of the *apo*-EntH structure (PDB 1VH9) with a α carbon atom RMSD value of 0.37 Å and 0.43 Å (Figure 3.3 A). Thus 2,4-dihydroxyphenacyl-CoA or phenacyl-CoA binding does not induce a significant change in the conformation of the protein backbone. On the other hand, several loops adopt slightly different conformations in the liganded structures as compared to the *apo* structure. Most noticeable is a small loop, consisting of His47-Leu52, which leads into the N-terminus of the central α helix and that flanks the ligand-binding site.

3.3.3 An Overview of the EntH Substrate Binding Site

The substrate binding sites of hotdog thioesterases are modular. The nucleotide unit of the acyl/aroyl-CoA substrate or the protein unit of the acyl/aroyl-CoA substrate is accommodated at the surface of the enzyme, as the phosphopantetheinyl arm threads through a channel or tunnel that leads to the acyl/aroyl thioester binding pocket. The hotdog thioesterase substrate-binding site is located at the interface of two subunits of the dimer. In the case of the type AB clade of hotdog thioesterases, one subunit of the

opposing dimer interacts (via a loop connecting two β -stands) with the nucleotide unit of the acyl/aryl-CoA substrate. As is illustrated in Figure 3.4, using the structure of EntH(phenacyl-CoA) complex as the example, the substrate binding site of EntH conforms to this model. The nucleotide units of the phenacyl-CoA or the 2,4-dihydroxyphenacyl-CoA “sit” on the surface of the protein as illustrated in Figure 3.5, while the phosphopantetheinyl arm threads through a narrow tunnel, which leads to the catalytic site. The pocket that forms the EntH catalytic site is confined, and is mostly closed to solvent (Figure 3.6).

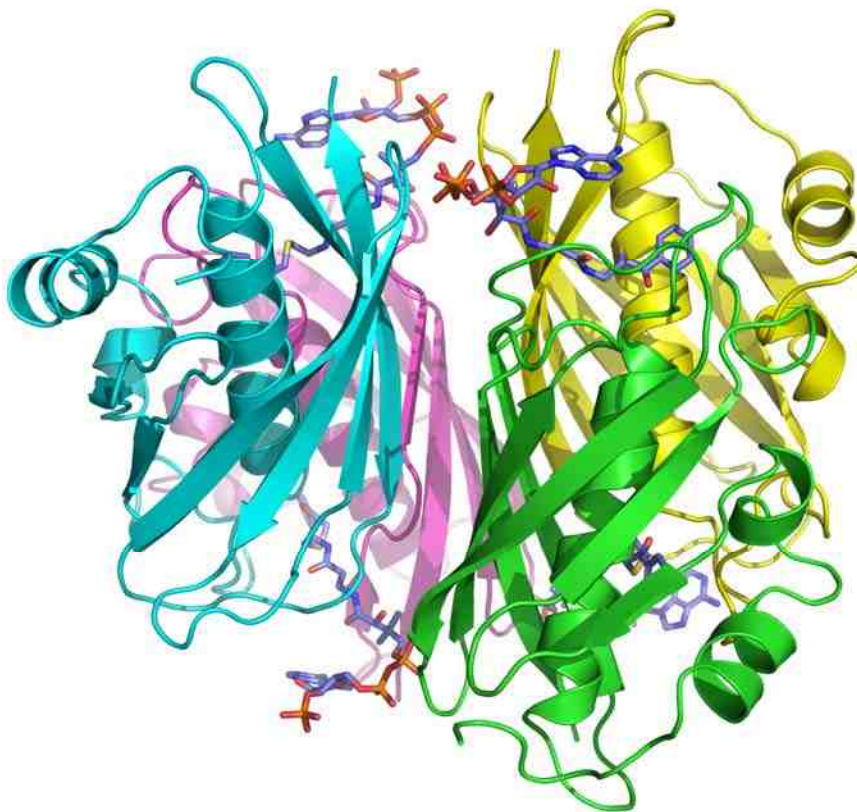


Figure 3.4 A cartoon representation of the structure of the EntH(phenacyl-CoA) complex showing the location of the phenacyl-CoA ligand relative to the three subunits that form the substrate binding site. The phenacyl-CoA ligand is shown in stick representation

(carbon atoms light blue, nitrogen blue, oxygen red, sulfur yellow and phosphorus orange).

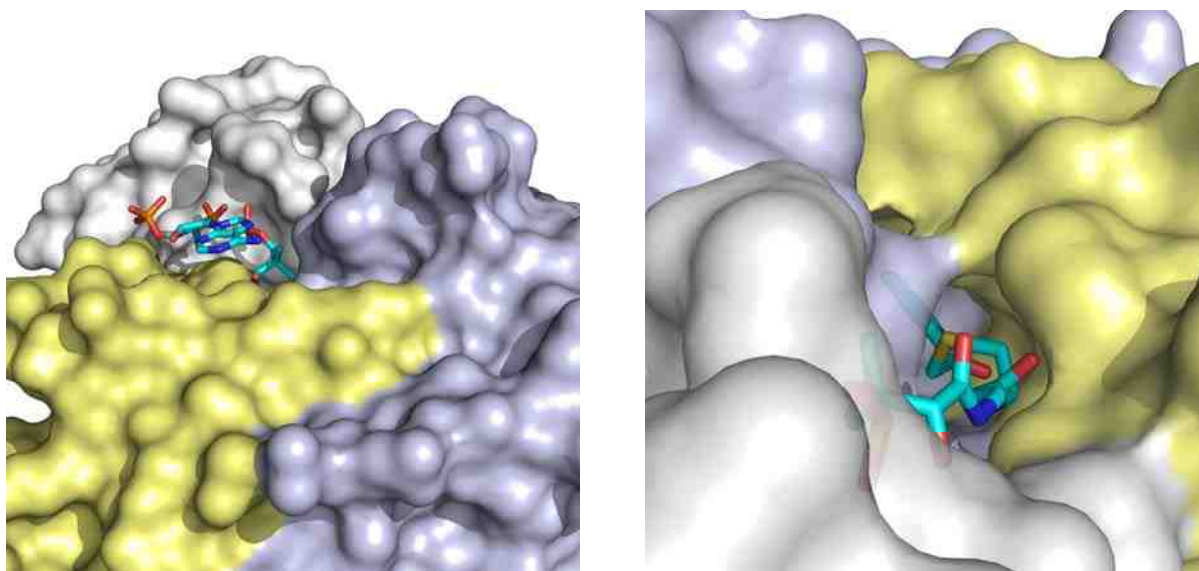


Figure 3.5 The structure of the EntH(phenacyl-CoA) complex with the phenacyl-CoA ligand shown in stick representation (carbon atoms cyan, nitrogen blue, oxygen red, sulfur yellow and phosphorus orange) and the surfaces of three subunits that contribute to the substrate binding site colored yellow, light gray and dark gray. The panel on the left depicts the phenacyl-CoA nucleotide moiety at the protein surface. The panel to the right depicts that tunnel through which the phosphopantetheinyl arm threads (for clarity the nucleotide moiety is not shown).

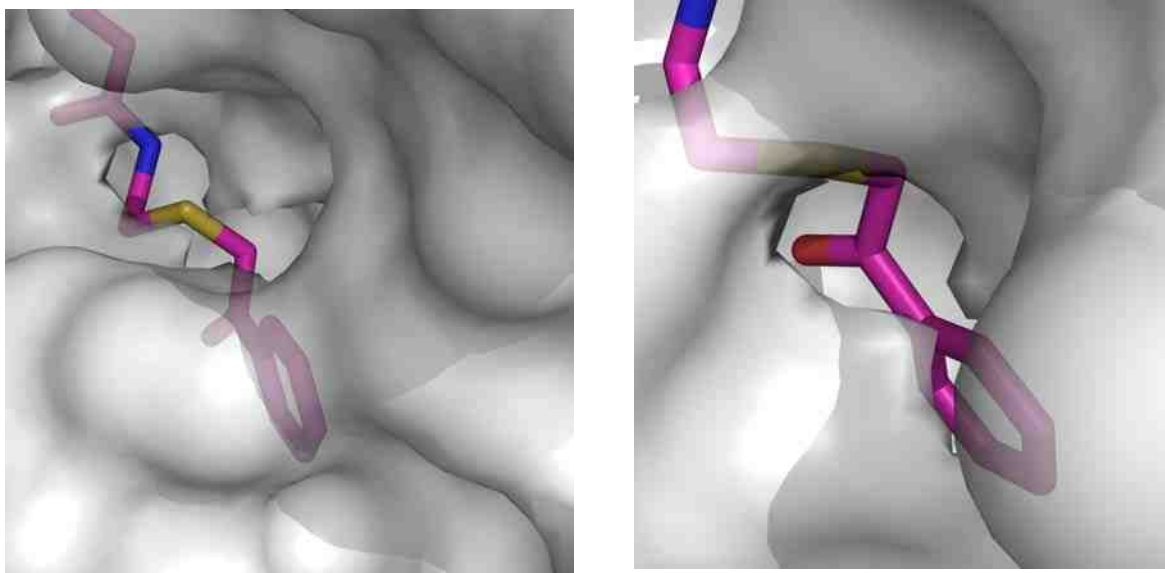


Figure 3.6 Two snapshots of the aroyl thioester binding pocket observed in the structure of the EntH(phenacyl-CoA) complex. The phenacyl-CoA ligand shown in stick representation (carbon atoms magenta, nitrogen blue, oxygen red, and sulfur yellow).

3.3.4. Comparative Analysis of the EntH ACP/CoA Binding Site.

As was reported in Chapter 2, EntH is highly active with hydroxybenzoyl-CoA and hydroxybenzoyl-EntB substrates. The structures of EntH bound with 2,4-dihydroxyphenacyl-CoA or phenacyl-CoA show that the ligand nucleotide moiety sits on the surface of the enzyme (Figure 3.5). A model built from docking the ArCP domain of EntB with the corresponding EntH site is shown in Figure 3.7. The model suggests that the surface, which accommodates the CoA nucleotide is also topologically compatible with the ArCP protein. Efforts in the protein X-ray structure laboratory at Boston University (Professor Karen Allen and Rui Wu) are currently directed at determining the structure of EntH bound with the holo-ArCP domain of EntB.

My work on the other hand, has focused on analysis of the EntH CoA binding site in the context of other type AB hotdog thioesterases, which specifically target acyl/aroyl-CoA substrates or like EntH, catalyze thioester ester hydrolysis in both acyl/aroyl-CoA and acyl/aroyl-CoA substrates. Figure 3.8 depicts the interactions observed between the CoA unit of phenacyl-CoA ligand and the EntH binding site.

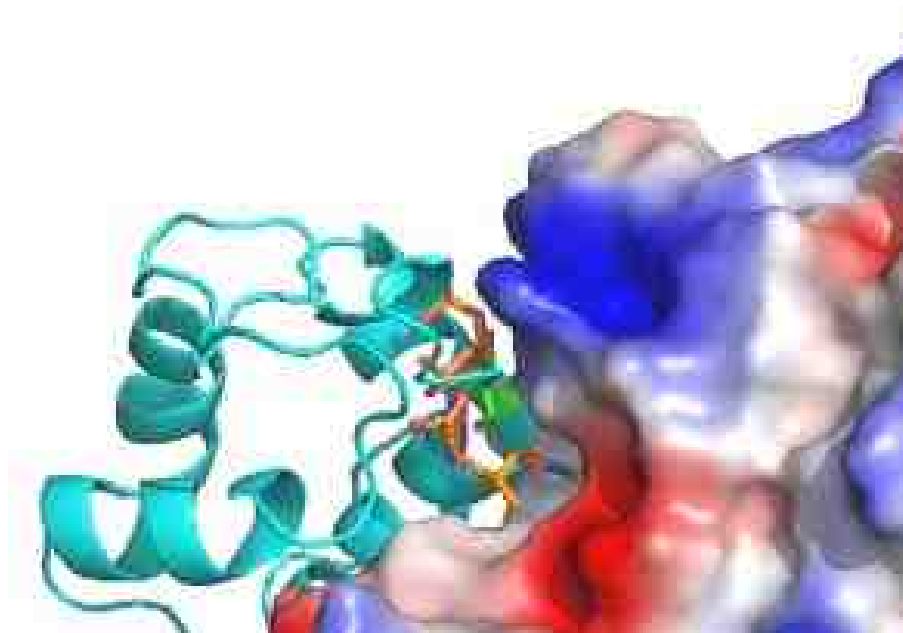


Figure 3.7 The EntH-EntB docking model. The cartoon representation of the EntB apo-ACP domain is colored cyan with the side chain of the Ser phosphopantetheinyl attachment residue shown in stick representation (colored carbon green, oxygen red). The EntH surface is colored red for oxygen, blue for nitrogen and white for carbon. The phenacyl-CoA ligand is shown in stick (carbon and phosphorus orange, oxygen red, nitrogen blue).

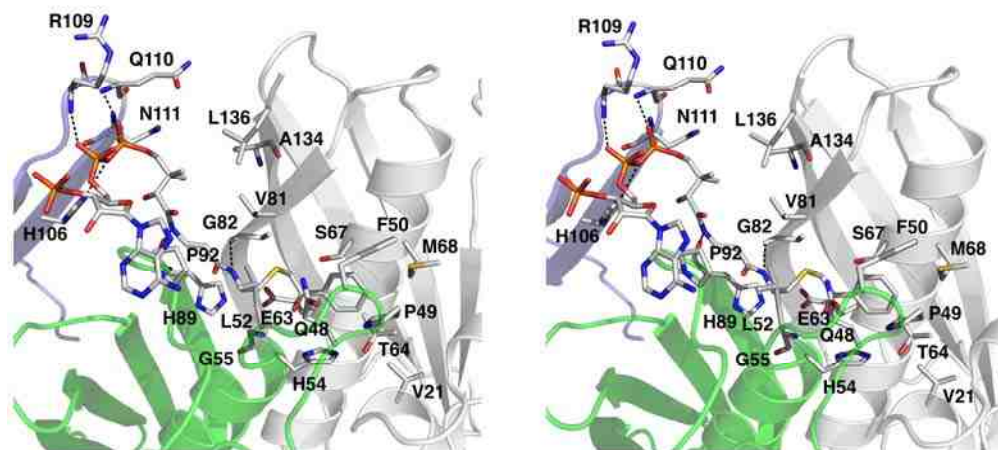


Figure 3.8. Stereoscopic representation of the ligand-binding interactions observed in the structure of the EntH (phenacyl-CoA) complex. The coloring is as follows: gray for carbon atom, blue for nitrogen atom, red for oxygen atom, yellow for sulfur atom and brown for phosphorus atom. The two colors (gray and light green) of the ribbon trace correspond to the two molecules within the dimer. The light blue ribbon corresponds to a subunit of a neighboring dimer that flanks the nucleotide-binding site.

The methyl groups of the phosphopantetheinyl phosphate unit interact with several hydrophobic residues: Leu136, Val81 and Ala134 from one subunit and Leu52 and Pro92 from the opposing subunit. The main chain amide groups form hydrogen bonds with the phosphopantetheinyl arm, one with a backbone carbonyl group of Gly82 of one subunit and the second with a His89 backbone carbonyl group of the opposing subunit. The 5'-pyrophosphate of the phenacyl-CoA ligand forms an extensive hydrogen bond network with His106, Arg109, Gln110 and Asn111 of a monomer from the opposing dimer. These latter interactions suggest the pyrophosphate unit CoA is anchored but the 3'-AMP moiety is not. The electron density maps of the 2,4-

dihydroxyphenacyl-CoA or phenacyl-CoA bound at the interface of two EntH subunits are presented in Figure 3.9 A and B. The nucleotide moiety of the ligands in both structures showed very poor electron density in contrast to the pantetheine arm and aromatic ring. This suggests that the nucleotide is not confined but rather assumes different orientations, even in the crystalline state. For comparison, the hotdog thioesterase of the same type AB clade hTHEM2 was selected. Like EntH, the human THEM2 (hTHEM2) displays broad substrate specificity towards aroyl and acyl-CoAs however, unlike EntH it prefers fatty acyl-CoA substrates over aroyl-CoA substrates and most importantly, it does not recognize the fatty acyl-ACP (ACP is the engineered cytosolic fatty acid synthase domain (19)). The X-ray structure of the hTHEM2 bound with the inert analog of the substrate decanoyl-CoA, undecan-2-one-CoA, shows much higher electron density for the ligand's nucleotide moiety (Figure 3.9 C), indicative of a conformational restriction. Likewise, the electron density (not shown) observed for the nucleotide unit of the 4-hydroxyphenacyl-CoA ligand bound to the hotdog thioesterase type AB clade member 4-hydroxybenzoyl-CoA thioesterase is well defined (14).

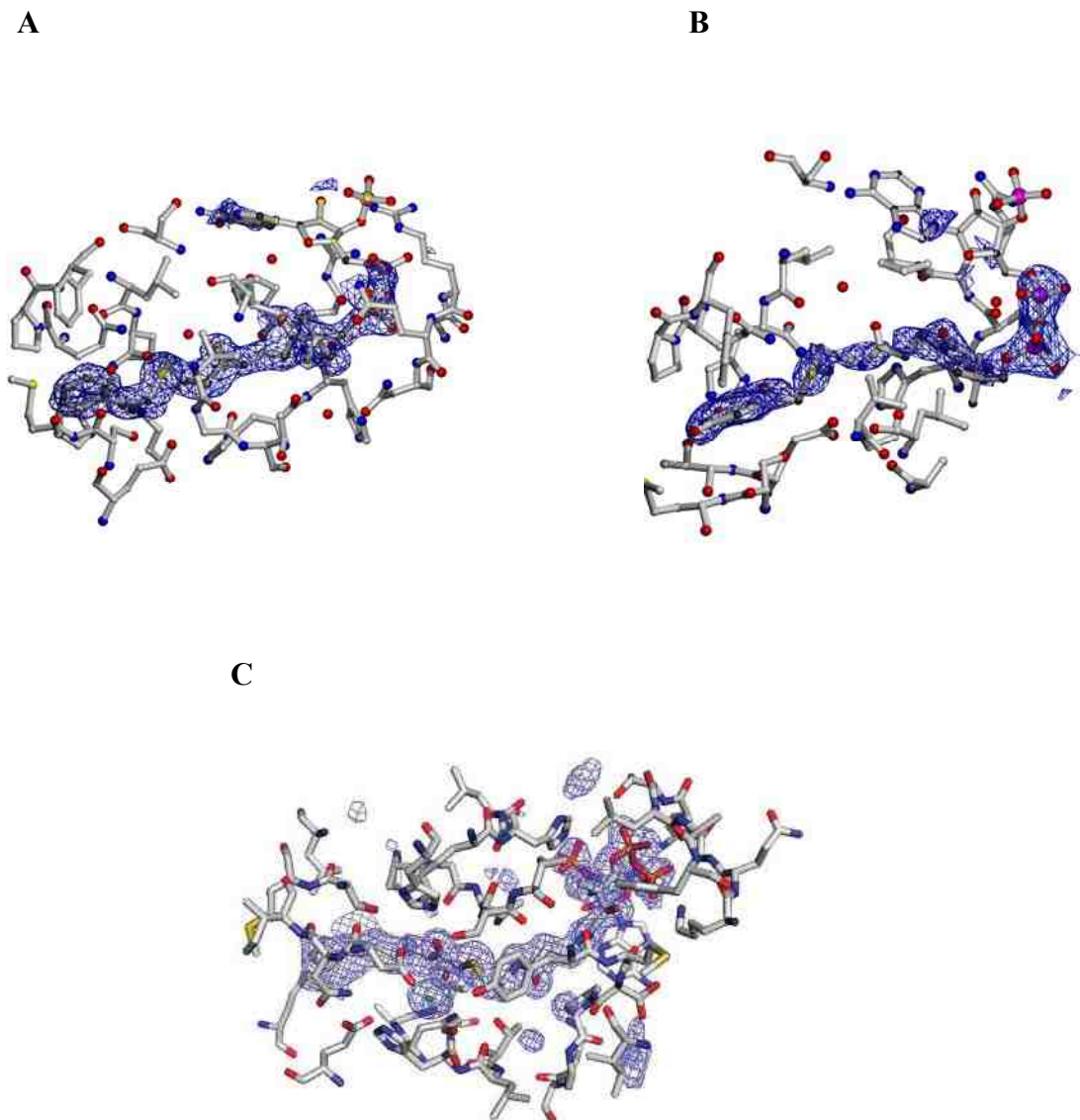


Figure 3.9 The composite omit map of **(A)** 2,4-dihydroxyphenacyl-CoA ligand in subunit A of EntH (2,4-dihydroxyphenacyl-CoA) complex, **(B)** phenacyl-CoA in subunit D of EntH (phenacyl-CoA) complex and **(C)** undecan-2-one-CoA in hTHEM2 (undecan-2-one-CoA) complex (PDB ID 3F5O). The density of the map (shown in blue) is contoured at the level of 1σ . The ligand carbon atoms are colored black, the EntH carbon atoms are colored white, and all oxygen, nitrogen sulfur and phosphorus atoms are colored red, blue, yellow and magenta respectively.

In order to test whether the phosphate groups of the nucleotide unit of the phenacyl-CoA bind to EntH, the ^{31}P -NMR spectra phenacyl-CoA measured in the presence and in the absence of excess EntH are compared in Figure 3.10. A difference in chemical shift and and/or linewidth of the 3' -P and 5' -PP resonances measured for the EntH(phenacyl-CoA) complex in buffer vs the phenacyl-CoA in buffer would signify some degree of association of ligand nucleotide unit with enzyme surface. ^{31}P -NMR resonances derived from the 5' -PP of bound phenacyl-CoA are not shifted but are broadened whereas the 3' -P resonance is both broadened and shifted 0.22 ppm downfield. In contrast, the same experiment carried out with CoA in place of the phenacyl-CoA showed no difference of the CoA 3' -P and 5' -PP resonances in the presence vs absence of the EntH (Figure 3.11). This finding is consistent with the comparatively weaker binding affinity that CoA has for EntH ($K_i = 225 \mu\text{M}$, a value ~ 45 fold higher than that of phenacyl-CoA).

As shown in Figure 3.5, the pantetheinyl arm of the bound phenacyl-CoA ligand “snakes” through a defined tunnel formed by opposing subunits of the EntH dimer. The relatively low B factors associated with the pantetheinyl atoms, compared to the B factors associated with the nucleotide atoms, suggest that the pantetheinyl arm is immobilized in the tunnel. As the tunnel defines a fixed distance between the surface portal and the catalytic site is no wonder that EntH shows no catalytic activity towards aroyl-(N-acetyl cystiene) thioesters. The cysteine side chain is not long enough to span the distance of the tunnel.

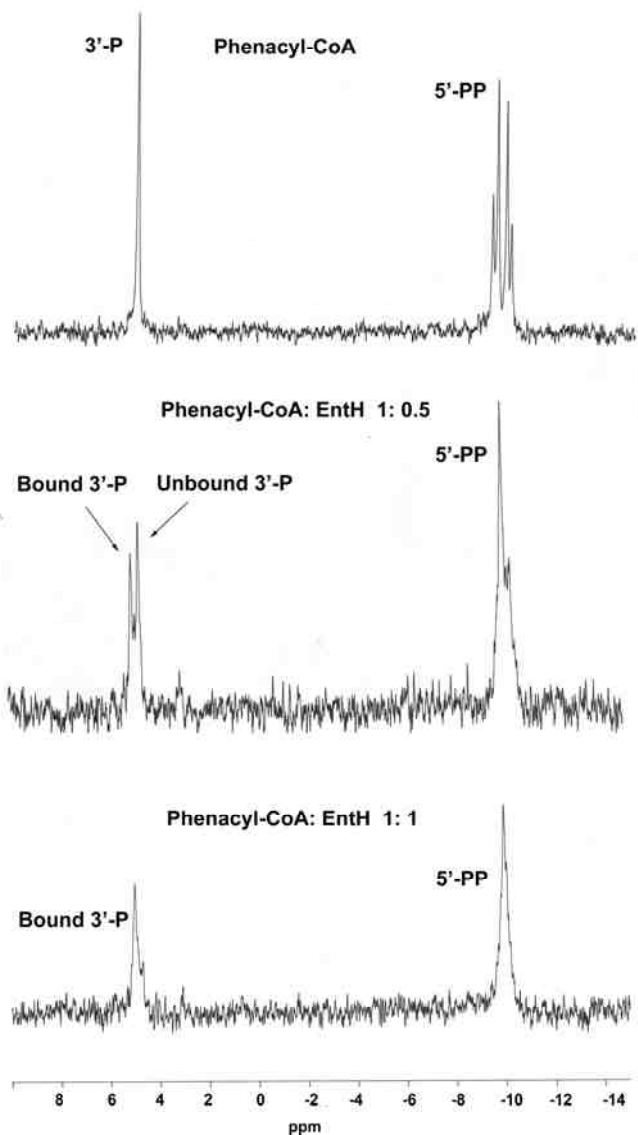


Figure 3.10 (Top) ^{31}P -NMR spectrum of a solution containing 1 mM phenacyl-CoA, 50 mM HEPES (pH 7.5) and 50 mM NaCl (singlet (3' -P) at 5.02 ppm and two doublets (5' -PP) at -10.02 to -9.60 ppm). **(Middle)** ^{31}P -NMR spectrum of a solution containing 1 mM phenacyl-CoA, 0.5 mM EntH, 50 mM HEPES (pH 7.5) and 50 mM NaCl. **(Bottom)** ^{31}P -NMR spectrum of a solution containing 1 mM phenacyl-CoA, 1.0 mM EntH, 50 mM HEPES (pH 7.5) and 50 mM NaCl.

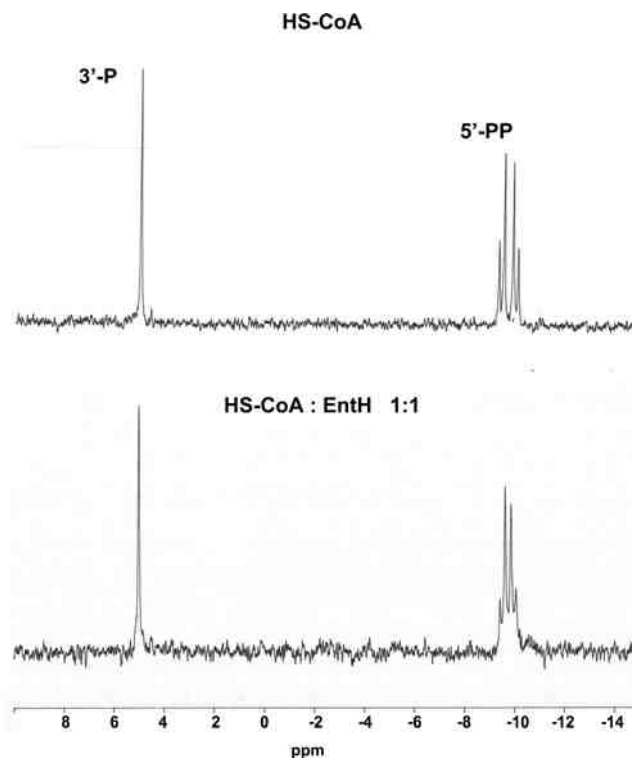


Figure 3.11 (Top) ^{31}P -NMR of a solution containing 1 mM CoA, 50 mM HEPES (pH 7.5) and 50 mM NaCl. **(Bottom)** ^{31}P -NMR of a solution containing 1 mM CoA, 1 mM EntH, 50 mM HEPES (pH 7.5) and 50 mM NaCl.

Indeed, as illustrated in the comparison of the pantetheinyl binding sites of EntH, hThem2 and *Arthrobacter* 4-hydroxybenzoyl-CoA (4-HBA-CoA) shown in Figure 3.12, the tunnel restricts the substrate range of hotdog thioesterases to aroyl/acyl-CoAs and aroyl/acyl-ACPs. Like EntH, hTHEM2 # and *Arthrobacter* 4-hydroxybenzoyl-CoA are not active with aroyl-(N-acetyl cystiene) thioesters .

Unlike EntH, which displays high thioesterase activity towards the aroylated holo-ACP domain of EntB, hTHEM2 (which is active with fatty acyl-CoAs) is not active with fatty acylated holo-ACPs (engineered from the human cytoplasmic fatty acid synthase ACP domain). Likewise, the *Arthrobacter* 4-HBA-CoA thioesterase is not active with 4-HBA-ACP (Hong Zhao, unpublished results). In order to gain insight into the structural basis for discrimination between CoA and ACP thioesters, we have examined the steric and electrostatic features of the surfaces surrounding the portals to the respective pantetheinyl binding tunnels (Figure 3.12). The shallow surface of EntH accommodates the EntB ArCP (model shown in Figure 3.7). The electrostatic surface of EntH (electropositive patch from Arg124 at 5.6 Å and Arg 91 at 4.4 Å) complements the 3'-phosphoryl group of the phenacyl-CoA ligand thus accounting for the downfield shift of the 3'-P resonance and the restriction of conformational freedom of both the 3'-P and 5'-PP groups that leads to the observed resonance line broadening (Figure 3.10).

In contrast, the liganded *Arthrobacter* 4-HBA-CoA thioesterase structure shows a well-defined binding site for the CoA nucleotide, which sits in crevice surrounded by electropositive walls (Arg 150, Arg 102 and Arg 135). Whereas its favorable interaction with the 5' -pyrophosphate/3' -phosphate groups of the CoA unit is evident, the topology of this binding site could not possibly complement the protein unit of an aroyl-ACP. The CoA nucleotide-binding site of hTHEM2 is not as specifically tailored, however it does hold distinct topological features, which suggest discrimination of an acyl-ACP substrate. Specifically, the entrance to the portal is circumscribed such that it is unlikely that the protein of an acyl-ACP could make an adequately close approach for productive binding of the phosphopantetheinyl arm.

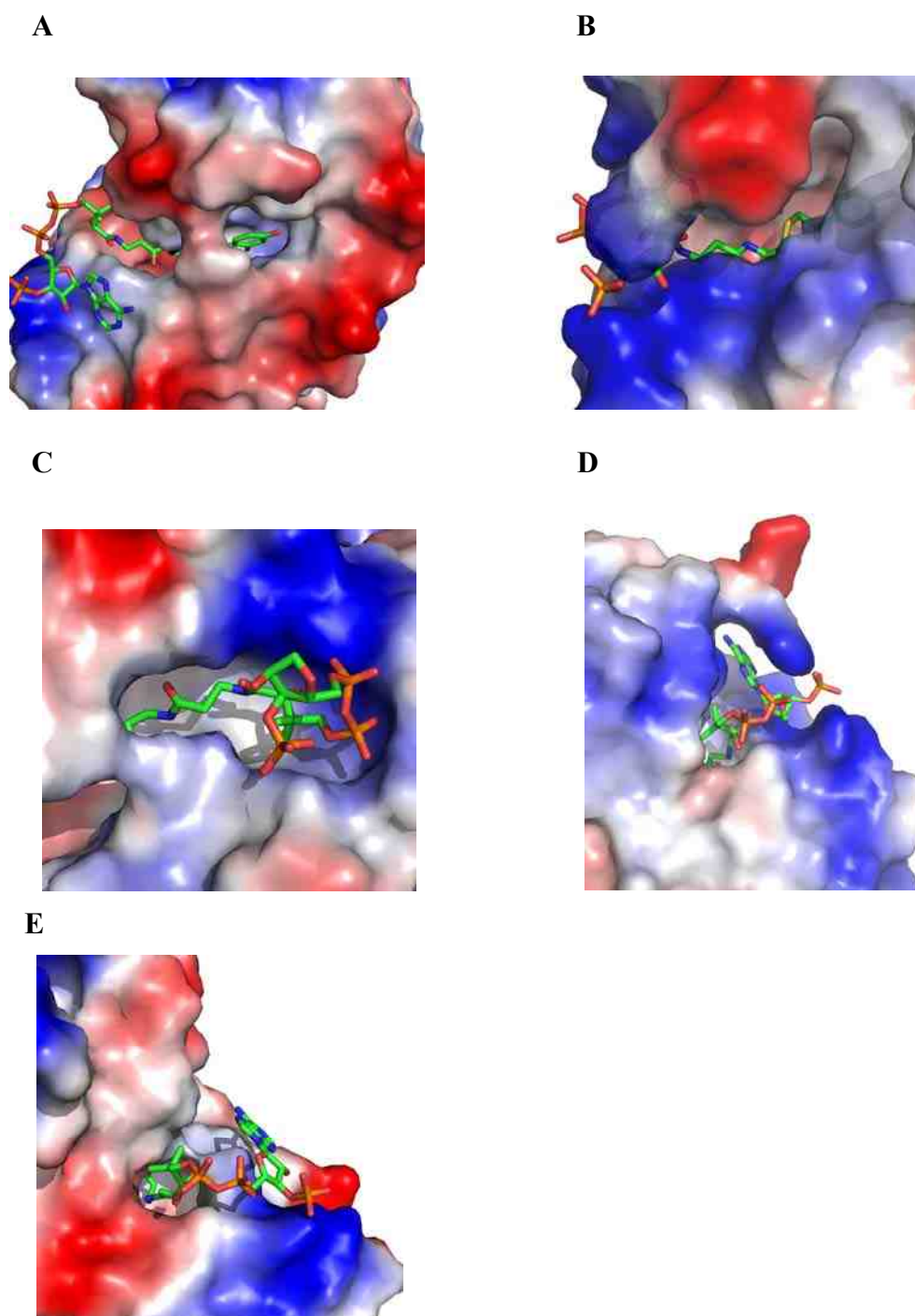


Figure 3.12 Topological and electrostatic features of the substrate binding site of (A) and (E) EntH bound with phenacyl-CoA, (C) hTHEM2 bound with undecan-2-one-CoA (the adenine moiety is not shown for clarity), (B) and (D) *Arthrobacter* 4-HBA-CoA

thioesterase bound with 4-hydroxyphenacyl-CoA. The electropositive area is represented in blue color, and the electron rich surface is represented in red color.

3.3.5. Comparative Analysis of the EntH Catalytic Site.

The hotdog thioesterases conserve fold, yet they do not conserve the residues that form the active site. One common feature to all hotdog thioesterases is a carboxylate residue, which mediates the hydrolysis reaction. Based on the location of this carboxylate residue on the catalytic scaffold, the family is divided into two clades. EntH is a member of the type AB clade. The thioesterases in this clade employ an Asp/Glu located on the central helix to mediate thioester hydrolysis. The prototype of the clade is the *Arthrobacter* 4-HBA-CoA thioesterase. Small alterations in the backbone conformation, which defines the catalytic site, as well as the residues that are positioned on the backbone, distinguish the thioesterases of this clade. EntH most closely resembles the *Arthrobacter* 4-HBA-CoA thioesterase. The catalytic site as viewed from the X-ray structures of EntH bound with phenacyl-CoA or 2,4-hydroxyphenacyl-CoA are shown in Figure 3.13.

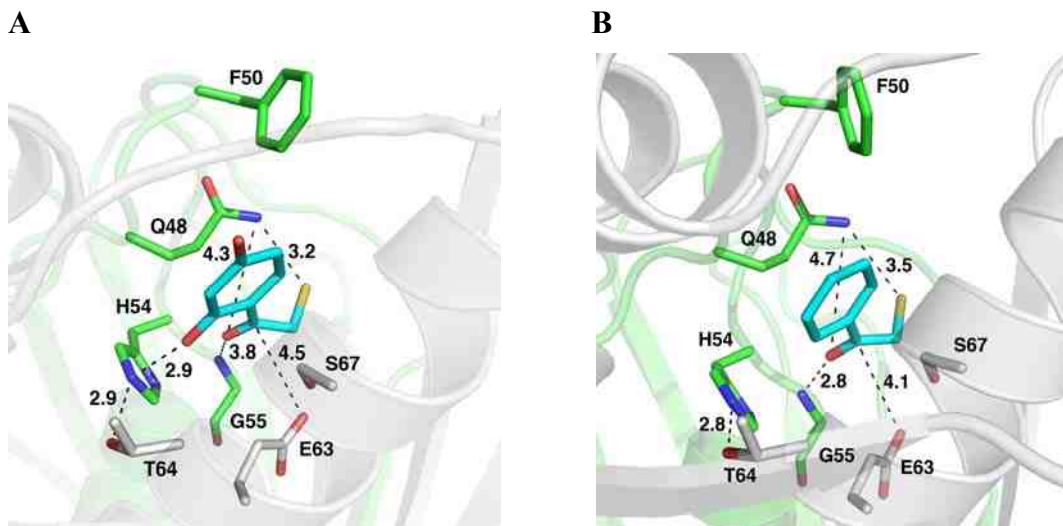


Figure 3.13 (A) and (B) The catalytic sites defined by the X-ray structures of EntH bound (A) 2,4-dihydroxyphenacyl-CoA and (B) phenacyl-CoA bound to EntH.

The Glu63 is positioned to collaborate with Ser67 in the positioning and activation of the water nucleophile. The structure also identifies residues that match the *Arthrobacter* 4-HBA-CoA thioesterase catalytic site residues. To access the contributions of these residues in EntH catalysis we measured the loss catalytic efficiency that occurs with the replacement by an amino acid that cannot perform the function. The results are reported in Table 3.2 along with the kinetic data reported by another laboratory. The data evidence the important roles played by Glu63, Gln4, His54 and Ser67 in catalysis. A model for the catalytic site is presented in Figure 3.14, which we posit is the most likely route to catalytic turn over.

Table 3.2 Steady-state constants for EntH wild type and mutant catalyzed hydrolysis of benzoyl-CoA or salicyl-CoA (if otherwise indicated) in 50 mM K+HEPES (pH 7.5 and

25 °C) and 1 mM DTNB^a.

Substrate	Benzoyl-CoA ^b			Salicylyl-CoA ^c		
Enzyme	k_{cat}	K_m	k_{cat}/K_m	k_{cat}	K_m	k_{cat}/K_m
WT	1.6±0.2	21±1	7.6x10 ⁴	4.6±0.2	176±7	2.7x10 ⁴
E63A	<10 ⁻⁴			NA ^d		
Q48A	<10 ⁻⁴			NA		
Q48N	ND ^e			(6.8±0.3)x10 ⁻²	729±19	93±2
H54A	<10 ⁻⁴			(1.9±0.1)x10 ⁻²	162±12	(1.2±0.1)x10 ²
T64A	ND			1.4±0.1	672±50	2.1x10 ³
S67A	ND			(5.7±0.2)x10 ⁻²	263±23	1.9x10 ²
M67A	ND			(7.8±0.1)x10 ⁻²	373±12	2.1x10 ²

^a The error limits shown reflect the fit to a data set and do not represent the standard deviation calculated from repetitive determination.

^b Enzyme activities were measured using Benzoyl-CoA.

^c Enzyme activities were measured using salicylyl-CoA as reported elsewhere (4).

^d No activities as reported elsewhere (4)

^e Not determined

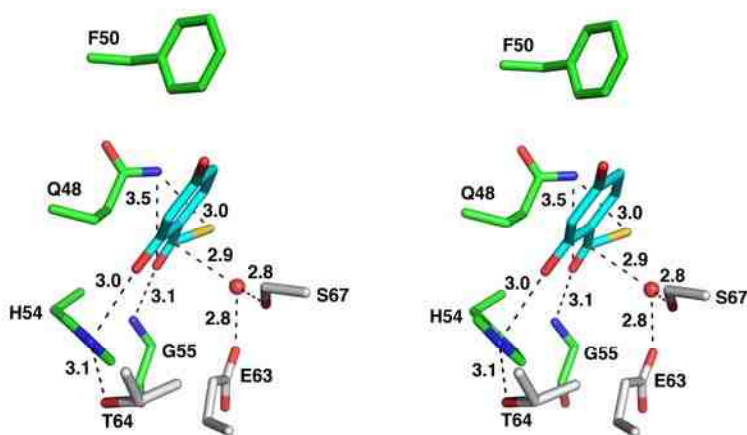


Figure 3.14 Stereoscopic representation of a model of 2,4-dihydroxybenzoyl-CoA bound to EntH.

3.4 References

1. Leduc, D., Battesti, A., and Bouveret, E. (2007) *J Bacteriol* **189**, 7112-7126
2. Chen, D., Wu, R., Bryan, T. L., and Dunaway-Mariano, D. (2009) *Biochemistry* **48**, 511-513
3. Badger, J., Sauder, J. M., Adams, J. M., Antonysamy, S., Bain, K., Bergseid, M. G., Buchanan, S. G., Buchanan, M. D., Batiyenko, Y., Christopher, J. A., Emtage, S., Eroshkina, A., Feil, I., Furlong, E. B., Gajiwala, K. S., Gao, X., He, D., Hendle, J., Huber, A., Hoda, K., Kearins, P., Kissinger, C., Laubert, B., Lewis, H. A., Lin, J., Loomis, K., Lorimer, D., Louie, G., Maletic, M., Marsh, C. D., Miller, I., Molinari, J., Muller-Dieckmann, H. J., Newman, J. M., Noland, B. W., Pagarigan, B., Park, F., Peat, T. S., Post, K. W., Radojicic, S., Ramos, A., Romero, R., Rutter, M. E., Sanderson, W. E., Schwinn, K. D., Tresser, J., Winhoven, J., Wright, T. A., Wu, L., Xu, J., and Harris, T. J. (2005) *Proteins* **60**, 787-796
4. Guo, Z. F., Sun, Y., Zheng, S., and Guo, Z. (2009) *Biochemistry* **48**, 1712-1722
5. Otwinowski, Z. a. M., W. (1997) *Methods in Enzymology* **276**, 307326
6. Emsley, P., Lohkamp, B., Scott, W. G., and Cowtan, K. (2010) *Acta Crystallogr D Biol Crystallogr* **66**, 486-501
7. Seeliger, D., and de Groot, B. L. (2010) *J Comput Aided Mol Des* **24**, 417-422
8. Kusunoki, M. (2002) *Tanpakushitsu Kakusan Koso* **47**, 1330-1332
9. Harris, M., and Jones, T. A. (2001) *Acta Crystallogr D Biol Crystallogr* **57**, 1201-1203

10. Zhuang, Z., Song, F., Zhao, H., Li, L., Cao, J., Eisenstein, E., Herzberg, O., and Dunaway-Mariano, D. (2008) *Biochemistry* **47**, 2789-2796
11. Song, F., Zhuang, Z., Finci, L., Dunaway-Mariano, D., Kniewel, R., Buglino, J. A., Solorzano, V., Wu, J., and Lima, C. D. (2006) *J Biol Chem* **281**, 11028-11038
12. Wei, J., Kang, H. W., and Cohen, D. E. (2009) *Biochem J* **421**, 311-322
13. Zhuang, Z., Song, F., Zhang, W., Taylor, K., Archambault, A., Dunaway-Mariano, D., Dong, J., and Carey, P. R. (2002) *Biochemistry* **41**, 11152-11160
14. Thoden, J. B., Holden, H. M., Zhuang, Z., and Dunaway-Mariano, D. (2002) *J Biol Chem* **277**, 27468-27476
15. Kunishima, N., Asada, Y., Sugahara, M., Ishijima, J., Nodake, Y., Miyano, M., Kuramitsu, S., and Yokoyama, S. (2005) *J Mol Biol* **352**, 212-228

CHAPTER 4.

KINETIC AND STRUCTURAL STUDIES OF HOTDOG FOLD

THIOESTERASE, E. COLI YDIH

4.1 Introduction

The *Escherichia coli* genome encodes 12 predicted hotdog fold proteins (1-6). One of these, YbdB (EntH) was described in Chapters 2 and 3. The paralog of EntH, YdiI will be described in this chapter. Hotdog thioesterase paralogs and even orthologs from divergent organisms typically share less than 25% sequence identity. Strikingly, alignment of the *E. coli* EntH and YdiI sequences revealed that they share 50% sequence identity and 75% sequence similarity. The *entH* gene is associated with enterobactin biosynthesis gene cluster and EntH optimizes the synthetic assemble of enterobactin by preferential removal of misacylated-holo-EntB. The *ydiI* gene, on the other hand, located between the *suf* operon, which encodes Fe-S biosynthetic enzymes, and the *ydiJ* gene, which encodes a predicted ketoreductase with unknown function (Figure 4.1).

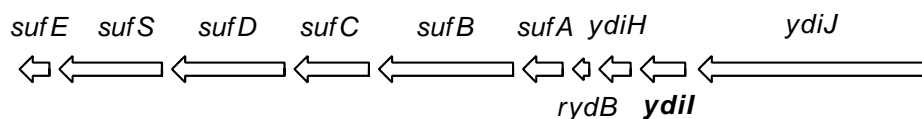


Figure 4.1 Co-location of *ydiI* gene with *suf* operon in *Escherichia coli* K-12 genome

So far, only limited insights into the cellular role of YdiI can be gleaned from the literature. Leduc *et al* showed that YdiI failed to supplement the function of EntH in

enterobactin biosynthesis. Although it might be surprising, taken into consideration of the high sequence similarity of the two proteins, it is suggestive that EntH and YdiI act on different substrates. The proximity of the *ydiI* gene to the *suf* operon, suggests a possible role in Fe-S cluster biosynthesis. However, deletion of the *ydiI* gene did not impede the biosynthesis of Fe-S clusters by the Suf system (7).

In this chapter, substrate screening and X-ray crystallographic structure determination of liganded YdiI were carried out with an aim of understanding the divergent functions of EntH and YdiI.

4.2 Experimental

4.2.1 YdiI Crystallization and X-ray Diffraction Data Collection

All three YdiI substrate analogue inhibitor complexes, YdiI (2,4-dihydroxyl phenacyl-CoA), YdiI (phenacyl-CoA) and YdiI (undecan-2-one-CoA), were crystallized by using the hanging-drop vapor diffusion method. Protein with a concentration of 10-15 mg/ml was combined with the substrate analogue 2,4-dihydroxylphenacyl-CoA, phenacyl-CoA or undecan-2-one-CoA to a final concentration of 2 mM.

YdiI (2,4-dihydroxylphenacyl-CoA), YdiI (phenacyl-CoA) and YdiI (undecan-2-one-CoA) complex crystals grew in 3-4 days at 18°C with typical overall dimensions 0.2 x 0.2 x 0.1 using 0.07-0.08 M sodium cacodylate trihydrate (pH 6.5), 0.18-0.20 M magnesium acetate tetrahydrate, 10-16% PEG8000, 20% glycerol. All crystals were cryoprotected by passing through 100% Paratone-N (Hampton Research) for approximately 5 minutes and freezing directly in a stream of N₂ gas at 100K. Diffraction data were collected to 1.89-1.97 Å resolution using an in-house Bruker AXS Proteum R

X-ray diffractometer outfitted with a MICROSTAR rotating-anode X-ray generator and Platinum 135 detector located at Boston University Department of Chemistry. Diffraction data were collected at 100K and indexed and scaled using Proteum 2 (Bruker AXS, Inc.).

4.2.2 YdiI (2,4-Dihydroxyphenacyl-CoA, Phenacyl-CoA or Undecan-2-one-CoA) Complex Structure Refinement.

All three YdiI-substrate analog complex structures were solved by molecular replacement using MOLREP of the CCP4 package using the prior YdiI structure (Protein Data Bank accession ID 1SBK) as a search model with all non-protein atoms removed. Iterative cycles of least squares refinement and manual model building were performed using PHENIX and the molecular graphics program COOT (8). The stereochemical quality of the final model was examined by PROCHECK (9). Figures were prepared with PyMol (DeLano Scientific), MOLSCRIPT, and POV-Ray (10).

4.2.3 Synthesis of Acyl-CoA analog inhibitors

2,4-dihydroxyphenacyl-CoA and phenacyl-CoA were synthesized as the following procedures. In N₂ atmosphere, 45 mg of 2-bromo-2',4'-dihydroxyacetophenone, 2-bromo-acetophenone or 2-bromo-undecanone was added dropwise within 1 hour to 30 mg CoA in 1.5 mL H₂O/THF (1:1) solution. PH was monitored by pH meter and maintained within a range of 7.2-8.0. After overnight incubation, reaction mixture was extracted with ethyl acetate twice. Then, crude product obtained after lyophilization was chromatographed to G-15 sephadex column. Purity of fractions was monitored by HPLC at dual wavelength 280 nm/260 nm. Fractions with

desirable purity were collected and combined to give target products. Structures were confirmed by $^1\text{H-NMR}$.

4.2.4 Enzymatic Assay

Steady-state kinetic methods were used to determine the k_{cat} and K_m for wild-type YdiI and EntH. The thioesterase activities was monitored at 412 nm ($\Delta\epsilon=13.6 \text{ mM}^{-1}\text{cm}^{-1}$) by coupling the reaction of 5,5'-dithio-bis(2-nitrobenzoic acid) (DTNB) with the CoA liberated from the hydrolysis of acyl-CoA substrates. All reactions were carried out in 50 mM HEPES buffer (pH 7.5) with 1 mM DTNB at 25°C.

The kinetic parameters of V_{max} and K_m were determined from initial velocity data, as measured as a function of substrate concentration, by using eq 1 and computer program KinetAsyst (IntelliKinetics)

$$V=V_{\text{max}}[A]/([A]+K_m) \text{ (eq. 1)}$$

Where [A] is the substrate concentration, V is the initial velocity, V_{max} is the maximum velocity, and K_m is the Michaelis constant. The reported error was computed for the data fitting. The k_{cat} was calculated from the ratio of V_{max} and the total enzyme concentration. The enzyme concentration was determined by using Bradford method.

The inhibition constant K_i for competitive inhibition was obtained by fitting the initial velocities of reactions of 4-hydroxybenzoyl-CoA (at varied concentration) in the absence or presence of inhibitor, 2,4-dihydroxyphenacyl, phenacyl-CoA or undecan-2-one-CoA to eq 2.

$$V=V_{\text{max}}[S]/[K_m(1+[I]/K_i)+[S]] \text{ (eq 2.)}$$

Where [I] is the concentration of the inhibitor, and K_i is the inhibition constant.

4.2.5 Sequence Alignment of YdiI, EntH, HThem2, PaaI and Arthrobacter sp. SU 4-HBA-CoA Thioesterase.

Protein sequences were from EntH (1VH9), *Arthr. sp* (*Arthrobacter sp* 4HBT 1Q4T), *E. coli* (*E. coli* PaaI 2FS2) and hTHEM2 (human thioesterase superfamily member 2 3F5O). Sequence accession numbers of the Swiss-Prot database are POA8Y8, Q04416, P76084 and Q9NPJ3 respectively. The secondary structure of YdiI is indicated above the alignment. The identical residues were colored in red and similar residues in yellow.

4.3 Results

4.3.1 Steady-state Kinetic Constants of YdiI Catalyzed Hydrolysis of Acyl-CoA and Acyl-holo-EntB.

Substrate screening of YdiI was carried out for the purpose of identifying acyl-CoA and acyl-holo-EntB thioesters with k_{cat}/K_M values that are considered “physiologically relevant”. The steady state kinetic parameters measured for the aryl- and acyl- CoAs with varied carbon chain lengths are summarized in Table 4.1. The aliphatic CoAs screened with YdiI include acetyl-CoA, propionyl-CoA, β -methylmalonyl-CoA, β -methylcrotonyl-CoA, hexanoyl-CoA (C_6), lauroyl-CoA (C_{12}), myristoyl-CoA (C_{14}), palmitoyl-CoA (C_{16}) and oleoyl-CoA (C_{18} , cis-9-ene). The YdiI k_{cat} values ranged from 0.1 to 0.7 s^{-1} and the K_M values ranged from 2-100 μM . In general, the catalytic efficiency of YdiI towards medium to long chain (C_{12-16}) acyl-CoAs is high enough ($k_{\text{cat}}/K_M > 1 \times 10^5 \text{ M}^{-1}\text{s}^{-1}$) to qualify these metabolites as potential physiological substrates. In contrast, EntH (YbdB) displays relatively low activity with these substrates ($k_{\text{cat}}/K_M \sim 1 \times 10^2 - 1 \times 10^3 \text{ M}^{-1}\text{s}^{-1}$).

Table 4.1 Steady-state kinetic constants for YdiI catalyzed hydrolysis of acyl-CoA and acyl-S-holo-EntB thioesters measured at pH 7.5 and 25 °C using DTNB spectrophotometric assay.

Substrate	k_{cat} (s^{-1})	K_{m} (μM)	$k_{\text{cat}}/K_{\text{m}}$ ($\text{M}^{-1}\text{s}^{-1}$)
Acetyl-CoA	$< 10^{-4}$		
Propionyl-CoA	0.21 ± 0.01	120 ± 10	1.7×10^3
Hexanoyl-CoA	0.30 ± 0.01	21 ± 1	1.4×10^4
Lauroyl-CoA	0.74 ± 0.01	2.2 ± 0.1	3.3×10^5
Palmitoyl-CoA	0.58 ± 0.01	1.9 ± 0.1	2.9×10^5
Myristoyl-CoA	0.62 ± 0.01	1.5 ± 0.2	4.1×10^5
β -methylmalonyl-CoA	0.67 ± 0.03	115 ± 7	5.8×10^3
Oleoyl-CoA	0.12 ± 0.00	1.3 ± 0.1	8.7×10^4
β -methylcrotonyl-CoA	0.50 ± 0.02	69.4 ± 0.4	7.3×10^3
4-HBA-CoA	5.2 ± 0.2	9 ± 1	5.9×10^5
Benzoyl-CoA	17.7 ± 0.7	25 ± 3	7.1×10^5
2,4-DHB-holo-EntB	0.071 ± 0.004	200 ± 20	3.6×10^3

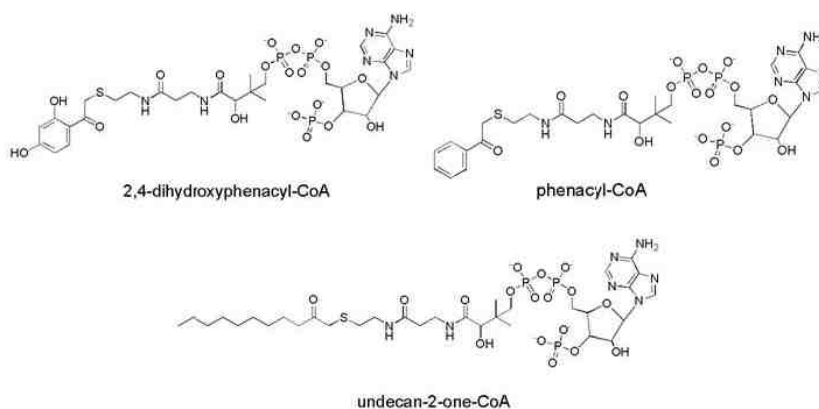
The k_{cat} values measured for the aroyl-CoA thioesters, 4-hydroxybenzoyl-CoA and benzoyl-CoA were 10-20 fold larger than those measured for the acyl-CoA thioesters, suggesting greater productive binding. The high substrate specificity constants measured for these substrates ($k_{\text{cat}}/K_{\text{M}} > 1 \times 10^6 \text{ M}^{-1}\text{s}^{-1}$) suggest that YdiI might function in the

hydrolysis of a yet to be discovered *E. coli* aroyl-CoA thioester metabolite. Unlike EntH (YbdB) ($k_{\text{cat}}/K_m = 1 \times 10^5 \text{ M}^{-1}\text{s}^{-1}$), YdiI does not catalyze the hydrolysis of 2,4-DHB-holo-EntB with high efficiency ($k_{\text{cat}}/K_m = 3.6 \times 10^3$).

4.3.2 Overall fold and oligomeric association of the YdiI-inhibitor complexes.

Based on the structures of the most active YdiI substrates inert substrate analogues were designed for co-crystallization with YdiI. The 2,4-dihydroxyphenacyl-CoA, phenacyl-CoA, and undecan-2-one-CoA (Figure 4.2) were prepared and shown to be strong competitive inhibitors with K_i values of $2.5 \pm 0.2 \mu\text{M}$, $7.8 \pm 0.4 \mu\text{M}$ and $12.2 \pm 0.5 \mu\text{M}$ respectively. The three dimensional structures of YdiI complexes of these inhibitors were determined at 1.9-2.0 Å resolution by our collaborators, Professor Karen Allen and Dr. Rui Wu at Boston University. As with EntH, YdiI is a dimer of dimer, associated back-to-back (i.e., sheet-to-sheet) (Figure 4.3).

A



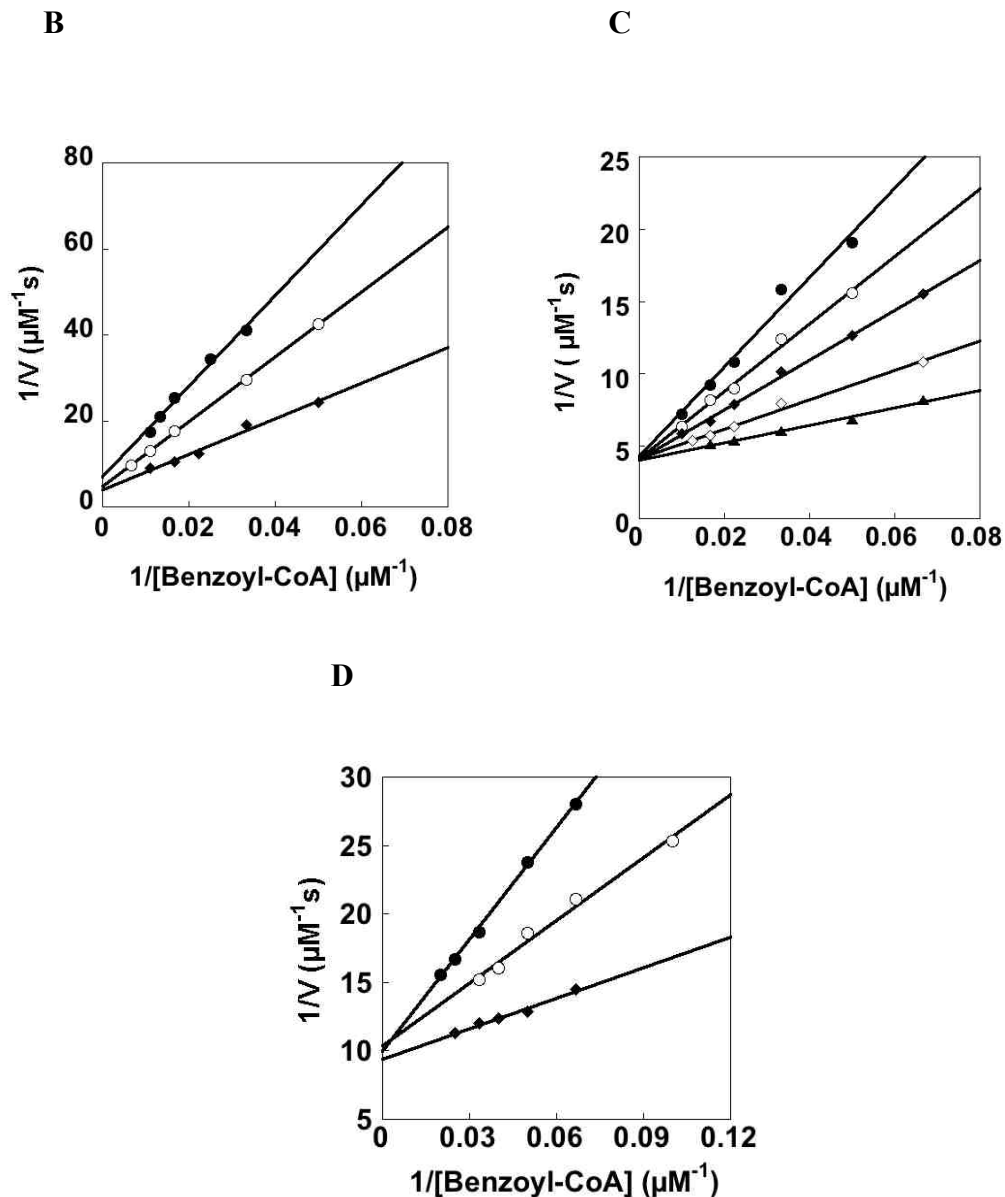
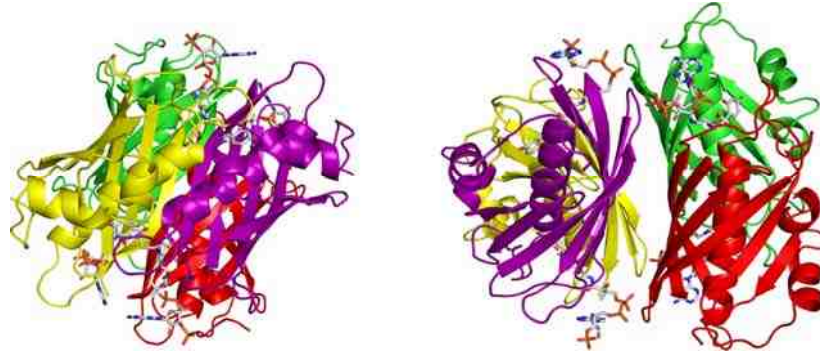
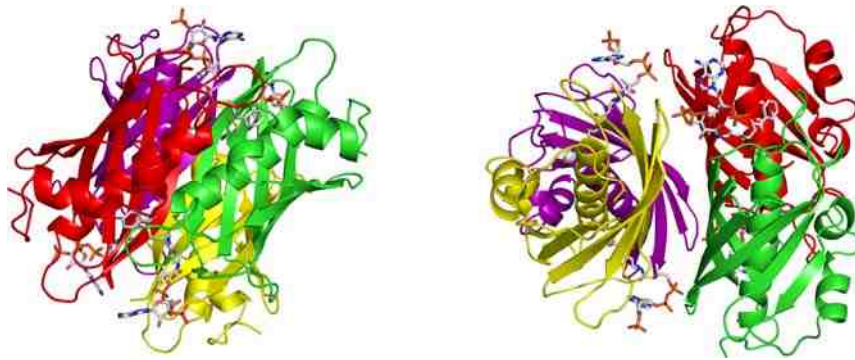


Figure 4.2 Chemical structures of the inert substrate analogs (A) and competitive inhibition plots for phenacyl-CoA (B), 2,4-dihydroxyphenacyl-CoA, (C) and undecan-2-one-CoA (D). YdiI catalyzed hydrolysis of benzoyl-CoA was monitored by DTNB assay using 0.0645 μM YdiI, (15 to 200 μM) benzoyl-CoA, (5 to 20 μM) phenacyl-CoA, or (2.5 to 10 μM) 2,4-dihydroxyphenacyl-CoA or (15 to 30 μM) undecan-2-one-CoA.

A



B



C

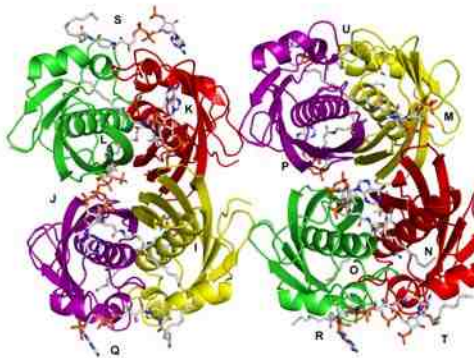


Figure 4.3 The asymmetric unit of YdiI bound with **(A)** 2,4-dihydroxyphenacyl-CoA **(B)** phenacyl-CoA and **(C)** undecan-2-one-CoA. Carbon atoms of ligands are colored white, and all oxygen, nitrogen, sulfur and phosphorus atoms are colored red, blue, yellow and magenta respectively.

The structures of the YdiI (2,4-dihydroxy phenacyl-CoA) and YdiI (phenacyl-CoA) complexes superimpose well with the structure of *apo* YdiI (PDB entry 1SBK) (α carbon atoms RMSD value of 0.58 Å and 0.56 Å, respectively). Compared with the *apo* YdiI structure, the structure of the YdiI (undecan-2-one-CoA) complex has several loops which adopt different conformations (RMSD value of 1.14-1.18 Å). The conformational difference between *apo* YdiI and YdiI (undecan-2-one-CoA) is most likely a direct result of the accommodation of long aliphatic chain of undecan-2-one-CoA in the active site.

4.3.3 The 2,4-Dihydroxyphenacyl-CoA, Phenacyl-CoA and Undecan-2-one-CoA ligands.

The electron density map of 2,4-dihydroxyphenacyl-CoA, phenacyl-CoA or undecan-2-one-CoA bound at the interface of two subunits of YdiI are presented in Figure 4.4. The electron density maps of 2,4-dihydroxyphenacyl-CoA and phenacyl-CoA are well defined (Figure 4.4 A and B), whereas the nucleotide moiety and the S-CH₂-C=O unit of the undecan-2-one-CoA ligands showed poor electron density (Figure 4.4 C).

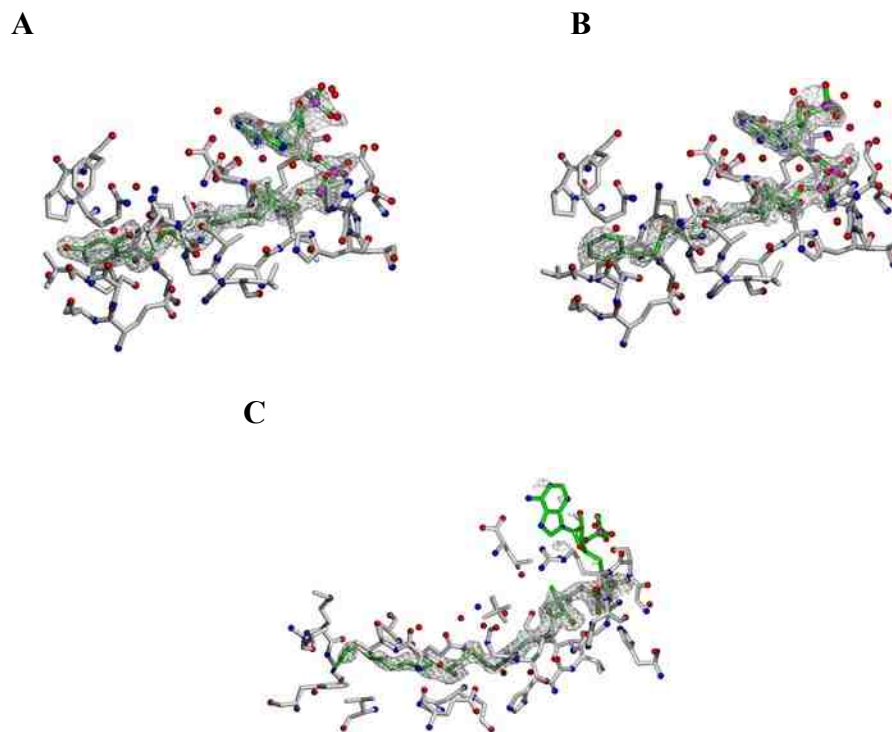


Figure 4.4 The electron density map of (shown in grey) of YdiI bound (A) 2,4-dihydroxy phenacyl-CoA, (B) phenacyl-CoA and (C) undecan-2-one-CoA ligands. The density maps are contoured at the level of 1σ .

4.3.4 The Substrate Binding Site of YdiI.

The ligand phosphopantetheine arm and the phenacyl/decanyl unit in all three YdiI liganded structures are wedged at the interface between the two monomers. The nucleotide binding site includes one monomer from the opposing dimer. The catalytic scaffold consists of the connecting loop and N-region of the central α -helix of one monomer and the midsection of the central helix from the opposing monomer (Figure 4.5 A).

(a) The Substrate Binding Site of YdiI (2,4-Dihydroxyphenacyl-CoA) Complex

A stereoscopic view of the 2,4-dihydroxyphenacyl-CoA binding site residues is shown in Figure 4.5 A and the electrostatic features of the binding site is shown in Figure 4.6 B, respectively. In the liganded YdiI dimer, the adenine 3'-monophosphate and 5'-pyrophosphate units bind to the protein surface at the entrance of the active site, while the pantetheine arm and acyl tail snake through a channel formed at the subunit interface.

The CoA 3'-phosphate is within hydrogen bond distance of the side chain of Arg91 of one monomer, whereas the 5'-pyrophosphate forms extensive hydrogen bonds with backbone amide NH and/or side chain of His106, Ser109, Arg110 and His111 from one subunit of the opposing dimer. The methyl groups of the pantetheine unit interact with several hydrophobic residues, Leu136, Val81, Leu83 and Ala134 from one monomer and Val90 from the opposing. In addition, two pantetheine unit amide NHs form hydrogen bonds with the backbone carbonyl groups of Gly82 on one monomer and His89 of the opposing monomer.

The 2,4-dihydroxyphenacyl group is positioned in a hydrophobic binding pocket located between the mid-section of central α helix from one monomer and a small loop (Gln48-Leu52) from the opposing monomer. The key side chains involved in binding the 2,4-dihydroxyphenacyl unit include the residues Met15, Val21, Val68, Ser64, Ser67 and Glu63 (from one monomer) and Phe50, Pro49, Leu52, His54, Gly55 and Gln48 (from the opposing monomer). In addition, the 2,4-dihydroxyl group lies within hydrogen bond distance of two ordered water molecules. The catalytic scaffold is very similar to that

observed in the *Arthrobacter* 4-HBA-CoA thioesterase and EntH. Among these active site binding residues are the highly conserved residues Ser64, Ser67, Glu63, His54, Gly55 and Gln48.

(b) The Substrate Binding Site of YdiI (Undecan-2-one-CoA) Complex

In the YdiI (undecan-2-one-CoA) complex, the pantetheine unit interacts with the enzyme through similar interactions as in YdiI (2,4-dihydroxyphenacyl-CoA) complex (Figure 4.6 B). The aliphatic chain acyl moiety binding site is formed by Glu63, Ser67, Ser64, Val68, Phe28, Leu12, Met15, Val21 and Asn19 of one monomer and Pro49, His54, Gln48 and Gly55 of the other monomer. The binding site is a long hydrophobic tunnel, which accounts for the tight binding of long chain fatty acyl-CoAs. The arrangement of catalytic residues around the thioester unit is similar to that observed with the YdiI (2,4-dihydroxyphenacyl-CoA) complex (Figure 4.6 B). Consistent with the type AB clade the YdiI Glu64 (monomer II)-Gln48 (monomer I) pair is positioned on opposite sides of thioester C=O. The general base/nucleophilic Glu64 is ~ 3.5 Å from carbonyl carbon, while Gln48 is within hydrogen bond distance to carbonyl oxygen. Gly55 from monomer II orients and polarizes the C=O group via hydrogen bonding (2.9 Å).

The accommodation of the long aliphatic chain of the undecan-2-one-CoA ligand site does not noticeably alter the positioning of the pantetheine arm however, nucleotide unit is not associated at the same site on the surface surrounding the portal to the active site.

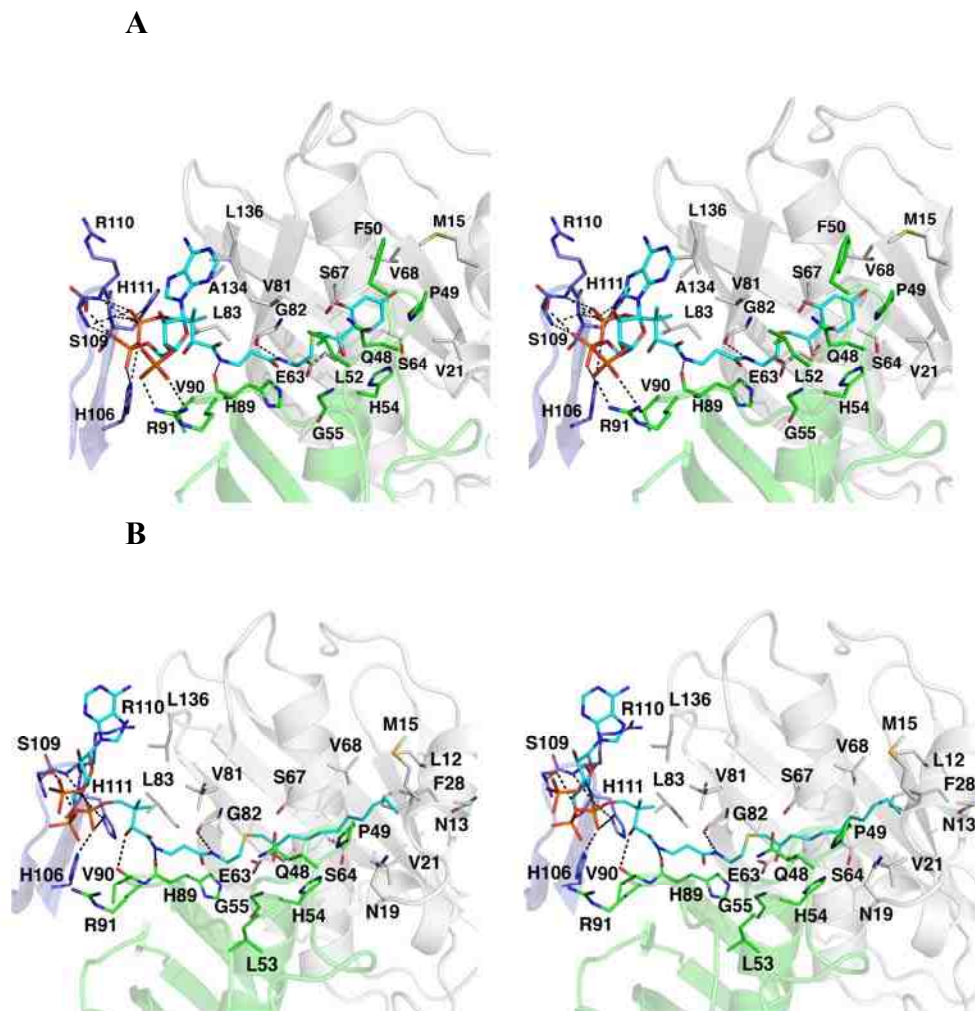


Figure 4.5 Stereoscopic representation of atomic interactions associated with the YdiI substrate-binding site. Key interactions of **(A)** 2,4-dihydroxyphenacyl-CoA and **(B)** undecan-2-one-CoA (ligand P in Figure 4.2 C) bound to YdiI. Atomic coloring is as follows: blue for nitrogen atom, red for oxygen atom, yellow for sulfur atom and brown for phosphorus atom. The two colors (white and light green) of the ribbon trace correspond to the two molecules within the dimer. The light blue ribbon corresponds to a subunit of a neighboring dimer that flanks the nucleotide binding site. The carbon atoms are colored according to which subunit the residue is located on, and the carbon atoms of the ligand are shown in cyan.

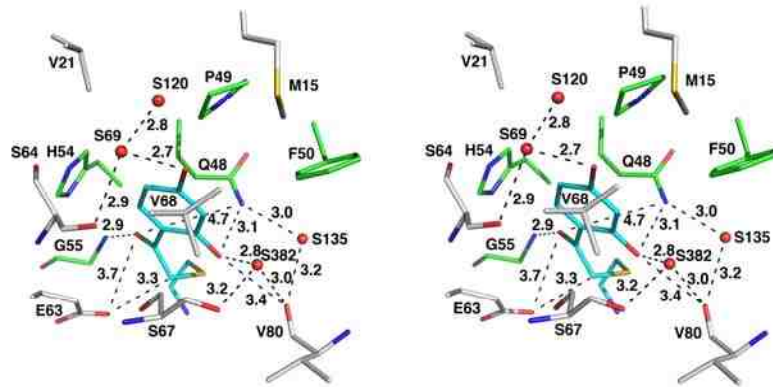
4.3.5 The Substrate Model.

Based on the investigation of the substrate binding interactions defined by YdiI (2,4-dihydroxyphenacyl-CoA) complex, YdiI (phenacyl-CoA) complex and YdiI (undecan-2-one-CoA) complex structures, the decanoyl-CoA is modeled in place of the undecan-2-one-CoA (Figure 4.6 C). The resulting model shows the substrate thioester bond carbonyl oxygen to be within hydrogen bond distance of the backbone nitrogen of Gly55 (3.1 Å), and with Gln48 amine NH (3.4 Å). Orientation of Glu63 carboxylate group allows a direct nucleophilic attack on the thioester bond carbonyl carbon.

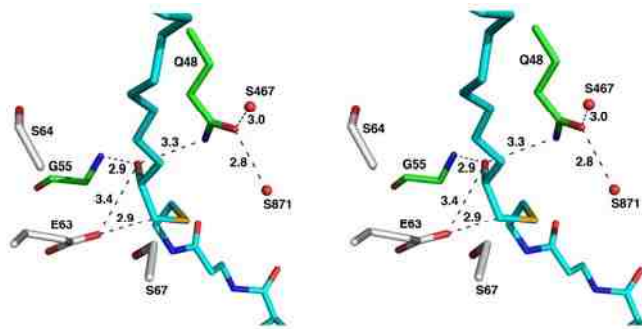
These key active site residues of YdiI, *E. coli* EntH, *Arthrobacter* sp 4-hydroxybenzoyl: CoA thioesterase, *E. coli* phenylacetyl-CoA thioesterase PaaI and human thioesterase superfamily member 2 (hTHEM2) are overlaid in Figure 4.7. Among these residues, Gly55, His54, Glu63 and Gln48 are conserved. Gly55 in YdiI is conserved as Gly 55 in *E. coli* EntH, Gly65 in *Arthrobacter* 4-HBA-CoA thioesterase, Gly53 in *E. Coli* PaaI and Gly57 in hTHEM2. His54 is conserved as His54 in *E. coli* EntH, His64 in *Arthrobacter* 4-HBA-CoA thioesterase, His52 in *E. coli* PaaI and His56 in hTHEM2. Glu63 and Gln48 are conserved as Glu63 and Gln48 in *E. coli* EntH, Glu73 and Gln58 in *Arthrobacter* 4-HBA-CoA thioesterase and conservatively substituted with Asp65 and Asn50 in hTHEM2 and Asp61 and Asn46 in *E. Coli* PaaI. The residues not conserved are Ser67 and Ser64. Ser67 is conserved as Ser67 in *E. coli* EntH and conservatively substituted with Thr77 in *Arthrobacter* 4-HBA-CoA thioesterase, and with Thr69 in hTHEM2, but replaced with Ala65 in *E. Coli* PaaI. Ser64 is conservatively

substituted with Thr64 in *E. coli* EntH, Thr62 in *E. Coli* PaaI, however replaced with Met74 in *Arthrobacter* 4-HBA-CoA thioesterase and Asn66 in hTHEM2.

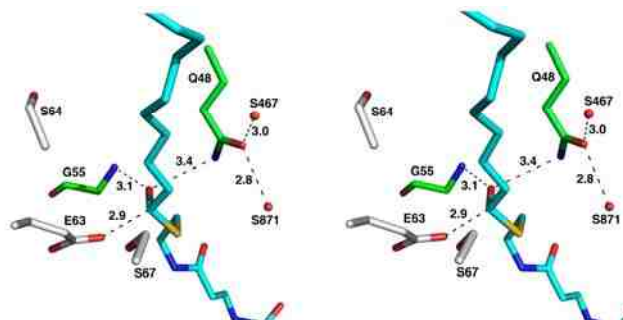
A



B



C



(magenta), hTHEM2 (brown) and *E. coli* PaaI (cyan). **(B)** Structure based sequence alignment of YdiI with other hot dog fold thioesterases created with ESPript. Protein sequences were from YdiI (YdiI (undecan-2-one-CoA) complex chain E&F), EntH (1VH9), *Arthr. sp* (*Arthrobacter* sp 4HBT 1Q4T), *E. coli* (*E. coli* PaaI 2FS2) and hTHEM2 (human thioesterase superfamily member 2 3F5O). Sequence accession numbers of the Swiss-Prot database are P77781, POA8Y8, Q04416, P76084 and Q9NPJ3 respectively. The secondary structure of YdiI, which is mainly defined by the analysis of the structure using DSSP program, is indicated above the alignment. The identical residues were colored in red and similar residues in yellow.

4.4 Discussion

Substrate screening and co-crystallization studies with inert substrate analogs of both EntH (YbdB) and YdiI allow pair-wise comparisons in terms of substrate preferences, specificities, catalysis mechanisms and structural bases of these aspects.

Substrate screening studies of YdiI was carried out and summarized in Table 4.1. In good agreement with its high sequence similarities to EntH, YdiI has a broad substrate specificity. Both medium to long chain fatty acyl and hydroxyl- or nonhydroxyl- aryl-CoAs can be efficiently hydrolyzed by YdiI with steady state kinetic parameters within the range of “physiologically relevant”, in contrast to the slow hydrolysis of short chain acyl-CoAs. The high k_{cat}/K_M values measured for the aryl-CoAs are common to both YdiI and EntH. Overall, YdiI has at less 10 fold (k_{cat}/K_M) higher efficiency towards hydrolysis of the aroyl/acyl-CoA thioester substrates than EntH. In contrast, EntH is 1000-fold more active towards hydrolysis of the hydroxybenzoyl-EntB, which explains why YdiI cannot

assume the function of EntH in the gene knockout mutant. The physiological role of YdiI might instead be to regulate the ratio of acyl-CoA to CoA in the cytoplasm (5,11).

As shown in Figure 4.7 C, YdiI and its clade-like members, including EntH, 4-HBA-CoA and hThem2, recruit similar residues with comparable orientation for construction of their respective unliganded active sites, which accounts for their similar preferences towards aryl-CoAs. For YdiI and EntH pair, in particular, their active sites are delineated by almost identical residues with the exception of position 64 (Threonine in the case of EntH and Serine in the case of YdiI).

As addressed in Chapter 3, EntH catalyzed hydrolysis is proposed to occur via a general base mechanism, in which the carboxylate side chain of Glu63 functions as the general base. In the case of YdiI, based on the liganded active site architecture of YdiI (undecan-2-one-CoA) complex, a modeled decanoyl-CoA is docked into the reaction center for the purpose of elucidating interactions relevant to reaction mechanism. Gly55 is 3.1 Å from the carbonyl oxygen of the ligand, while Gln48, oriented on opposite face of the ligand, is 3.4 Å away from the same oxygen atom. Similar interactions are also observed in liganded EntH active site, suggesting the well-conserved functions for these two residues. Notably, variation in liganded orientation in YdiI active site to that of EntH is the carboxylate of Glu73. In contrast to >4 Å distance from carbonyl carbon of the ligand, Glu 73 carboxylate is within nucleophilic attack range (2.9 Å). At this point, it is untenable to point out that these two proteins may catalyze hydrolysis of thioester in distinct mechanism. However, undoubtedly, detailed mechanistic studies in these directions would be valuable in understanding the divergence of hotdog thioesterases.

Especially, these two hotdog fold proteins are highly homologous in sequence and structure.

Comparisons of the active sites of YdiI and EntH provide little insight into the biased substrate range of the two proteins. Thus, we turn to examine the binding site for adenine moiety of the two proteins. As described in Chapter 3, EntH has an open and solvent exposed binding surface of adenine moiety of CoA ligand, thus this part of EntH bound 2,4-dihydroxyphenacyl-CoA, and phenacyl-CoA ligands show low electron density. In contrast, adenine moieties of YdiI bound 2,4-dihydroxyphenacyl-CoA, and phenacyl-CoA ligands show high electron densities. In the light of the binding of 2,4-dihydroxyphenacyl-CoA, and phenacyl-CoA ligands in YdiI, together with its kinetic constants, we postulate that YdiI efficiently targets the CoA thioesters by accommodating their nucleotide unit on protein surface through several favorable interactions. These interactions are absent in liganded EntH structures, which might explain its lack of discrimination between CoA and holo-EntB thioesters.

4.5 References

1. Dillon, S. C., and Bateman, A. (2004) *BMC Bioinformatics* **5**, 109
2. Zhuang, Z., Song, F., Martin, B. M., and Dunaway-Mariano, D. (2002) *FEBS Lett* **516**, 161-163
3. Leduc, D., Battesti, A., and Bouveret, E. (2007) *J Bacteriol* **189**, 7112-7126
4. Angelini, A., Cendron, L., Goncalves, S., Zanotti, G., and Terradot, L. (2008) *Proteins* **72**, 1212-1221

5. Zhuang, Z., Song, F., Zhao, H., Li, L., Cao, J., Eisenstein, E., Herzberg, O., and Dunaway-Mariano, D. (2008) *Biochemistry* **47**, 2789-2796
6. Chen, D., Wu, R., Bryan, T. L., and Dunaway-Mariano, D. (2009) *Biochemistry* **48**, 511-513
7. Butland, G., Babu, M., Diaz-Mejia, J. J., Bohdana, F., Phanse, S., Gold, B., Yang, W., Li, J., Gagarinova, A. G., Pogoutse, O., Mori, H., Wanner, B. L., Lo, H., Wasniewski, J., Christopolous, C., Ali, M., Venn, P., Safavi-Naini, A., Sourour, N., Caron, S., Choi, J. Y., Laigle, L., Nazarians-Armavil, A., Deshpande, A., Joe, S., Datsenko, K. A., Yamamoto, N., Andrews, B. J., Boone, C., Ding, H., Sheikh, B., Moreno-Hagelseib, G., Greenblatt, J. F., and Emili, A. (2008) *Nat Methods* **5**, 789-795
8. Emsley P, C. K. (2004) *Acta Crystallographica Section D-Biological Crystallography* **60**, 2126-2132
9. 5. Laskowski, R. A., MacArthur, M.W., Moss, D.S. & Thornton, J.M. (1993) *J. Appl. Crystallogr* **26**, 283-291
10. Kraullis, P. J. (1991) *J. Appl. Crystallogr.* **24**, 946–950
11. Willis, M. A., Zhuang, Z., Song, F., Howard, A., Dunaway-Mariano, D., and Herzberg, O. (2008) *Biochemistry* **47**, 2797-2805

CHAPTER 5

HUMAN BROWN FAT INDUCIBLE THIOESTERASE *VARIANT 2* (BFIT2): CELLULAR TARGETING AND CATALYTIC FUNCTION

5.1 Introduction

Two forms of adipose tissue exist in humans, including white adipose tissue (WAT) and brown adipose tissue (BAT). WAT serves as the sites for storing neutral lipid as energy resources, while BAT undergoes enhancement in fuel pool consumption for heat production in response to low temperature and fat-rich diet (termed collectively as adaptive thermogenesis) (1-3). This energy dissipation in form of heat is mainly promoted by uncoupling protein 1 (UCP1), an integral membrane protein in the mitochondrion of BAT (4-6). UCP1 enables the conversion of energy to heat release through a process involving proton leaking across mitochondrial inner membrane of BAT, a feature that causes distinction between BAT and ordinary tissues, which store energy gained from fatty acid metabolism in the form of ATP by oxidative phosphorylation. (7-8). Owing to these features, UCP1 is a unique protein that serves as a focus of the research on adaptive thermogenesis transition of BAT.

Awareness has arisen that although essential, UCP1 together with its homologues does not meet the demands of drastic energy expenditure profile change as adaptive thermogenesis. Supporting observations have come from differential gene expression analyses, which reveal that a remarkable metabolic shift takes place concomitantly with heat production during adaptive thermogenesis (9-10). Under cold exposure/food

restriction conditions, expression activation occurs to genes encoding proteins for *de novo* fatty acid synthesis, β -oxidation, lipogenesis, glucose transportation and catabolism. Clearly, effective and lasting adaptive thermogenesis relies on the coordination of an array of biological pathways. Interestingly although paradoxically, a simultaneous amplification of fatty acid synthesis and β -oxidation takes place during thermogenesis, events that further underscores the necessity of a finely tuned and systematic regulation in adaptive thermogenesis.

Among many of the genes induced in cold/fat-diet is an acyl-CoA thioesterase, known as brown fat inducible thioesterase (BFIT). The unique characteristics of BFIT is demonstrated by its domain organization, which is comprised of tandem “hot dog” thioesterase domains and a steroidogenic acute regulatory protein related lipid transfer (START) domain from N- to C- terminus (Figure 5.1 A). This multi-domain protein and cytoplasmic acetyl-CoA hydrolase (CACH) are categorized as thioesterase START group for their similarity in domain organization (11-12). Two splicing variants of BFIT exist in human, namely BFIT1 and BFIT2, whereas mouse only has BFIT2 isoform. The two isoforms only differ at their C-terminal START domains, where only BFIT2 has the α 4-helix, which functions as a “gatekeeper” in facilitating lipid binding of START superfamily. Recent reports showed that tandem thioesterase domains within a single peptide indicate the occurrence of “half-of-sites” regulation, involving cooperative association of two domains into functional dimer, which only half of the active sites is reactive (13-15). When recruited in a special event, the thioesterase domains of BFIT may have the influential impact on cellular concentration of fatty acid, CoA species profiling, and as a consequence regulation of fatty acid synthesis and metabolism,

although its thioesterase activity has not been rigorously investigated. Due to the presence of START domain (STARD14) in BFIT, it is tempting to speculate that BFIT potentially involves in modulating lipid metabolism. Furthermore, previous studies showed its potential connection with metabolic syndromes, including obesity and diabetes (16). Taken together, the information gathered thus far suggests that BFIT2 plays an intriguing, complicated but largely unclear role in adaptive thermogenesis.

In the studies described below, we have characterized of *Homo sapiens* BFIT variant 2 (herein simplified as BFIT2) expressed using an eukaryotic expression system, including the identification of posttranslational modification at the N-terminal region of BFIT2, its subcellular location, substrate range of its thioesterase hydrolysis. As part of comprehensive investigation of the structure and function of BFIT, these studies have led to a more precise assessment of its cellular function in thermogenic transition of BAT.

5.2 Experimental

5.2.1 Materials

HEK293T/17 cell line, Dulbecco's modified eagle's medium, fetal bovine serum and Trypsin-EDTA reagent were purchased from American Type Culture Collection (ATCC) (Manassas, VA). Transfection and stable cell line selection reagents, including OptiMEM, Zeocin, Lipofectamine 2000 and phosphate buffered saline (PBS), were purchased from Invitrogen (Carlsbad, CA). Mitochondria isolation kit for cultured cells and direct immunoprecipitation kit were purchased from Pierce (Rockford, IL). Mammalian expression vectors, including pcDNA4 HisMax TOPO, pcDNA3.1 CT-GFP TOPO and pcDNA3.1 NT-GFP TOPO, were from Invitrogen. Anti-BFIT2 antibody was

purchased from Abcam (Cambridge, MA), anti-mouse mHSP70 and anti-mouse AIF antibodies were purchased from Thermo Scientific. Other antibodies, including mouse anti-Xpress, anti-HisG, anti α -tubulin, and anti-mouse HRP, were purchased from Invitrogen. Nuclear extraction kit was purchased from Active Motif (Carlsbad, CA). Cellular compartment dyes, including mitotracker Red CMXRos, WGA/Alexa Fluor 594 conjugate and Hoechst 33342, were purchased from Invitrogen. All other reagents were purchased from Sigma-Aldrich unless otherwise indicated.

5.2.2 Expression Vectors and Transfection of HEK293T/17 Cells.

PCR was used to amplify the full length, Δ N-terminal 34 amino acids and N terminal (1-42) amino acids only of BFIT2 cDNAs. After addition of 3' end adenine hangover by platinum *Taq* polymerase (Invitrogen), the cDNAs were cloned into TOPO expression vectors. The resulting plasmids, including *BFIT2/pcDNA4 HisMAX*, *BFIT2 (Δ N 33)/pcDNA4 HisMAX*, *BFIT2 START/pcDNA4 HisMAX*, *BFIT2/pcDNA3.1 NT-GFP*, *BFIT2/pcDNA3.1 CT-GFP* and *BFIT2 (N 1-42)/pcDNA3.1 CT-GFP*, were used to transfect HEK293T/17 cells. For transient transfection, HEK293T/17 cell lines were plated in 24-well or 6-well plates and transfected in the following day at 80% confluency according to the manufacturer's instruction. For *BFIT2/pcDNA4 HisMAX* stably transfected cell lines, selections were carried out in culture medium with 500 μ g/mL Zeocin.

5.2.3 Immunoblot Analysis.

Transfected or untransfected HEK293T/17 whole cell extract lysates were

prepared by using nuclear extraction kit or an immunoprecipitation lysis/wash buffer (Pierce) following manufacturer's instructions. Total protein concentration was measured by Bradford assay. The lysate was separated by 12% polyacrylamide gel before being transferred to nitrocellulose membrane, which was then treated with primary antibody at 4°C for overnight. After extensive washing, membrane was treated with secondary antibody. Finally, imaging was carried out using enhanced chemical luminescent reagents (Sigma-Aldrich).

5.2.4 Immunoprecipitation of Recombinant BFIT2 from HEK293T/17 Stably Transfected Cells.

Anti-Xpress or anti-HisG antibody covalently immobilized agarose beads were prepared following the manufacturer's instruction. In brief, 10-50 µg of primary antibody was incubated with 20-100 µL of an agarose bead slurry in the presence of NaCNBH₃ at room temperature for at least 2 hours. Uncoupled antibody was removed by extensive washing and uncoupled reactive sites in agarose beads were quenched by incubation in solution with primary amine and NaCNBH₃. The lysate of BFIT2/pcDNA4 HisMAX stably transfected HEK cell was incubated with antibody immobilized agarose beads at 4°C for overnight. Unbound proteins were removed in washing steps. Finally, the antigen, BFIT2, was eluted by washing agarose beads with elution buffer (10 mM Tris with 100 mM glycine pH 7.4).

5.2.5 Confocal Imaging of GFP Fusion Protein.

HEK293T/17 cells were cultured on the coverslip and transfected with GFP fusion expression plasmids. At 12-24 hours post-transfection, cells were fixed using 4% formaldehyde in PBS buffer at 37°C for 15 mins. Subsequently, the fixed cells on coverslip were incubated with 1 μM Hoechst 33342, and either 50 μM Mitotracker Red CMXRos or 5 μg/mL WGA/Alexa Fluor 594 conjugate at room temperature for 20 mins. Finally, the coverslip was rinsed with PBS, mounted onto slides, and subject to confocal microscope imaging analysis (Cancer Center Fluorescence Microscopy Facility, University of New Mexico School of Medicine).

5.2.6 Mitochondrial Fraction Isolation from HEK293T/17 Cells.

Mitochondria isolation was performed following the manufacturer's instructions. In brief, *ca.* 2×10^7 of untransfected or BFIT2 stably transfect HEK293T/17 cells was pellet by centrifuge and resuspended in mitochondria isolation reagent A. Cells were lysed by a Dounce tissue grinder. Following the addition of Mitochondria isolation reagent C to the cell lysate, Cytosolic and mitochondrial fractions were separated by centrifuge. The resulting mitochondria pellet was washed once with reagent C and resuspended in suitable buffer for down stream processing.

5.2.7 In vitro Processing and Importing of BFIT2 by Freshly Isolated Mitochondria from Untransfected HEK293T/17 cells.

Precursor (full length) BFIT2 were isolated by immunoprecipitation. A posttranslational processing assay was performed as follows: 6 mg of freshly isolated

mitochondria was resuspended in 150 μ L of 2x mannitol buffer (450 mM mannitol, 50 mM sucrose, 20 mM Tris, 0.2 mM EDTA, 100 mM ATP, 10 mM $MgCl_2$, 160 mM KCl, 0.2 mM pyruvate and 360 mM malate). After addition of 0.6 μ g BFIT2 in an equal volume and incubation at room temperature for 10 mins, reaction mixture was incubated at 37°C. The 80 μ L aliquots were extracted at 5 mins, 30 mins, and 60 mins respectively, which were then heated in boiling water for 10 mins before loading to SDS page gel. Finally, after protein transfer to nitrocellulose membrane, the membrane was subject to western blot analysis.

For *in vitro* mitochondrial import assay, precursor BFIT2 were incubated with mitochondria as described above. Aliquots of reaction mixtures were extracted at 5 min, 30 mins and 60 mins. Supernatant and pellet (mitochondria) were separated and subject to western blot analysis respectively.

5.2.8 In silico Analysis of N terminal Region of BFIT2.

An emboss pepwheel helical wheel plot was utilized to investigate the amphipathy of the N-terminal region (No. 1-42 residues) of BFIT2. Sequence alignment was generated by ClustalW program. Tertiary structure prediction of the N-terminal region of BFIT2 was generated by employing Protein homology/analogy recognition engine (PHYRE).

5.2.9 MALDI-TOF Mass Spectrometry Analysis of BFIT2

A sample of precursor BFIT2 for mass spectrometry was obtained by immunoprecipitation. Sample was loaded and separated by 12% polyacrylamide SDS-gel.

Protein sample recovered from SDS-gel was subjected to trypsin digestion and subsequent MALDI-TOF mass spectrometry analysis. Signature peptides were searched against those in protein data bank by MASCOT server for protein identity confirmation (Protein Analysis Facilities, Skirball institute at the New York University School of Medicine)

5.2.10 HPLC Analysis of Thioesterase Activities of BFIT2 towards Acyl-CoAs.

Palmitoyl-, myristoyl, lauroyl- and butyryl-CoAs were tested as substrates of BFIT2. The 200 μ L reaction solution containing 50 mM HEPES (pH 7.5), 50 mM NaCl, 50 μ M one of the above-mentioned acyl-CoAs and 0.34 μ M BFIT2 was incubated at 37°C. Aliquots of 50 μ L were extracted at 30 min and 60 min respectively, and passed through a Amicon YM-10 (Millipore) filter. The control sample omitted BFIT2 was incubated at 37°C for 2h. The filtrates were analyzed by reverse phase HPLC (UV detection wavelength at 260 nm) using a C18 reverse phase column (5 μ m, 4.6 mm \times 25 cm) and a linear gradient of 0 to 65% acetonitrile in 20 mM (NH₄)₂HPO₄ buffer (pH 6.7).

5.2.11 Digitonin Treatment of Mitochondrial Fraction Freshly Isolated from BFIT2 Stably Transfected HEK293T/17 Cell.

Mitochondrial fractions of BFIT2 stably transfected HEK293T/17 cell was isolated as described previously. Mitochondrial fractions, resuspended in mannitol buffer with a concentration of 1 mg/mL, were treated with digitonin in the range of concentration from 0 to 0.6 mg/mL at 4°C for 45 mins. After spinning down at 20,000 \times g

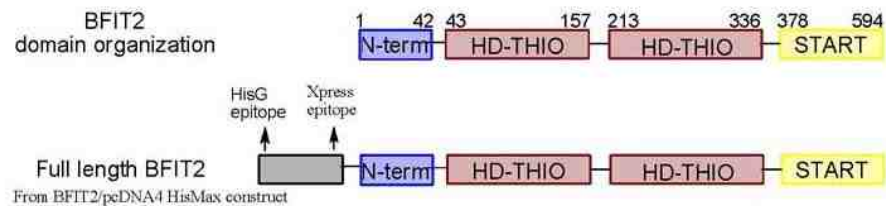
by centrifuge, pellet (mitochondria) fraction was subjected to western blot analysis using anti-mHSP70, anti-AIF and anti-Xpress (for BFIT2 detection) antibodies respectively.

5.3 RESULTS

5.3.1 Immunoblot of Recombinant BFIT2.

The expression of full length BFIT2 in *BFIT2/pcDNA4 HisMax* transfected HEK cells yielded recombinant BFIT2 (Figure 5.1 A) with 35 additional amino acids at its N-terminal region. Introduced by expression vector, this additional sequence with MW of *ca.* 3.7 kDa includes 2 epitopes, HisG (HHHHHHHG) and Xpress (DLYDDDDDK), which facilitate protein purification and immunoblot detection. In the immunoblots of transfected cell lysates, all primary antibodies detected specifically a *ca.* 71 kDa protein, which matches the predicted value of recombinant full length BFIT2 (Herein referred as precursor BFIT2) (Figure 5.1 B). Interestingly, immunoblot with anti-BFIT2 antibody showed the presence of an additional protein with *ca.* 66 kDa (Herein referred as mature BFIT2), which was evaded from detection by either anti-Xpress or anti-HisG antibody. The two forms of BFIT2 could be a consequence of an unknown posttranslational modification.

A



B

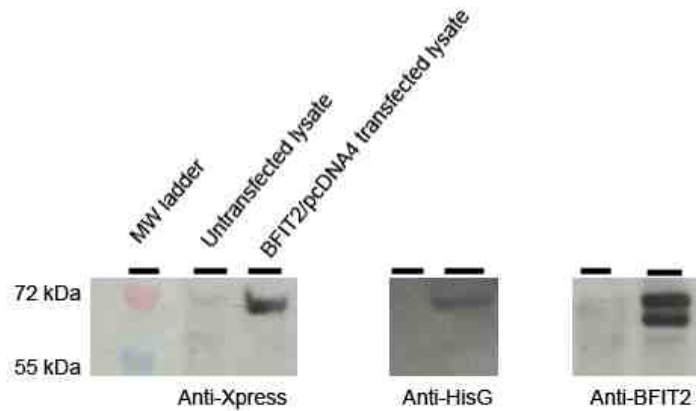


Figure 5.1 Immunoblot of untransfected and *BFIT2/pcDNA4 transfected* HEK cell lysates. Various BFIT2 construct utilized are illustrated in *panel A*. BFIT2 includes two hotdong thioesterase domains (HD-THIO), a steriodogenic acute regulatory protein related lipid transfer domain (START) and a N-terminal region (N-term). Full length (precursor) BFIT2 and BFIT2 lacking the 1-34 amino acids at its N-terminal region (Δ N33 BFIT2) were introduced 35 amino acids to their N-terminus by expression vector, which includes Xpress and HisG epitopes. Shown in *panel B* is Immunoblot of both untransfected and *BFIT2/pcDNA4 transfected* HEK cell lysates using anti-Xpress, anti-HisG and anti-BFIT2 antibody respectively.

5.3.2 Isolation of BFIT2 by Immunoprecipitation and Subsequent MALDI-TOF

Mass Spectrometry Analysis.

For the purpose of protein identity confirmation, isolation of both forms of BFIT2 was attempted by employing immunoprecipitation. To this end, BFIT2 stably transfected HEK293T/17 cells were isolated through antibiotic selection. Similar to the case of the transient transfection, western blot of the whole cell lysate of stable cells by anti-BFIT2

antibody showed 1:1 ratio of precursor to mature BFIT2 (data not shown). Immunoprecipitation was carried out using anti-Xpress antibody immobilized agarose beads. Immunoblot of the resulting immunoaffinity purified protein using anti-Xpress and anti-BFIT2 antibody respectively confirmed that only precursor BFIT2 was pulled down (Figure 5.2). Identity of the isolated precursor BFIT2 was confirmed subsequently by MALDI-TOF mass spectrometry analysis. The proteolytic digestion treatment of the protein sample yields 25 fingerprint peptides, covering 33% total sequence of BFIT2, which includes its intact N-terminal region following the Xpress epitope, parts of internal sequence and C-terminal (data not shown). Taken together, these results confirm the identity and sequence integrity of the purified precursor BFIT2.

Isolation of mature BFIT2 by immunoprecipitation using anti-BFIT2 antibody coupled beads, was not successful. This is most likely that anti-BFIT2 antibody interacts with internal sequence of BFIT2, which is folded and inaccessible for antibody interaction under the immunoprecipitation experimental conditions.

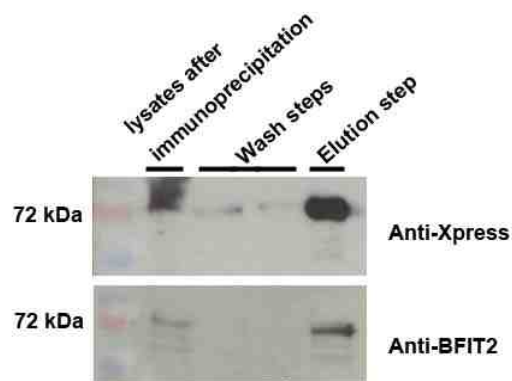


Figure 5.2 Immunoblot monitoring of isolation of precursor BFIT2 by immunoprecipitation. Immunoprecipitation was carried out using anti-Xpress antibody immobilized agarose beads (see experimental for detail). Eluents obtained from each step were loaded to SDS-page and then nitrocellulose membrane, which subsequently subject to immunoblot using anti-Xpress antibody and anti-BFIT2 antibody for antigen detection respectively. Immunoblots confirmed that only the precursor BFIT2 was pulled down by anti-Xpress coupled beads.

5.3.3 Posttranslational Modification of Precursor BFIT2 Occurs at Its N-terminal Region.

As suggested above, it is possible that precursor BFIT2 undergoes posttranslational modification to form mature protein. Studies were designed to locate where the cleavage occurs. Immunoblot analysis showed that mature BFIT2 evaded detection by either anti-Xpress or anti-HisG antibody, whose interacting antigen sequence was introduced to the N-terminus of recombinant BFIT2 by its expression vector. This result strongly indicates the cleavage occurs at the N-terminal region of BFIT2. Consequently, mature BFIT2 lacks the two artificially introduced epitopes, which make it undetectable by either anti-Xpress or anti-BFIT2 antibody.

Building on this hypothesis, $\Delta N 34$ BFIT2/*pcDNA4 HisMax* was constructed and utilized in transfection of HEK cells, yielding BFIT2 lacking 1-34 amino acid of its N-terminal region (ΔN -34 BFIT2) (Figure 5.3). Immunoblot of the transfected cell lysate using either anti-Xpress or anti-BFIT2 antibody led to the detection of only one protein with a MW of *ca.* 68 kDa, a value close to that predicted for ΔN -34 BFIT2. This

observation indicates that posttranslational modification process has been abolished and that modification involves cleavage in the N-terminal region of BFIT2 (Figure 5.3).

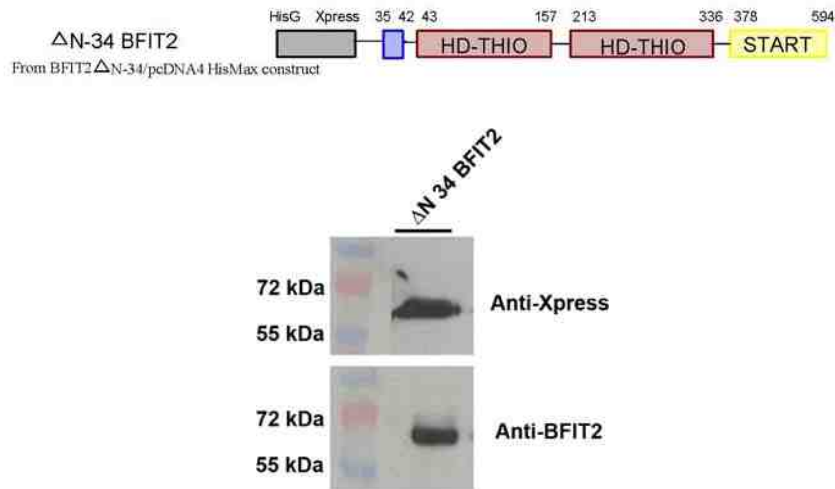


Figure 5.3 Immunoblot of $\Delta N 34$ BFIT2/*pcDNA4 HisMax* transfected HEK cell lysates. Transfection yielded BFIT2 lacking 1-34 amino acid of its N-terminal region (ΔN -34 BFIT2). Immunoblot of transfected cell lysate only detect one protein corresponding to ΔN -34 BFIT2 using either anti-Xpress or anti-BFIT2 antibody, indicating the abolishment of posttranslational modification after removal of part of the N-terminal region of BFIT2.

5.3.4 In silico Analysis of N terminal Region of BFIT2 reveals the presence of amphipathic helix.

In order to investigate the functional significance of posttranslational cleavage at the N-terminal region, a modeled tertiary structure of this region was examined. The N-terminal region of BFIT2 contains the first 1-42 amino acids. Based on the model, the overall structure is best described as turn-coil-helix-coil-turn (Figure 5.4 A). The 1-13

and 24-42 amino acid regions are shown as turn/coil, while residues 14-24 were predicted to be helical. The extensive coil structure from residue 25-42 indicates the region may function independently to the remaining domains of BFIT2. Thus, one possible purpose of the posttranslational modification at this region is for protein localization.

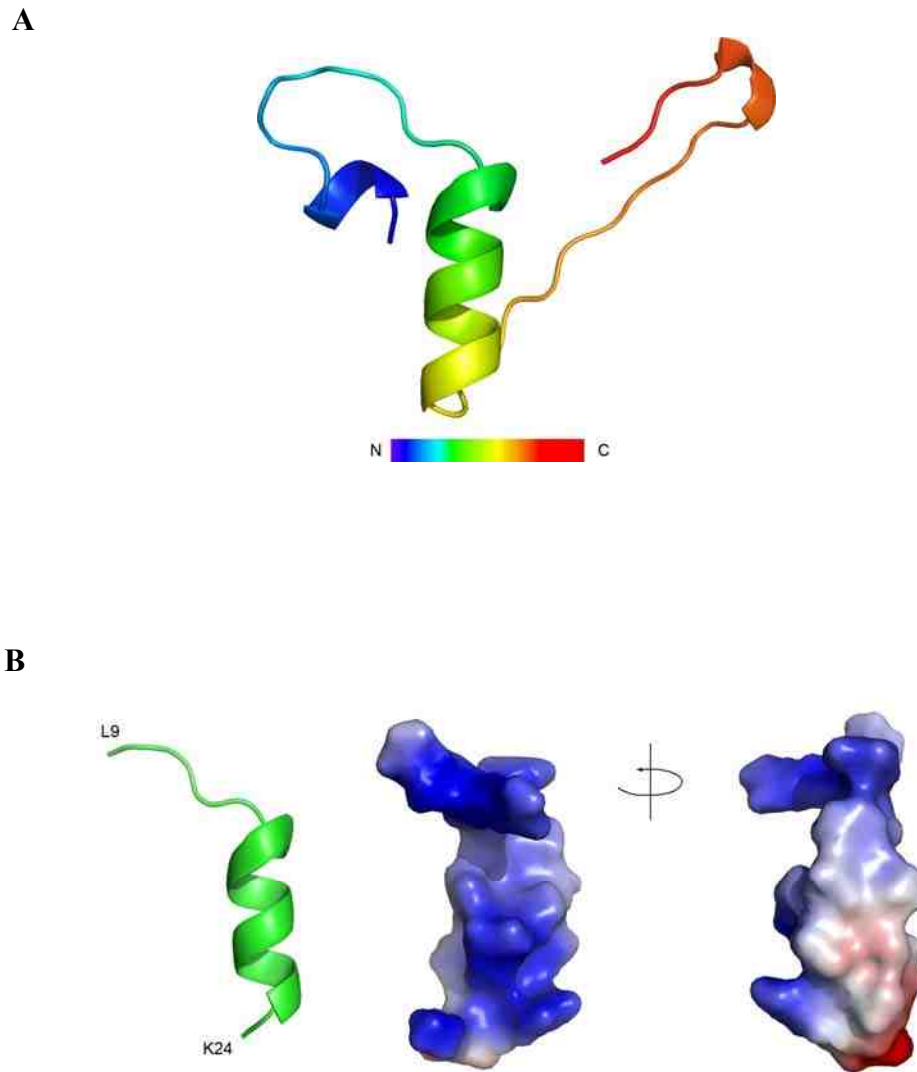
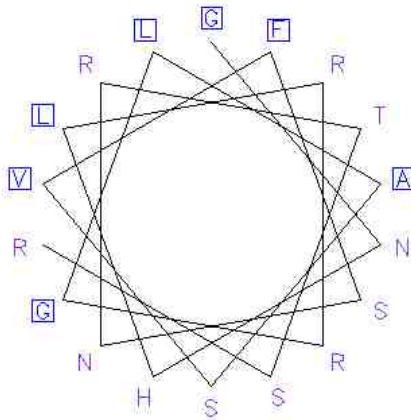


Figure 5.4 Modeled tertiary structure of N-terminal region of BFIT2 and the electrostatic vacuum surface of the central helix. A. predicted tertiary structure of BFIT2 N-terminal region; B. electrostatic vacuum surface of entral helix together with its adjacent region

(residues 9-24). One face of the surface is exclusively positively charged, while the opposite face is mainly uncharged/hydrophobic.

A



B

	pI	
Human	11.34	MIQN V GNHLRRGLASVFSNRTSRKSALR--AGNDSAMADGEGYR
Mouse	9.96	MIQN V GNHLRRGFASMF S NRTSRKSI S HPE S GD P PTMA E EGEGYR
Bovine	10.74	MIQ T VGNHLRRGLASVFSNRTSRKSASR--TDSD-NMADGEGYR
		*** * * ** *

Figure 5.5 Amphipathy of N-terminal region of BFIT2 and sequence alignment of the region from various species. A. Emboss pepwheel hexical wheel plot of residue 5-24 of the N-terminal region of BFIT2; B. Sequence alignment and estimated pI value of N-terminal region of BFIT2 from various species. Identical residues are shade in black, residues common to at least two of the three species in gray. Conserved residues with positively charged side chain are marked by asterisk.

The central α -helix, together with its adjacent coil (residues 9-24) has amphipathic feature. In the vacuum electrostatic surface of the region (Figure 5.4 B), one face of the surface is exclusively hydrophilic/positively charged, whereas the opposite

face is mainly nonpolar/hydrophobic. A close examination of the amino acid sequence further reveals the amphipathy. As shown in the emboss pepwheel plot representation of the central helix (Figure 5.5 A), residues with polar/non-polar side chain localize at opposite faces of the helix, a feature that is in good consistence with the modeled tertiary structure. Sequence alignment of N-terminal regions of BFIT2 from various mammalian species indicates that a high degree of conservation exists across species. Residues Leu9, Arg10, Arg11, Arg20, Arg23 and Lys24 with positive charge are identical in this group of protein and they impart a net positive charge in this region, as reflected in the predicted high pI values. Furthermore, this region is also enriched with well conserved hydroxyl group carrying residues, including Ser15, Ser18, Thr21, Ser22 and Ser25. Taken together, the amphipathic, helical together with preponderance of positive charge and hydroxyl carrying residues of the N-terminal region are characteristic of mitochondrial targeting sequence. Thus, it is expected that this protein carrying the sequence is imported to mitochondria promoted by mitochondrial translocases and subsequently processed by matrix processing peptidase (17-19).

5.3.5 Cellular Localization of BFIT2 N-terminal region-GFP Expression Constructs.

GFP-BFIT2 and BFIT2-GFP expression constructs were initially utilized to transfect HEK cells for the purpose of determining cellular location of BFIT2. GFP-BFIT2 showed cytosolic location (data not shown), which can be explained by the blocking of mitochondrial targeting sequence by the N-terminal GFP (20-21). On the other hand, BFIT2-GFP misbehaved and formed irregular punctuate-like inclusion body, a phenomona that is possibly a result of the attachment of GFP to the $\alpha 4$ helix of the C-

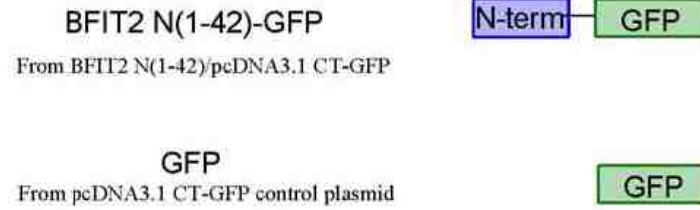
terminal end of the START domain of BFIT2, whose mobility is vital for the function of the domain (11). To avoid road blocks in labeling precursor BFIT2 by GFP, *BFIT2 N(1-42)/pcDNA 3.1 CT-GFP* construct was transfected into HEK 293T/17 cell using either Mitotracker or CMXRosWGA/Alexa Fluor 594 conjugate together with Hoechst 33342 stains to determine cellular localization of chimeric protein BFIT2 N(1-42)-GFP (Fig. 6B). GFP alone displayed distributive localization in cell (Figure 5.6 B). In contrast, BFIT2 N (1-42)-GFP exhibited patterned punctuate localization (Figure 5.6 B), which overlaps with a mitochondrial specific stain Mitotracker Red CMXRos, (Figure 5.6 C).

5.3.6 Mitochondrial Processing and Import of Precursor BFIT2.

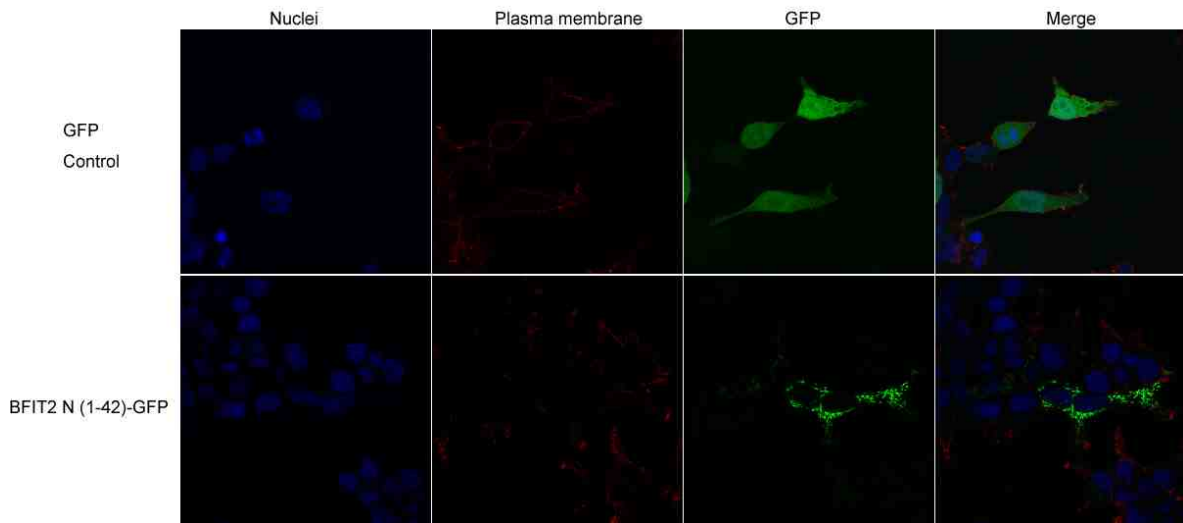
To confirm the posttranslational cleavage at the N-terminal region of BFIT2 is a consequence of proteolytic process following its translocation to mitochondria, affinity purified precursor BFIT2 was treated with freshly isolation mitochondria from HEK 293T/17 cells. Isolated mitochondria fraction is free of cytosolic fraction (Figure 5.7 A). Initially free of mature BFIT2, the precursor BFIT2 sample was processed and mature BFIT2 formed after incubation with mitochondria at 37°C (Figure 5.7B). The importing of BFIT2 to mitochondria is confirmed by western blot analysis of the mitochondria fraction after incubation of precursor BFIT2 and mitochondria isolated from untransfected HEK cells (Figure 5.7 C). Furthermore, nearly all the precursor BFIT2 was imported to mitochondria, as suggested by both the absence of BFIT2 in the supernatant and clear detection of BFIT2 in the mitochondria pellet in the blot. Only part of the precursor BFIT2 imported to mitochondria was processed to mature form. This is likely

due to the additional sequence to the N-terminal of precursor BFIT2, which might retard the cleavage of leader sequence by mitochondrial processing peptidase.

A



B



C

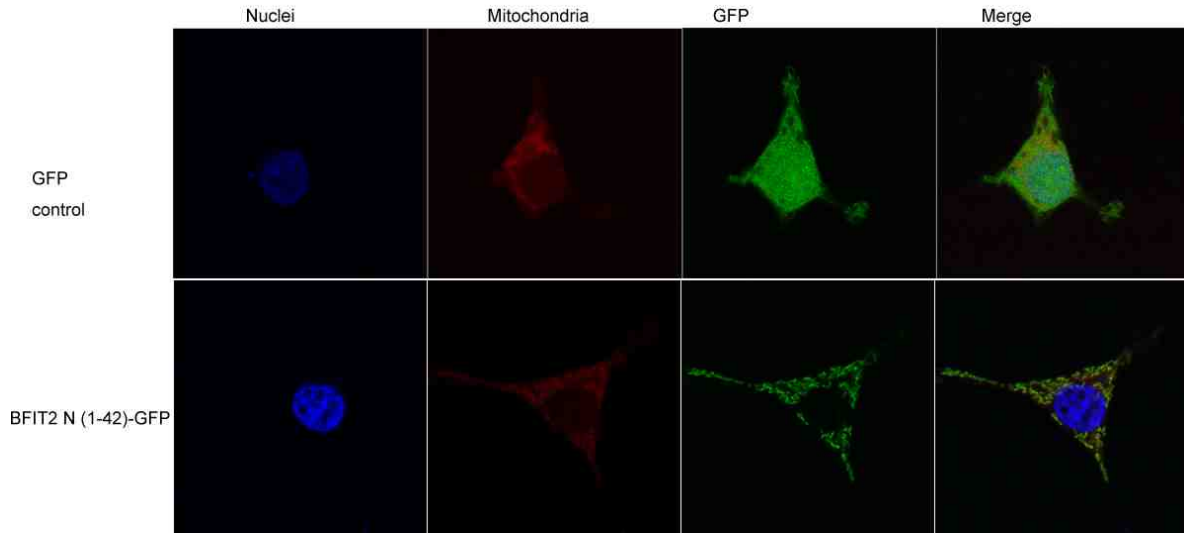
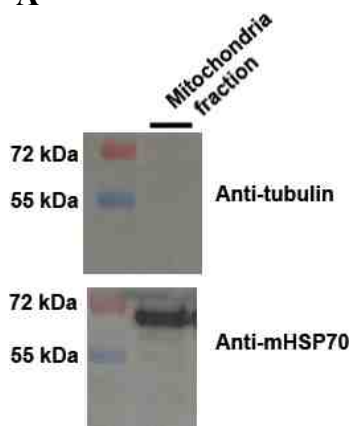
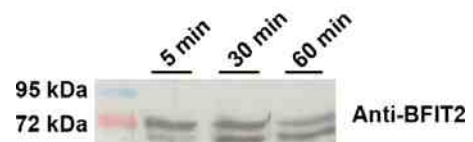


Figure 5.6 Confocal imaging of BFIT2 N-terminal region-GFP reveals its mitochondrial localization. A. domain organization representation of BFIT2 N(1-42)-GFP and GFP; B. BFIT2 N(1-42)-GFP showed punctuate localization in HEK 293T/17 cells, whose nuclei were stained by Hoechst 33342 and plasma membrane by WGA/Alexa Fluor 594 conjugate; C. The localization of BFIT2 N(1-42)-GFP overlapped with mitotracker CMXRos, giving a merged image with yellow color.

A



B



C

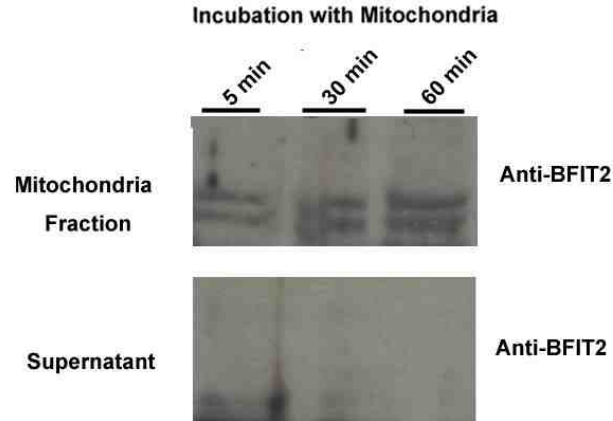


Figure 5.7 In vitro mitochondrial processing and importing of precursor BFIT2. **A.** isolated mitochondria fraction is free of cytosolic component and enriched with mitochondria. mHSP70 is a mitochondria marker (71 kDa) and α -tubulin, the cytosolic marker (50 kDa); **B.** Affinity purified precursor BFIT2 was incubated with mitochondria in mannitol buffer. Aliquots of reaction mixture were extracted at 5 min, 30 min and 60 min respectively and subject to immunoblot analysis. Precursor BFIT2, free of its mature form, was posttranslationally processed after incubation with mitochondria; **C.** Affinity purified precursor BFIT2 was incubated with mitochondria in mannitol buffer. Aliquots were extracted as described above. Supernatant and pellet with mitochondria were separated and subject to immunoblot analysis. BFIT2 were relocated to mitochondria and processed into its mature form.

5.3.7 BFIT2 Localizes in the Matrix of Mitochondria.

To identify submitochondrial location of BFIT2, digitonin treatment of mitochondria isolated from BFIT2 stably transfected HEK293T/17 was carried out. Low

concentration (0.05-0.20 mg/mL) of digitonin would permeate outer membrane of mitochondria and release protein localized in intermembrane space, while higher concentration (0.04-0.60 mg/mL) of digitonin is needed to release matrix proteins, since they are protected by inner membrane of mitochondria. AIF and mHSP70 are protein markers for intermembrane and matrix respectively.

Low concentration of digitonin brings about the release of AIF, while mHSP70 is insensitive to this concentration range (Figure 5.8), but begins to release when concentrations of digitonin are in the range of 0.40 to 0.60 mg/mL. BFIT2 responds to digitonin treatment in similar manner as does mHSP70, suggesting that it is a mitochondrial matrix protein.

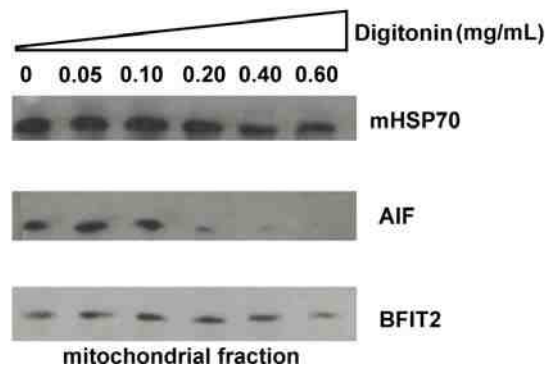
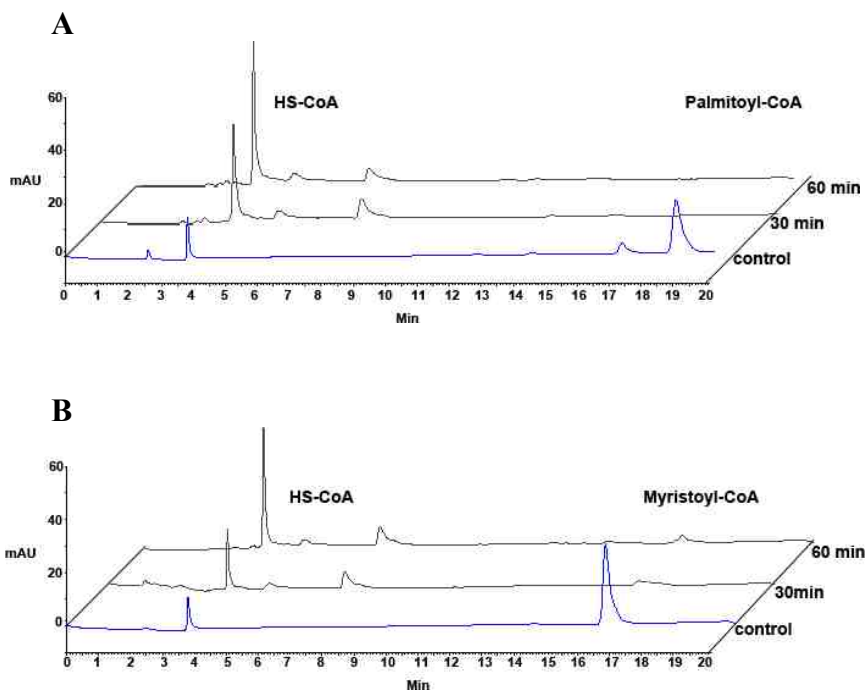


Figure 5.8 Digitonin treatment of mitochondria isolated from BFIT2 stably transfected HEK293T/17 cell. Mitochondria fractions after treated with various concentrations of digitonin at 4°C for 45 mins were subject to western blot by anti-mHSP70, anti-AIF and anti-Xpress (for BFIT2 detection) antibodies. mHSP70 is mitochondrial matrix protein, while AIF is intermembrane space marker. Similar to mHSP70, BFIT2 is insensitive to

digitonin treatment and released at high digitonin concentration at 0.6 mg/mL, suggesting its matrix submitochondrial location.

5.3.8 BFIT2 Has Thioesterase Activities towards Medium to Long Chain Acyl-CoAs.

BFIT2 clearly shows hydrolytic activities to all four of acyl-CoAs tested (Figure 5.9 A to C). In the presence of BFIT2, palmitoyl-, myristoyl- and lauroyl-CoAs were hydrolyzed to free CoA within 30 mins, as shown in HPLC spectra. In contrast, controls that do not have BFIT2 present, only showed minor spontaneous decomposition under the experimental conditions. In particular, BFIT2 is much less reactive towards the short chain butyryl-CoA, with >60% of this substrate remaining unreacted after 1 hour incubation with BFIT2, compared to the other three medium to long chain fatty acyl-CoAs (Figure 5.9 D).



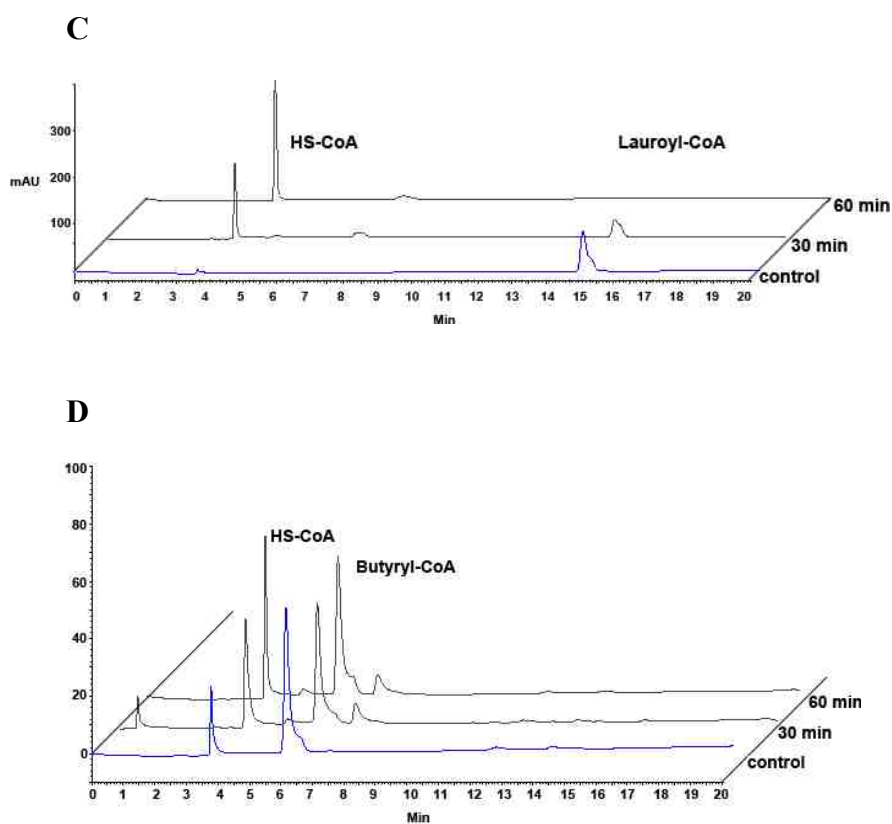


Figure 5.9 HPLC analysis of BFIT2 thioesterase activities to acyl-CoAs. Palmitoyl-CoA (A), myristoyl-CoA (B), lauroyl-CoA (C) and butyryl-CoA (D) were tested as substrates for affinity purified BFIT2. 50 μ M of one of the acyl-CoA in 50 mM HEPES buffer (pH 7.5) was incubated with 0.34 μ M BFIT2 (see experimental for details). Reactions were monitored by HPLC (UV detection wavelength at 260 nm) at 30 min and 60 min respectively at 37°C. Palmityl-, myristoyl-, lauroyl- and butyr- CoAs have retention time of 19.3 min, 16.9 min, 15.1 min and 6.4 min respectively. As control experiment, 50 μ M of one of the acyl-CoA was incubated in 50 mM HEPES in the absence of BFIT2 for 2 hours at 37°C, before subject to HPLC analysis.

5.4 Discussion

Activation of adaptive thermogenesis is concomitant with patterned changes in gene expressions. Upregulated expression takes place for those encoding proteins involving in glucose uptake, *de novo* lipogenesis, and long chain fatty acid oxidation etc. This feature suggests that a systematic regulation and coordination of a variety of metabolic processes occur. Thus, effective thermogenesis not only requires UCP1, but it is also associated with recruitment of a variety of other proteins, many of which including BFIT have unknown functions.

In addition to the upregulated RNA level of BFIT during thermogenetic transition of BAT, other information about BFIT can be garnered from proteomic study of mouse mitochondria, the result of which showed that mouse BFIT (mBFIT) is not detectable in mitochondria, but its function is coordinated with mitochondrial process (22). Recent mitochondrial proteomic analysis of mouse BAT upon cold exposure, however, identified the presence of mBFIT in mitochondria. Strikingly, these proteomic analyses also showed that some other Hotdog fold thioesterases, including Acot7, Acot8, and hThem2 (Acot13), are present in BAT mitochondria (23), a finding that conflicts with their assignment to cytosolic/perixosomal cellular locations made in biochemical studies and the fact that none of these proteins possess mitochondrial targeting sequence (MTS) (13,24-25). Furthermore, BFIT was also proposed to be a cytoplasmic protein (26). However, whether BFIT possesses MTS and localizes in mitochondria or not remains to be addressed. Therefore, a rigorous investigation of these major questions is an essential component of characterization and ultimate functional assignment of BFIT.

In the present study, BFIT2 was selected for this effort owing to the fact that the only conserved isoform for both human and mouse BFIT is its variant 2. The N-terminal

region of BFIT2 with no annotated function was examined. This region, composed of the first 1-42 amino acids, is well conserved among analogous protein in various mammalian species, although it does not have domain or sequence homolog. Posttranslational modification at N-terminal region was identified. Sequence and structural prediction showed that this region contains a central helix with features common to MTS. Consistent with this finding is the cellular localization of chimeric protein, N (1-42)-GFP, which is confirmed to be in mitochondria and suggestive of a region containing MTS. The result of in vitro posttranslational modification assay showed that the processing of precursor BFIT2 is mitochondrial dependent. Furthermore, observation made by using mitochondrial importing assay showed that BFIT2 completely relocates from supernatant to mitochondria of untransfected HEK cells in mitochondria, again providing support for the proposal that it is localized in mitochondria for its possession of MTS.

With evidences demonstrating the identity of BFIT2 as mitochondrial resident, we next set out to find out the sub-mitochondrial location of this soluble protein. Based on the differences in sub-mitochondrial locations and localization mechanisms, mitochondrial proteins are generally categorized into 4 classes (19). Only two types of mitochondrial localized proteins, class I intermembrane space (IMS) proteins and mitochondrial matrix proteins, contain typical MTS for transportation and processing. Unlike mitochondrial matrix protein, Class I IMS proteins in general have not only the typical positively charged amphipathic MTS, but they also have a hydrophobic sorting domain following MTS. This sequence facilitates the insertion of protein in lipid bilayer, where it anchors to the inner membrane within intermembrane space after removal of MTS.

The identification of the sub-mitochondrial location would provide further insights into protein functions. Thioesterase superfamily member 4 (also known as Them4 and CTMP), for example, resides at both mitochondrial inner membrane and IMS as do the other apoptic proteins. In deed, Them4 is released into cytosol on early apoptosis where it inhibits Akt phosphorylation and sensitizes cell to apoptosis (27-28).

The sub-mito location of BFIT2 was probed by digitonin treatment of freshly isolated BFIT2 stably transfected HEK cells to avoid the formidable task of isolation of mitochondria from mouse BAT exposed to cold or fat diet. Recombinant BFIT2 was found to release from mitochondria in a fashion that is similar to matrix protein mHSP70. Moreover, amino acid sequence of BFIT2 lacks the hydrophobic sorting domain of type 1 IMS protein. Together, these observations suggest that BFIT2 localizes in the BAT mitochondrial matrix.

Fatty acid oxidation, carried out in mitochondrial matrix in BAT, is coupled with heat production during thermogenesis. It has been proposed that BAT mitochondria contain large quantities of medium to long chain fatty acyl-CoAs owing to their high availability and the upleveled mitochondrial transportation of fatty acids. The matrix location of BFIT2 leads us to believe BFIT is involved in the metabolism of these CoA species. To test this hypothesis, palmitoyl- (C16), myristoyl-(C14) and lauroyl-(C12) CoAs were used as potential substrates for recombinant precursor BFIT, expressed and purified from HEK cells. The results show that BFIT2 has good catalytic activity towards these long to medium acyl-CoAs, but not for shorter chain, such as butyryl-CoA.

Observations in several investigations suggest that *Homo sapiens* Hotdog fold thioesterases are involved in acetoacetate formation, fatty acid synthesis termination,

lipid and membrane biosynthesis, conversion of free CoA from acetyl-CoA and disease developments (29-32). However, partially complicated by their tissue and event specific expression, and interacting partners, molecular details needed to assign their cellular role(s) remain elusive. In the case of BFIT2, the results of this effort which have lead to the assignment of its mitochondrial matrix location and substrate range, together with previous report of its inducible behavior, enables a more precise definition of the functional role played by BFIT in BAT thermogenesis transition.

Upon activation of thermogenic transition in response to low temperature and fat-rich diet, BAT releases neurotransmitters, which induce upregulation of lipolysis. In the cytosol, the fatty acid pool increases accordingly, providing fuel for energy production in mitochondria (33). For recruitment of fatty acid in mitochondrial, expressions of genes of acyl-CoA synthetases are also activated for converting fatty acids to fatty acyl-CoA in cytosol during the thermogenic shift. The activated long chain fatty acids are then transported as carnitine esters to mitochondrial matrix, where they are converted back to acyl-CoAs by consuming free CoA pool in mitochondrial (34) and cause an increase of the ratio of acyl- to free CoA in mitochondrial matrix. It has long been known that the control of free CoA concentration in mitochondrial is critical for its proper function (31). From this perspective, thioesterase activity of BFIT towards medium to long chain fatty acyl-CoAs, the predominant acyl-CoA species in the location, would help maintaining ratio of acyl- to free CoA in mitochondria. Thus, its thermogenesis induced expression can be considered as a response to additional need in regulating CoA thioester concentration in order to maintain the proper function of mitochondria under the highly challenged metabolic situation as thermogenesis. Moreover, free CoA released by the

action of BFIT would then participate critical metabolic processes in BAT mitochondria, including tricarboxylic acid cycle (TCA), formation of acetyl-CoA from pyruvate and conversion of carnitine ester back to acyl-CoA (35).

Previous studies showed certain thioesterases may provide a mechanism in controlling free CoA level in mitochondria by aiming at a variety of CoA species (31). In the case of BFIT, however, in vitro hydrolysis assay reveals that BFIT2 targets at long to medium chain fatty acyl-CoAs. These substrates, through β -oxidation, serve as the energy source for heat production of BAT during thermogenesis. An unregulated rapid hydrolytic thioesterase would be detrimental to this functional end. Thus, a question arises regarding the regulation of BFIT thioesterase activities. One possible scenario is that regulation takes place through feed back inhibition of HS-CoA. “Half-of-sites”, a novel induced-fit mode, was reported as a mechanism to negatively regulate thioesterase activities for Hotdog fold thioesterase *Thermus thermophilus* PaaI (36). In this mode, only one of the two identical active sites formed from homodimer of PaaI is catalytic, while the other one only has regulatory function. Compared to prokaryotic counterpart, *Homo sapiens* Hotdog thioesterases, including Acot7, 8, 9, 10, 11 (BFIT) and 12, all have innovative tandem Hotdog domains. Structural studies of Acot7 and Acot12 show that they contain only one active site that is bound with free CoA in a heterodimer, suggesting the operation of “Half-of-sites” regulation. Supporting this proposal is the observation that conversion of the catalytic silent active site of Acot7 to active form is associated with increase in hydrolytic activities and relief of an inhibition effect (13). The organization of tandem hotdog domains of BFIT (both variant 1 and variant2) strongly suggests that their thioesterase activities might be subject to same mode of regulation.

In addition to its thioesterase activities, it is speculated that BFIT might involve in lipid metabolism or synthesis owing to the presence of its C-terminal START domain. Studies on some of the other 14 *Homo sapiens* START proteins from revealed that they have their lipid trafficking capabilities (37-40). The physiological lipid passengers for most of these proteins, which includes BFIT2, are largely unknown. The intriguing combination of tandem Hotdog fold with a START domain is suggestive of an interplay between the two protein folds for the purpose of mutual regulations of activities.

In summary, in this investigation, we demonstrated for the first time cellular localization of BFIT, and proposed its cellular function, which is to maintain appropriate free CoA level in BAT mitochondria. Future studies in clarifying function of BFIT will focus on its domains communication in regulation of thioesterase activities, the potential ligand(s) and the involvement of START domain in lipid transfer and metabolism.

5.5 References

1. Rothwell, N. J., and Stock, M. J. (1979) *Nature* **281**, 31-35
2. Cannon, B., and Nedergaard, J. (2004) *Physiol Rev* **84**, 277-359
3. Wijers, S. L., Saris, W. H., and van Marken Lichtenbelt, W. D. (2009) *Obes Rev* **10**, 218-226
4. Nicholls, D. G., and Locke, R. M. (1984) *Physiol Rev* **64**, 1-64
5. Arechaga, I., Ledesma, A., and Rial, E. (2001) *IUBMB Life* **52**, 165-173
6. Krauss, S., Zhang, C. Y., and Lowell, B. B. (2005) *Nat Rev Mol Cell Biol* **6**, 248-261
7. Garlid, K. D., Jaburek, M., and Jezek, P. (1998) *FEBS Lett* **438**, 10-14

8. Klingenberg, M., and Huang, S. G. (1999) *Biochim Biophys Acta* **1415**, 271-296
9. Yu, X. X., Lewin, D. A., Forrest, W., and Adams, S. H. (2002) *FASEB J* **16**, 155-168
10. Watanabe, M., Yamamoto, T., Mori, C., Okada, N., Yamazaki, N., Kajimoto, K., Kataoka, M., and Shinohara, Y. (2008) *Biol Pharm Bull* **31**, 775-784
11. Soccio, R. E., and Breslow, J. L. (2003) *J Biol Chem* **278**, 22183-22186
12. Alpy, F., and Tomasetto, C. (2005) *J Cell Sci* **118**, 2791-2801
13. Forwood, J. K., Thakur, A. S., Guncar, G., Marfori, M., Mouradov, D., Meng, W., Robinson, J., Huber, T., Kellie, S., Martin, J. L., Hume, D. A., and Kobe, B. (2007) *Proc Natl Acad Sci U S A* **104**, 10382-10387
14. Willis, M. A., Zhuang, Z., Song, F., Howard, A., Dunaway-Mariano, D., and Herzberg, O. (2008) *Biochemistry* **47**, 2797-2805
15. Zhuang, Z., Song, F., Zhao, H., Li, L., Cao, J., Eisenstein, E., Herzberg, O., and Dunaway-Mariano, D. (2008) *Biochemistry* **47**, 2789-2796
16. Adams, S. H., Chui, C., Schilbach, S. L., Yu, X. X., Goddard, A. D., Grimaldi, J. C., Lee, J., Dowd, P., Colman, S., and Lewin, D. A. (2001) *Biochem J* **360**, 135-142
17. von Heijne, G. (1986) *EMBO J* **5**, 1335-1342
18. Omura, T. (1998) *J Biochem* **123**, 1010-1016
19. Herrmann, J. M., and Hell, K. (2005) *Trends Biochem Sci* **30**, 205-211
20. Simpson, J. C., Wellenreuther, R., Poustka, A., Pepperkok, R., and Wiemann, S. (2000) *EMBO Rep* **1**, 287-292

21. Chartrain, I., Blot, J., Lerivray, H., Guyot, N., and Tassan, J. P. (2007) *Cell Biol Int* **31**, 196-201
22. Mootha, V. K., Bunkenborg, J., Olsen, J. V., Hjerrild, M., Wisniewski, J. R., Stahl, E., Bolouri, M. S., Ray, H. N., Sihag, S., Kamal, M., Patterson, N., Lander, E. S., and Mann, M. (2003) *Cell* **115**, 629-640
23. Forner, F., Kumar, C., Lubner, C. A., Fromme, T., Klingenspor, M., and Mann, M. (2009) *Cell Metab* **10**, 324-335
24. Jones, J. M., and Gould, S. J. (2000) *Biochem Biophys Res Commun* **275**, 233-240
25. Kanno, K., Wu, M. K., Agate, D. S., Fanelli, B. J., Wagle, N., Scapa, E. F., Ukomadu, C., and Cohen, D. E. (2007) *J Biol Chem* **282**, 30728-30736
26. Kirkby, B., Roman, N., Kobe, B., Kellie, S., and Forwood, J. K. (2010) *Prog Lipid Res*
27. Parcellier, A., Tintignac, L. A., Zhuravleva, E., Cron, P., Schenk, S., Bozulic, L., and Hemmings, B. A. (2009) *Cell Signal* **21**, 639-650
28. Zhao, H., Martin, B. M., Bisoffi, M., and Dunaway-Mariano, D. (2009) *Biochemistry* **48**, 5507-5509
29. Aragon, J. J., and Lowenstein, J. M. (1983) *J Biol Chem* **258**, 4725-4733
30. Svensson, L. T., Kilpelainen, S. H., Hiltunen, J. K., and Alexson, S. E. (1996) *Eur J Biochem* **239**, 526-531
31. Hunt, M. C., and Alexson, S. E. (2002) *Prog Lipid Res* **41**, 99-130
32. Yang, J. W., Czech, T., Yamada, J., Csaszar, E., Baumgartner, C., Slavic, I., and Lubec, G. (2004) *Amino Acids* **27**, 269-275
33. Jezek, P. (1999) *J Bioenerg Biomembr* **31**, 457-466

34. Sahlin, K., and Harris, R. C. (2008) *Acta Physiol (Oxf)* **194**, 283-291
35. Constantin-Teodosiu, D., Carlin, J.I., Cederblad, G., Harris, R.C. & Hultman, E. (1991) *Acta Physiol Scand* **143**, 367-372
36. Kunishima, N., Asada, Y., Sugahara, M., Ishijima, J., Nodake, Y., Miyano, M., Kuramitsu, S., and Yokoyama, S. (2005) *J Mol Biol* **352**, 212-228
37. Arakane, F., Kallen, C. B., Watari, H., Foster, J. A., Sepuri, N. B., Pain, D., Stayrook, S. E., Lewis, M., Gerton, G. L., and Strauss, J. F., 3rd. (1998) *J Biol Chem* **273**, 16339-16345
38. Ponting, C. P., and Aravind, L. (1999) *Trends Biochem Sci* **24**, 130-132
39. Tsujishita, Y., and Hurley, J. H. (2000) *Nat Struct Biol* **7**, 408-414
40. Durand, S., Angeletti, S., and Genti-Raimondi, S. (2004) *Placenta* **25**, 37-44