

Investigating the Role of Ligand Binding Residues in
Escherichia coli Phosphopantetheinyl Transferase
EntD

Vasken Parsekhian

A Thesis

in

The Department

Of

Chemistry and Biochemistry

Presented in Partial Fulfillment of the Requirements
for the Degree of Master of Science (Chemistry) at
Concordia University
Montreal, Quebec, Canada

Fall 2010

© Vasken Parsekhian, 2010

CONCORDIA UNIVERSITY
School of Graduate Studies

This is to certify that the thesis prepared

By: **Vasken Parsekhian**

Entitled: **Investigating the Role of Ligand Binding Residues in
Escherichia coli Phosphopantetheinyl Transferase EntD.**

and submitted in partial fulfillment of the requirements for the degree of

Master of Science (Chemistry)

complies with the regulations of the University and meets the accepted standards
with respect to originality and quality.

Signed by the final examining committee:

_____ Chair
Dr. J. Turnbull

_____ Examiner
Dr. J. Powlowski

_____ Examiner
Dr. M. Sacher

_____ Supervisor
Dr. P. D. Pawelek

Approved by _____
Chair of Department or Graduate Program Director

Dean of Faculty

Date _____

Abstract

Investigating the Role of Ligand Binding Residues in *Escherichia coli* Phosphopantetheinyl Transferase EntD

Vasken Parsekhian

This project focuses on understanding how bacteria employ proteins in order to survive in iron-restricted conditions. Iron is an essential nutrient for most organisms. Under conditions of iron stress, *Escherichia coli* synthesizes and secretes enterobactin, a low-molecular-weight iron-chelating siderophore. The *E. coli* enterobactin biosynthetic apparatus is composed of six enzymes, EntA-EntF. The latter part of enterobactin biosynthesis, non-ribosomal peptide synthesis (NRPS), involves the condensation of three molecules each of 2,3-dihydroxybenzoic acid and *L*-serine. NRPS involves the enzymes EntD, EntE, EntB (peptidyl carrier domain), and EntF (Lai *et al.*, 2006). The enzyme EntD is a phosphopantetheinyl transferase that facilitates NRPS by adding phosphopantetheine cofactor to the peptidyl carrier domains of EntB and EntF. The overall goal of this project is to directly characterize interactions between EntD-EntB and between EntE-EntB *in vitro*.

The major portion of this project involved characterization of recombinant *E. coli* EntD function and conformation in the presence and absence of its cofactors: magnesium and coenzyme A. Biophysical experiments indicate that EntD undergoes a conformational change upon the addition of both cofactors. Site-directed mutagenesis helped explore the importance of a conserved active site residue (E155) and a potential ligand binding residue (H44). ESI-MS and gel-based assays were used to determine EntD

(WT, variants) activity by observing the covalent attachment of phosphopantetheine to EntB. These assays prove that our recombinant hexahistidine-tagged EntD (WT, H44A) and EntB proteins are both functional *in vitro*.

This project also involved the characterization of EntB variants in which residues predicted to interact with EntE were mutated. The genes encoding the two EntB variants, W277A and R282A, were obtained by site-directed mutagenesis and used to overexpress recombinant variant proteins. Native polyacrylamide gel electrophoresis and mass spectroscopy demonstrated that these single-site variants were unstable, resulting in proteolysis of the C-terminal peptidyl carrier protein domain in which these residues are found.

Acknowledgements

I owe my deepest gratitude to my supervisor, Dr. Peter D. Pawelek, for his endless support, enthusiasm, and patience throughout my Master's studies. His guidance helped me in all the time of research and writing of this thesis. I could not have imagined a better advisor and mentor for my M.Sc. study.

I am also indebted to my committee members Dr. Justin Powlowski and Dr. Michael Sacher for their valuable time and suggestions. I would also like to sincerely thank Dr. Christopher Wilds and Dr. Anne Noronha for their HPLC attempts.

I would like to thank France Desrochers for helping me with the homology model. A special thanks to all the members of the Pawelek lab encountered over the years (Sofia Khalil, Michael Lowden, Chris Ablenas, Sam Aleksanian, Abbas Ghazi etc...). I would also like to thank my friends for their helpful advice, and also giving me a much needed break, namely; Francis McManus, Peter K. Quashie., Said Hassounah, and Natasha Hotz.

I would like to give a warm thank you to my parents, Daniel and Silva Parsekhian, for their constant source of moral and of course financial support. Their belief in me was a driving force during my graduate studies. Special thanks to my loving and funny brother, Paren, to who I have only one thing to say....yadess! I would also like to thank my sister, Lucy, and Marc for their love and support.

No words could describe the role my fiancée, Rebecca Essiminy, played during my studies. Her unconditional emotional and moral support for 8 years was my foundation throughout my academic years. As much as I am eager to finish my graduate studies, it pales in comparison to how much I am eager to marry you and build our family.

Table of Contents

LIST OF FIGURES.....	x
LIST OF TABLES	xii
ABBREVIATIONS	xiii
1. Introduction.....	1
1.1. Iron Requirement	1
1.2. Iron Bioavailability	2
1.3. Siderophore Classes	3
1.4. Structure of Enterobactin.....	6
1.5. Enterobactin Biosynthesis.....	7
1.6. Discovery of the Bifunctional Activity of EntB	12
1.7. Crystal Structure of EntB	12
1.8. Phosphopantetheinyl Transferases.....	14
1.9. Crystal Structure of Sfp	16
1.10. Sequence Alignment of Phosphopantetheinyl Transferases.....	19
1.11. EntH: Prevention and Repair Protein.....	21
1.12. Protein-Protein Interactions in the NRPS Module.....	23
1.13. Fur Protein.....	27
1.14. Enterobactin Secretion and Ferric Enterobactin Uptake	28
1.15. Mammalian Siderophores.....	30
1.16. Research Objective.....	31
2. Material and Methods.....	32
2.1. Strains, Growth Media, and Buffers.....	32
2.2. Preparation of Competent Cells.....	33
2.3. <i>E.coli</i> Transformation	33
2.4. Agarose Gel.....	34
2.5. Plasmid DNA Extraction	34
2.6. Site-Directed Mutagenesis	35
2.7. PCR Clean-up	36
2.8. Sodium dodecyl sulfate – polyacrylamide gel electrophoresis (SDS-PAGE).....	37

2.9.	Native PAGE.....	38
2.10.	Protein Expression.....	38
2.11.	Cell Lysis.....	39
2.12.	Protein Purification Under Denaturing Conditions on Ni-NTA Spin Columns	39
2.13.	Purification of H6-EntB, H6-EntD, and H6-EntE by IMAC Chromatography	40
2.14.	EntB Anion-Exchange	41
2.15.	Measurement of Protein Concentration	41
2.16.	Dialysis.....	42
2.17.	Buffer Exchange on NAP-5 Columns	42
2.18.	Synthesis of Alexa Fluor 488-CoA and BODIPY-TMR-CoA.....	43
2.19.	Thin-Layer Chromatography	43
2.20.	Biophysical Experiments	44
2.20.1.	Far-UV Circular Dichroism	44
2.20.2.	Near-UV Circular Dichroism	44
2.20.3.	Temperature Denaturation Monitored by Circular Dichroism	45
2.20.4.	Fluorescence Spectroscopy	45
2.20.5.	Fluorescence Anisotropy	46
2.20.6.	Analytical Ultracentrifugation.....	46
2.21.	Electrospray Ionization Mass Spectrometry	47
2.21.1.	Determination of Protein Mass.....	47
2.21.2.	Activity Assay.....	48
2.21.3.	Time-Course Assay	49
2.22.	Gel-based Activity Assay	49
2.23.	Gel-based Time Course Assay	50
3.	Results	51
3.1.	Characterization of EntD Cofactor-Binding Residues.....	51
3.1.1.	Homology Model	51
3.1.2.	Site-Directed Mutagenesis	53
3.1.3.	Protein Purification Under Denaturing Conditions.....	55
3.1.4.	Large-Scale Protein Purification of EntD and Variants	56
3.1.5.	Secondary Structure of EntD and Variants Assessed by Far-UV Circular Dichroism.....	59
3.1.6.	Tertiary Structure of EntD and Variants Assessed by Fluorescence Spectroscopy	63

3.1.7.	Secondary Structure Change upon Cofactor Binding Determined by Far-UV CD.....	64
3.1.8.	Thermal Stability of EntD and Variants in the Presence of Cofactors Monitored by Circular Dichroism	67
3.1.9.	Cofactor Binding Assessed by Near-UV Circular Dichroism	69
3.1.10.	Sedimentation Velocity Analytical Ultracentrifugation.....	71
3.1.11.	Bioconjugation of Alexa Fluor 488 and BODIPY-TMR with CoA	72
3.1.12.	Determination of BODIPY-TMR-CoA K_d by Fluorescence Anisotropy.....	76
3.1.13.	Determination of Phosphopantetheinyl Transferase Activity by ESI-MS.....	78
3.1.14.	Determination of EntE Activity on EntB WT by ESI-MS	80
3.1.15.	Time-Course Assay with ESI-MS.....	80
3.1.16.	Determination of Activity by SDS-PAGE Using Fluorescent CoA Conjugates	82
3.1.17.	Gel-based Time-Course Assay of EntD and Variants with EntB.....	85
3.2.	Characterization of EntB Variants	88
3.2.1.	Site-Directed Mutagenesis	88
3.2.2.	Purification of EntB WT and Variants.....	90
3.2.3.	Native Gel of EntB WT and Variants	92
3.2.4.	Molecular Weight Determination of EntB WT and Variants	93
3.2.5.	Determination of EntD Activity on Variants	95
4.	Discussion	96
4.1.	Research Objective Revisited	96
4.2.	Structural Characterization of Wild-Type EntD and Variants	97
4.3.	Overall Structural Effect of Cofactor Binding.....	100
4.4.	Measurement of EntD Enzyme Activity	104
4.5.	Structural Characterization of EntB and Variants	106
5.	Summary and Future Work	108
6.	References	111

LIST OF FIGURES

Figure 1	Representative examples of different siderophores.	4
Figure 2	Schematic (A) and 3D structure (B) of enterobactin and its ferric complex.	7
Figure 3	Schematic diagram of enterobactin biosynthesis DHB module.	9
Figure 4	Schematic diagram enterobactin biosynthesis NRPS module.	11
Figure 5	Crystal structure of EntB monomer.	14
Figure 6	Ribbon representation of Sfp in the presence of Mg ²⁺ and CoA.	18
Figure 7	Sequence alignment of Sfp with EntD and 3 other PPTases.	20
Figure 8	Partial sequence alignment of PPTases of the Sfp-type (top) and AcpS-type (bottom).	21
Figure 9	Proposed model of enterobactin transport in <i>E. coli</i>	30
Figure 10	Sequence alignment of EntD with Sfp performed using ClustalW2.	52
Figure 11	Homology-based model of <i>E. coli</i> EntD.	53
Figure 12	pCA24N plasmid with <i>entD</i> gene (b).	55
Figure 13	12% SDS-PAGE of EntD protein expression under different conditions.	56
Figure 14	Elution profile of wild-type EntD from the Imac Column.	58
Figure 15	SDS-PAGE of purification fractions of EntD wild-type.	58
Figure 16	SDS-PAGE of EntD wild-type, H44A, and E155A.	59
Figure 17	Overlay of the far-UV circular dichroism spectra of EntD, H44A, and E155A.	61
Figure 18	EntD E155A secondary structure prediction by PSIPRED.	62
Figure 19	Fluorescence emission scans of EntD WT and variants.	64
Figure 20	Far-UV circular dichroism of EntD WT in the presence of cofactors.	65
Figure 21	Far-UV CD spectra of EntD variants in the presence of Mg ²⁺ and CoA.	66

Figure 22 Thermal denaturation curves for EntD WT in the absence and presence of cofactors.	68
Figure 23 Near-UV CD spectra of EntD and variants in the presence of cofactors.	70
Figure 24 Sedimentation Velocity AUC of wild-type EntD in the presence of cofactors.....	72
Figure 25 Fluorescence emission spectra of BODIPY-TMR-CoA and 7 ethyl acetate washes.	73
Figure 26 Thin-layer chromatography for conjugation reaction of Alexa Fluor 488 with CoA. ...	75
Figure 27 Thin-layer chromatography of conjugation reaction of BODIPY-TMR with CoA.	75
Figure 28 Fluorescence anisotropy of BODIPY-TMR-CoA with EntD proteins.	77
Figure 29 Mass spectrum of <i>apo</i> -EntB.	79
Figure 30 Mass spectrum of <i>holo</i> -EntB.	79
Figure 31 Time-course assay of EntD WT, and H44A determined by ESI-MS.	82
Figure 32 10% SDS-PAGE demonstrating incorporation of fluorescent into EntB substrates.	84
Figure 33 10% SDS-PAGE of time course assay of EntD WT with BODIPY-TMR-CoA.	86
Figure 34 Time course of the reaction of <i>apo</i> -EntB with increasing time.	87
Figure 35 Ribbon Representation of EntB ArCP domain.	89
Figure 36 Amino acid sequence of EntB with secondary structure assignment.	90
Figure 37 12% SDS-PAGE of EntB WT purification after Profinity Imac Column.	91
Figure 38 12% SDS-PAGE of EntB WT, W277A, and R282A after anion exchange chromatography.	92
Figure 39 6% Native polyacrylamide gel of EntB WT and variants.	93
Figure 40 Schematic representation of cleaved residues of EntB variants.	94

LIST OF TABLES

Table 1 Recipes for growth media and buffers	32
Table 2 Secondary structure percentages determined by K2D server and PSIPRED.	61
Table 3 Melting temperature of EntD and variants determined by circular dichroism ..	68
Table 4 Sedimentation coefficient and frictional ratio of wild-type EntD in the presence of cofactors as determined by SEDFIT analysis	72
Table 5 Apparent K_d of BODIPY-TMR-CoA with EntD proteins determined by fluorescence anisotropy.	78
Table 6 Summary of end-point activity for wild-type EntD and variants	80
Table 7 Summary of EntE activity on <i>holo</i> -EntB determined by ESI-MS.	80
Table 8 Molecular weight determination of EntB WT and variants by ESI-MS.	94
Table 9 Summary of EntB (WT, variants) mass determinations after reaction with EntD as measured by ESI-MS.....	95

ABBREVIATIONS

2,3-DHB	2,3-Dihydroxybenzoate
APS	Ammonium persulfate
Ala (A)	Alanine
Arg (R)	Arginine
Å	Angstrom unit
ACN	Acetonitrile
ArCP	Aryl carrier protein
AUC	Analytical Ultracentrifuge
CoA	Coenzyme A
CD	Circular Dichroism
Da	Dalton
DNA	Deoxyribonucleic acid
EntA	2,3-dihydro-2,3-dihydroxybenzoate dehydrogenase
EntB	Isochorismate lyase/Aryl carrier protein
EntC	Isochorismate synthase
EntD	Phosphopantetheinyl transferase
EntE	2,3-dihydroxybenzoate AMP ligase
EntF	Enterobactin synthetase component F
EDTA	Ethylenediamine tetra-acetic acid
ESI-Q-TOF	Electrospray ionization quadrupole time of flight
FPLC	Fast protein liquid chromatography
Glu (E)	Glutamate

Hepes	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
His (H)	Histidine
IPTG	Isopropyl- β -D-thiogalactopyranoside
k_{cat}	Turnover number
K_d	Dissociation constant
K_m	Michaelis constant
LB	Luria broth
MW	Molecular weight
NRPS	Non-ribosomal peptide synthesis
MS	Mass spectrometry
PCP	Peptidyl carrier protein
PCR	Polymerase chain reaction
PDB	Protein data bank
PMSF	Phenylmethylsulfonyl fluoride
PPant	Phosphopantetheine
PPTase	Phosphopantetheinyl transferase
SDS-PAGE	sodium dodecyl sulfate polyacrylamide gel electrophoresis
TCEP	tris(2-carboxyethyl)phosphine
TE	Thioesterase
TEMED	Tetramethylethylenediamine
TMR	Tetramethylrhodamine
Tris	tris(hydroxymethyl)aminomethane
Trp (W)	Tryptophan
UV	Ultraviolet

1. Introduction

1.1. Iron Requirement

Iron is an essential nutrient for virtually all organisms. It is required for the growth of nearly all organisms such as bacteria, fungi, and plants. The necessity for iron arises by a vast variety of iron-dependent processes. Iron-sulfur proteins such as glutamate synthase and dihydroorotate dehydrogenase are required in amino acid and pyrimidine biosynthesis, respectively (Earhart, 2004). The tricarboxylic acid cycle is also dependent on iron-sulfur proteins such as aconitase and succinate dehydrogenase (Kurtz *et al.*, 1979, Salerno *et al.*, 1976). Another iron-sulfur protein dependent process is electron transport, where ferredoxin is required as an electron carrier and/or electron donor (Otaka *et al.*, 1989). Higher organisms have heme-containing proteins such as cytochromes which require iron for electron transport. Other heme-containing proteins are cytochrome oxidase required for the activation of oxygen, and peroxidase required for the hydrogen peroxide reduction (Earhart, 2004). Other iron-containing proteins, which are non-heme and non iron-sulfur, are involved in many biological processes. Ribonucleotide reductase, superoxide dismutase, and 3-deoxy-D-arabinoheptulosonic acid 7-phosphate synthase are required in different processes such as DNA synthesis, protection for superoxide radicals, and aromatic amino acid biosynthesis, respectively (Earhart, 2004).

1.2. Iron Bioavailability

Iron is the fourth most abundant element in the Earth's crust (Morgan *et al.*, 1980) but acquiring it has proven to be a challenge for microorganisms. This challenge is not due to its scarcity but due to the state of which iron is found in the environment. The soluble Fe(II) is only found under anaerobic conditions (Grass, 2006). Therefore, it is only accessible for cells growing anaerobically. Under aerobic conditions, free iron is mainly found in its oxidation state Fe(III). Ferric iron is complexed with oxides or hydroxides rendering it very insoluble (Earhart, 2004). The solubility product of Fe(OH)₃ has been determined to be 10⁻³⁹ (Hider *et al.*, 2009). Other conditions that determine the state of iron are the pH and the temperature. For example, at pH 7, it would take one hour to oxidize ferrous iron into ferric iron at 21 °C, while it could take up to 10 hours at 5 °C. Lowering the pH also elongates the time required to complete the oxidation reaction. At pH 6, it would take 100 hours to oxidize Fe²⁺ (Vance, 1994). Iron is also greatly limited to bacteria because in host animals, iron is bound by carrier proteins such as lactoferrin and transferrin which reduce its bioavailability. It is estimated that 1 μM of intracellular iron is required for microbial growth (Raymond *et al.*, 2003). However, these iron binding proteins greatly reduce the bioavailability of iron to 10⁻¹⁸ M (Raymond *et al.*, 2003). In order to overcome this obstacle, many organisms evolved a common strategy to scavenge and absorb iron from their environment such as soil, water, and their host organisms. This common strategy is the employment of siderophores for iron accumulation. Siderophores need to be capable of coordinating ferric iron to meet its chemical and geometrical requirements. Ferric iron has 5 valence electrons and to reach

an energetically stable thermodynamic state requires 12 donor electrons (Raymond *et al.*, 2003). An octahedral configuration is required since it provides the lowest energy needed for these electrons to coordinate ferric iron (Raymond *et al.*, 2003). The octahedral configuration can be visualized as two pyramids that have a common base and the iron atom in the center.

1.3. Siderophore Classes

Siderophores are low molecular-weight molecules that bind iron with a very high affinity. Over 500 siderophores have been identified ever since the first siderophores (mycobactin, ferrichrome, and coprogen) were isolated and identified as growth factors in the years 1949-1952 (Hider *et al.*, 2010). Siderophores differ from one another by their chemical structure, ligand denticity, and biosynthesis. The ligand denticity describes the number of coordinating atoms per molecule and in siderophores it ranges from bidentate to hexadentate. The approximately 500 different siderophores so far identified are categorized into different classes, which are distinguished by the chemical structure of the metal-binding group. As Figure 1 demonstrates, the five main classes of siderophores are hydroxamates, catecholates, hydroxycarboxylates, phenolates, and mixed types. Other metal-binding functional groups that are found in siderophores are oxazoline, thiazoline, α -keto acid, hydroxypyridinone, and α - and β -hydroxy acids. Hydroxamate is a functional group where the hydroxylamine is inserted into a carboxylic acid. Ferrioxamines and ferrichromes were the first hydroxamate siderophores that were recognized in bacteria (Raymond *et al.*, 2004). Desferrioxamine is another example

of a hydroxamate siderophore that is found in terrestrial and aquatic systems (Carrasco *et al.*, 2009). Hydroxamate siderophores are found in both bacteria and fungi. Hydroxamate siderophores are primarily hexadentate and originate from L-ornithine.

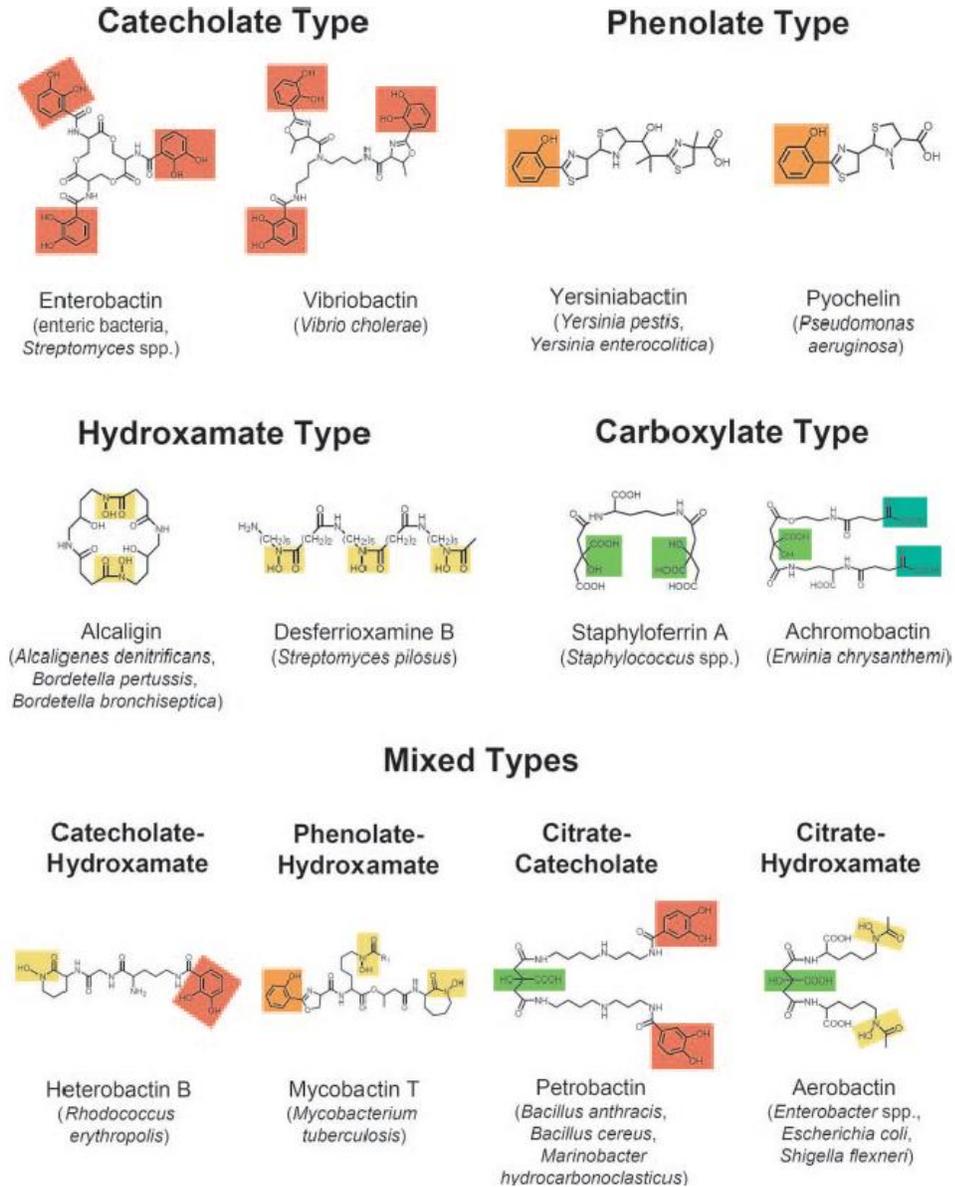


Figure 1 Representative examples of different siderophores.

The functional groups involved in iron coordination are color coded as follows: catecholates as red, phenolates as orange, hydroxamates as yellow, and carboxylates as green. Figure from Miethke *et al.*, (2007).

Catecholate siderophores are the other major class that includes enterobactin, which is the most studied of all the catecholate siderophores and the proteins involved in its synthesis are the focus of this paper. Enterobactin structure as well as its biosynthesis will be discussed in the following sections. Catecholates are produced only by bacteria. Catecholate siderophores can be cyclic as enterobactin or they can be linear such as azotobactin, protochelin, and aminochelin (Hider *et al.*, 2010). All linear catecholate siderophores are produced by *Azotobacter vinelandii*. Not unlike enterobactin, protochelin from *B. anthracis* is composed of three units of 2,3-dihydroxybenzoic acid. However, these units are attached to a linear backbone formed by the amide of lysine diaminobutane. Carboxylate siderophores are produced by a few bacteria and the fungi *Mucorales*. The bacteria that produce these siderophores are *Rhizobium meliloti* (rhizobactin, and rhizoferrin) and *Staphylococcus hyicus* (staphyloferrin A and B). The donor groups carboxylate and α -hydroxy are responsible in chelating Fe^{3+} in these siderophores. Carboxylates have also been determined to be produced by halophilic archaea (Dave *et al.*, 2006). However, the structures of these siderophores have not yet been determined. Mixed type siderophores contain a combination of two different structural functionality groups such as mycobactin T which is a phenolate-hydroxamate siderophore that has two hydroxamates and a single phenolate metal-binding group. Some siderophores have also been determined to be fluorescent. These siderophores include pyoverdines and pseudobactins which are produced by *Pseudomonas*. Siderophores in the pyoverdine family are heteropeptides composed of a linear or cyclic peptide. They also have a small dicarboxylic acid and a

quinolinic chromophore (Hamdan *et al.*, 1991). The fluorescence of these siderophores has not been related to any functional use.

1.4. Structure of Enterobactin

Enterobactin has a threefold symmetry and is composed of three 2,3-dihydroxybenzoic acid (DHB) molecules and three L-serines. The three L-serines form the trilactone macrocycle backbone (Figure 2A). At neutral pH, when metal coordination occurs through the six catecholate oxygens, the ferric-enterobactin complex has a net charge of -3 (Figure 2B). At low pH, iron chelation occurs by salicylate rather than catecholate complexing (Earhart, 2004). Under acidic conditions, the carbonyl oxygens of the amide groups replace the *m*-hydroxyl oxygens for ligation which have been protonated and not capable of metal-binding. Therefore, under acidic conditions, ferric-enterobactin is neutral. The ability of enterobactin to bind Fe³⁺ is immensely strong (approximate K_d 10⁻⁵²M) (Thulasiraman *et al.*, 1998) and is capable of removing iron from proteins, insoluble complexes, and even other siderophores (Zeng *et al.*, 2009). The linear forms of enterobactin constitutes of DHB-Ser (DBS) moiety. DBS₃, DBS₂, DBS and even DHB can serve as siderophores. However, these secondary siderophores have a lower affinity for Fe³⁺ (Earhart, 2004).

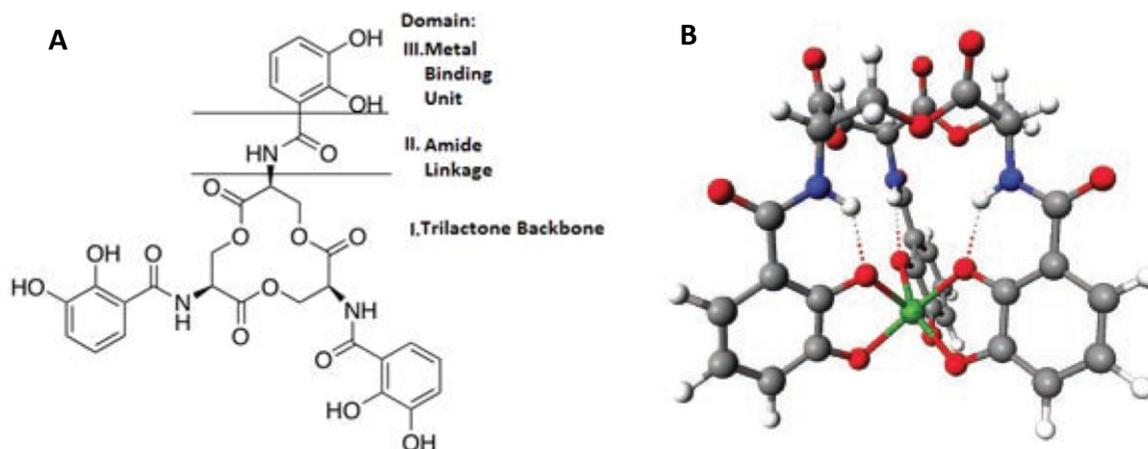


Figure 2 Schematic (A) and 3D structure (B) of enterobactin and its ferric complex.

A) Schematic diagram shows the three domains of enterobactin. B) Enterobactin folds to allow metal coordination of the six catecholate oxygens (red) with Fe^{3+} (green). Nitrogen atoms are shown as blue and hydrogen as white. Figure from Raymond *et al.* (2003).

1.5. Enterobactin Biosynthesis

Biosynthesis of enterobactin occurs in *E. coli* cytoplasm by seven enzymes; EntC, EntB, EntA, EntE, EntF, EntD, and EntH. The first five enzymes are directly involved in enterobactin biosynthesis and they are divided into the two functional modules. EntD and EntH are indirectly involved by either post-translationally modifying these enzymes or as a repair system. The first module of the enterobactin biosynthesis consists of the DHB module. This module is responsible for the conversion of chorismate into DHB and composed of EntC, the N-terminal domain of EntB, and EntA (Figure 3). Chorismate is a product of the shikimate pathway and an aromatic amino acid precursor. Under iron starvation, chorismate is directed away from the aromatic amino acid production pathway and converted into the siderophore precursor, DHB. Isochorismate synthase

(EntC), is the first enzyme in the enterobactin biosynthesis pathway and is responsible for the conversion of chorismate into isochorismate. EntC is a monomer with a molecular weight of 42.9 kDa. The crystal structure of EntC was recently solved by Sridharan *et al.* (2010) at 2.3 Å resolution. It belongs to a superfamily of enzymes that utilizes chorismate as a substrate and similar to these enzymes it adopts an $\alpha+\beta$ fold. The overall structure contains nine α -helices and 18 β -strands. EntB is a bifunctional enzyme and two-thirds of the N-terminal domain contains the isochorismate lyase activity. It hydrolyzes and releases the pyruvate group of isochorismate to form 2,3-dihydro-2,3-dihydroxybenzoate. The structure of EntB will be further discussed in section 1.7. The product of the isochorismate lyase domain is then converted into DHB by the activity of EntA, a 2,3-dihydro-2,3-dihydroxybenzoate dehydrogenase. EntA is a tetramer with a subunit size of 26 kDa and it requires NAD^+ . EntA is a member of the short alcohol-polyol-sugar dehydrogenase family.

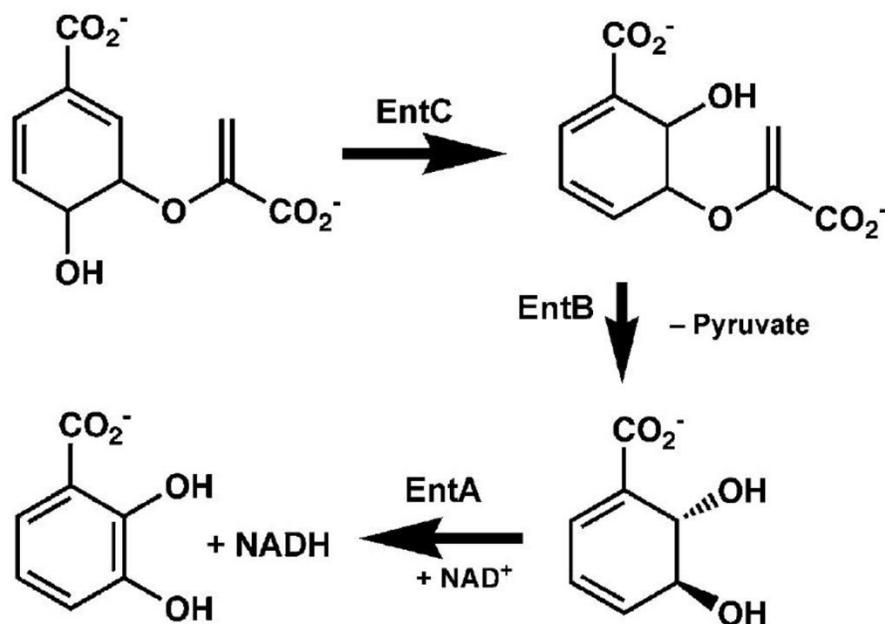


Figure 3 Schematic diagram of enterobactin biosynthesis DHB module.

Biosynthesis of 2,3-DHB from chorismate by EntC, EntB, and EntA. Under iron-limiting conditions, EntC converts chorismate into isochorismate. EntB then converts isochorismate into 2,3-dihydro-2,3-dihydroxybenzoate and releases a pyruvate molecule. In the presence of NAD^+ , EntA then converts the product of the ICL domain of EntB into 2,3-dihydroxybenzoate. Figure from Drake *et al.* (2006).

Non-ribosomal peptide synthetase (NRPS) is the second module and converts three molecules of DHB and three molecules of L-serine into enterobactin (Figure 4). The energy of this reaction is derived from hydrolysis of six ATP molecules. EntD is a phosphopantetheinyl transferase (PPTase), it post-translationally modifies EntB and EntF by transferring the 4'-phosphopantetheine (PPant) of coenzyme A (CoA) to a conserved serine residue on the ArCP domain of *apo*-EntB and *apo*-EntF converting them to *holo*-EntB and *holo*-EntF, respectively (Figure 4a). EntE is a DHB-AMP ligase, which will activate DHB and transfer it to the new PPant group on *holo*-EntB (Figure 4b). EntE is a monomer with a molecular weight of 59.2 kDa. The C-terminal one third of

EntB, which consists of approximately 100 amino acids, is referred to as the ArCP domain. Since the new PPant group tethers an aromatic carboxylic acid instead of an amino acid, the domain is termed aryl carrier protein instead of peptidyl carrier protein (PCP). EntF is at the center of the NRPS module where most of the reactions occur. It has a large size of 142 kDa and contains four domains: condensation (C), adenylation (A), PCP, and thioesterase (TE). The A domain activates L-serine and attaches it to the PPant group of the PCP domain (Figure 4b). The C domain will then transfer DHB from the ArCP domain of EntB and will catalyze amide bond formation with the serine (Figure 4c). Therefore, the C domain is the peptide bond forming catalyst. This reaction will yield a DBS moiety which is then transferred to the TE domain (Figure 4d). DBS is bound to a serine residue in an oxoester linkage. The empty C domain is then occupied by another DBS and an ester bond is formed between this DBS and the former in the TE domain (Figure 4d). Another round of DBS synthesis and ester bond formation to the DBS dimer yields a DBS trimer on the TE domain. The DBS trimer is then cyclized and released from EntF (Figure 4e).

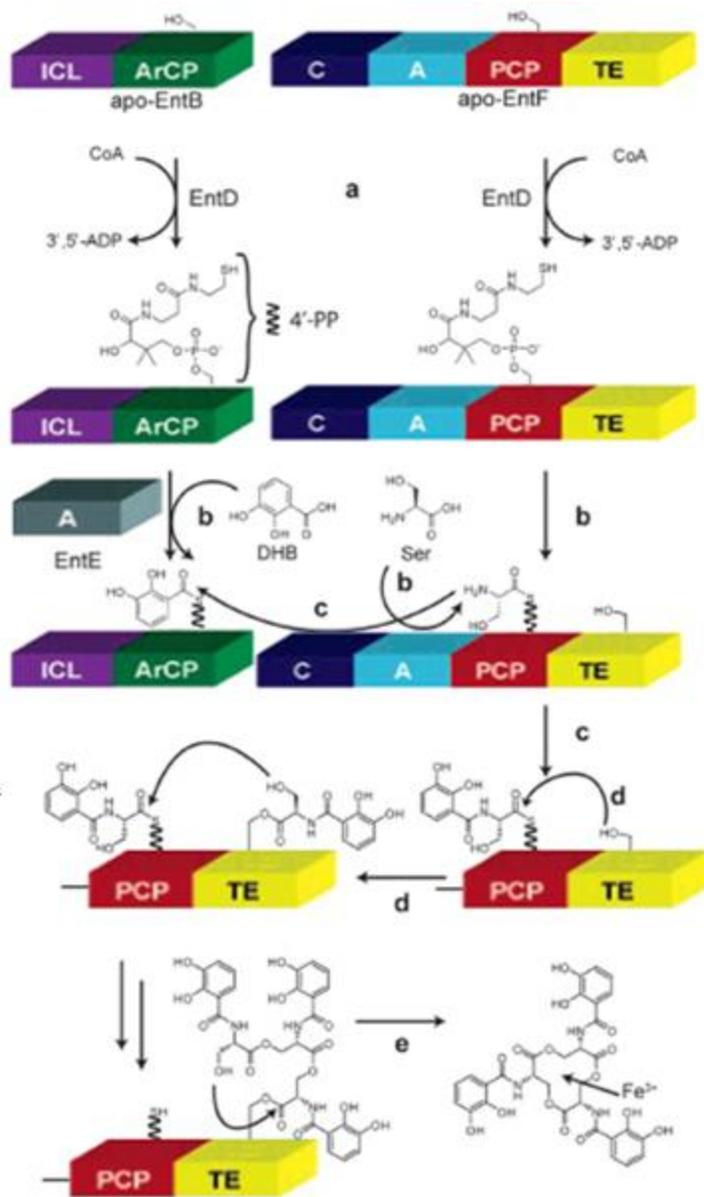


Figure 4 Schematic diagram enterobactin biosynthesis NRPS module

The first step in the NRPS module is the phosphopantetheinylation of ArCP domain of EntB and PCP domain of EntF (a). The adenylation activity of EntE and of the A domain of EntF attach DHB and L-serine onto this newly introduced PPant arm on EntB and EntF, respectively (b). Condensation of DHB with serine (DBS) occurs by means of the C domain of EntF. DBS is transferred to a serine residue on EntF TE domain (d). Steps b-d are repeated two more times to produce DBS₃ which is then cyclized and released by the EntF TE domain to produce enterobactin (e). Figure modified from Frueh *et al.* (2008) and Lai *et al.* (2006).

1.6. Discovery of the Bifunctional Activity of EntB

Initially, EntB was believed to be actually two separate proteins due to the activity of its two domains. The N-terminal ICL domain was termed EntB and the C-terminal ArCP domain was named EntG. One of the first indications that EntB was a bifunctional protein came from Staab *et al.* (1990) by using genetic, biochemical, and immunological studies. Their sequencing data failed to show a separate EntG gene. Also, analysis of EntG mutants determined that EntG corresponds to the C-terminus of EntB. Additionally, Gehring *et al.* (1997) proved that EntB is a bifunctional enzyme by employing genetic and biochemical studies. They generated three different EntB constructs to test as substrates for EntD: full length EntB, EntB Δ C where they deleted the C-terminal 25 amino acids, and EntB ArCP consisting of the C-terminal 98 amino acids. Full-length EntB and EntB ArCP had comparable catalytic efficiency to be phosphopantetheinylated by EntD with a k_{cat}/K_m of 0.8 and 0.9 $\mu\text{M}^{-1} \text{min}^{-1}$, respectively. However, EntB Δ C had a 16,000 fold decrease in catalytic efficiency resulting in a k_{cat}/K_m value of $5 \times 10^{-5} \mu\text{M}^{-1} \text{min}^{-1}$. This was due to the omission of the C-terminal 25 amino acids which contain the site for phosphopantetheinylation, serine 245. Furthermore, this showed that the two domains can function independently.

1.7. Crystal Structure of EntB

The crystal structure of EntB was solved by Drake *et al.* (2006) with a resolution of 2.3 Å. They used a single domain of the 207 residue long 2-amino-2-deoxyisochorismate lyase (PhzD) protein as the search model. PhzD is 45.5 % identical to the ICL domain of

EntB. The crystal structure of EntB shows that it is a dimer where the two ICL domains form the dimerization interface in the asymmetric unit and the ArCP domains are pointing out. Figure 5 illustrates the domains of EntB with the N-terminal ICL domain colored as magenta and the C-terminal domain colored green. The two domains are separated by the linker depicted as cyan which consists of 5 amino acids Pro-Ala-Pro-Ile-Pro. The ICL domain consists of a central 6 stranded parallel β -sheet. On one side of the sheet is a 20 residue long α -helix ranging from Pro51-Asn71. On the opposite side of the sheet are three α -helices that form the subunit interface of the EntB dimer. The C-terminal ends of the β -sheet are capped by a short single-turn helix. The α -helix asp85-Met94 in the ICL domain makes hydrophobic interactions with the neighboring subunit. Leu90 and Leu91 of one subunit interact with Trp24 and Tyr165 on the neighboring subunit of the dimer. Another set of interactions occurs between Met94 and Trp95 of one subunit and the methylene carbons of Lys21 and Arg196 of the neighboring subunit. The ArCP domain consists of four helices. The site of phosphopantetheinylation, Ser245, is found at the start of the second helix. Ser245 is found on the ArCP face that points away from the ICL domain and is directed into the solvent. This position allows easy access for the cofactor and to the EntE and EntF proteins with which it will interact. The two domains of EntB do not have extensive interdomain interactions and they can function independently from one another as described in section 1.6.

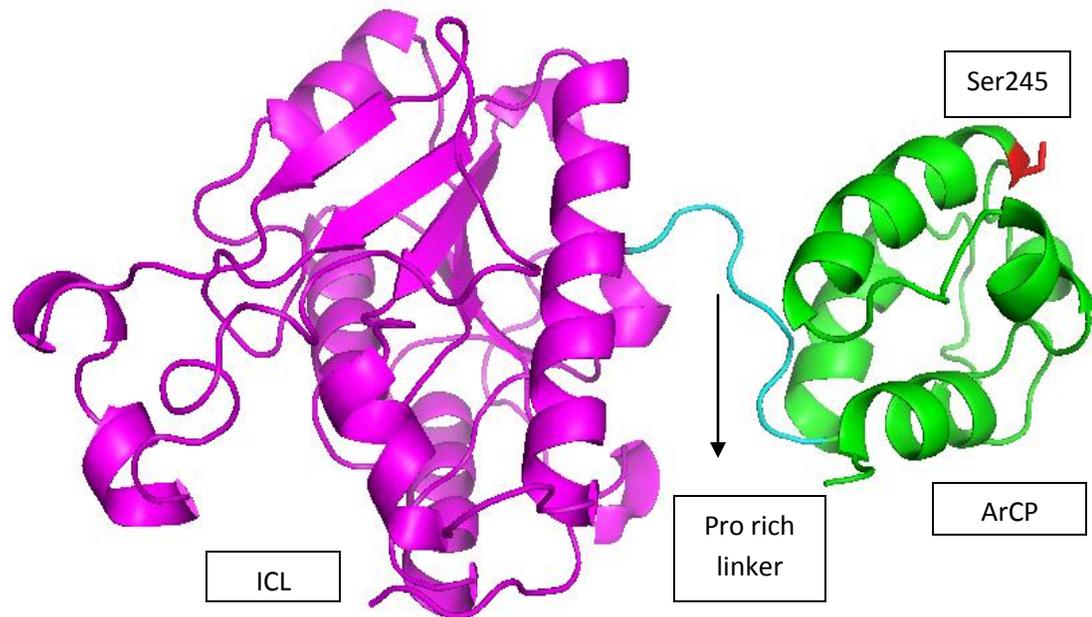


Figure 5 Crystal structure of EntB monomer.

The ICL domain (magenta) and ArCP domain (green) are separated by a 5 amino acid linker region (cyan). The conserved site of phosphopantetheinylation (Ser245) is highlighted in red. (PDB code: 2FQ1).

1.8. Phosphopantetheinyl Transferases

Phosphopantetheinyl transferases (PPTases) are responsible for the activation of carrier proteins. They play an essential role in the activation of fatty acid, polyketide, and nonribosomal peptide biosynthetic pathways by catalyzing the covalent attachment of 4'-phosphopantetheinyl prosthetic group of coenzyme A to a conserved serine residue within the carrier protein domains. This post-translational modification thus converts these proteins from their inactive *apo*-form into their active *holo*-form. The introduction of a covalently attached nucleophilic thiol becomes the site where the synthesis of the enterobactin is based. Within bacteria, PPTases are divided into one of two major enzyme classes in the PPTase superfamily; the AcpS- type and the Sfp-type.

Falling into these two classes is dependent on the primary sequence of the protein. The representatives of these two classes, AcpS from *Escherichia coli* and Sfp from *Bacillus subtilis* are structurally distinct (Roberts *et al.*, 2009). AcpS from *E. coli* was the first transferase to be cloned and characterized (Lambalot *et al.*, 1995). It activates the fatty acid ACP by converting it to its *holo*-form. *E. coli* AcpS is a 28 kDa dimer of two 125 amino acids with a k_{cat} of 80-100 min^{-1} and a K_m of 10^{-6} M for *apo*-ACP (Lambalot *et al.*, 1996). It has been shown that *E. coli* AcpS is capable of activating the *apo*-forms of several ACP homologues such as the DCP, lactobacillus casei D-alanyl carrier protein involved in D-alanylation of lipoteichoic acid and the *Rhizobia* protein, NodF, involved in the acylation of the oligosaccharide-based nodulation factors, and the *Streptomyces* ACPs involved in frenolicin, granaticin, oxytetracycline, and tetracenomycin polyketide antibiotic biosynthesis (Lambalot *et al.*, 1996). AcpS-type PPTases are capable of activating ACPs from polyketide synthases and fatty acid synthesis. However, AcpS is not able to efficiently phosphopantetheinylate *apo*-EntB and *apo*-EntF from the enterobactin biosynthesis module (Lambalot *et al.*, 1996; Gehring *et al.*, 1997). The representative PPTase of the second class is named from the broad range activity PPTase, Sfp. Sfp-type PPTases activate NRPS/PKS multienzymes. Sfp is required for the production of the lipoheptapeptide antibiotic surfactin by converting the inactive *apo*-forms of the seven PCP domains of surfactin synthase (SfrA-ABC) to their active *holo*-forms (Lambalot *et al.*, 1996). It has been shown that Sfp is non-specific since it is able to efficiently catalyze the modification of many PCP and ACP domains other than its substrates. It can efficiently phosphopantetheinylate the *apo*-PCP domains from

enterobactin synthetase (Lambalot *et al.*, 1996), bacitracin, gramicidin S, and tyrocidine synthetases. Also, it is capable of converting the *apo*-ACP domains or subunits from fatty acid synthases such as the *E. coli* type II fatty acid synthase *apo*-ACP subunit, and polyketide synthases. EntD is only specific to the enterobactin synthetase proteins EntB and EntF with a k_{cat}/K_m of 0.8 and 0.3 $\mu\text{M}^{-1} \text{min}^{-1}$ respectively (Lambalot *et al.*, 1996, and Gehring *et al.*, 1997). It can only inefficiently modify *apo*-ACP into *holo*-ACP. Some Sfp-type PPTases include the *Saccharomyces cerevisiae* Lys5 which activates the alpha-aminoadipate semialdehyde reductase Lys2, *E. coli* o195 whose physiological role is still unclear, Sppt from the cyanobacterium *Synechocystis* sp. PCC6803 which activates its cognate fatty acid synthesis carrier protein SACP, and the focus of this paper, *E. coli* EntD.

1.9. Crystal Structure of Sfp

The crystal structure of Sfp was determined in 1999 (Reuter *et al.*) in the presence of its substrate CoA and Mg^{2+} (Figure 6). The Sfp-CoA complex was refined at 1.8 Å with an R factor of 21.3%. It was determined that Sfp is a monomer with a α/β -fold exhibiting an intramolecular pseudo 2-fold symmetry. This pseudosymmetry divides the protein into two similarly folded halves of roughly identical size. The N-terminal half ranges from Met1 to Pro 103, while the C-terminal half extends from Ile104 to Pro209. The 15 other remaining amino acids from Asp210 to Leu224 have no counterpart in the N-terminal half. One half is comprised of a three stranded antiparallel β -sheet (β_1 - β_3 , β_4 - β_7 - β_8), and a long α -helix (α_3 , α_5) packed diagonally against the β -sheet. The structure of the β -

sheets of both halves is arranged like a barrel. The C-terminal half includes a small two-stranded antiparallel β -sheet ($\beta 6$ and $\beta 7$). The major difference between the two halves is found in the position of the loops C-terminally following the long α -helix. Loop $\alpha 3$ - $\beta 2$ in the N-terminal half sticks to the protein body while the corresponding loop $\alpha 5$ - $\beta 7$ in the C-terminal half projects from the rest of the protein body. The substrate CoA is bound in a pocket formed by the interface of the two Sfp halves. CoA is bound in a bent conformation within the pocket which is formed by the residues from strands $\beta 2$ and $\beta 4$. The amino acids around the CoA-binding pocket may be involved in binding of the PCP substrate. The pantetheinyl group of CoA had no defined conformation in the electron density map except for the two atoms next to the pyrophosphate. Therefore, this suggests that the PPant group rises from the CoA-binding pocket into the solvent and makes no interactions with the Sfp. However, the 3'-phospho-5'-ADP moiety is well defined in the electron density map. The 3'-phosphate is held firmly to Sfp through many interactions. It forms hydrogen bonds with the side chain hydroxyl of Thr44 and it salt bridges to the δ -amino groups of Lys28 and Lys31 and to the ϵ -nitrogen of the imidazole group of His90. Additionally, His90 binds to α -phosphate of the pyrophosphate through a salt bridge by δ -imidazole nitrogen and a hydrogen bond by the main chain amide. The δ -amino group of Lys155 is salt bridged to the α -phosphate. The α -phosphate of the pyrophosphate is also hydrogen bonded to the side chain hydroxyl of Ser89. The amino acids Tyr73, Gly74, Lys75, and Pro76 hold the adenine base of CoA in place, where it stacks against the side chain of Lys75. Tyr73 hydrogen bonds with the amino group of the adenine base. Metal complexation occurs via the α -

and β -phosphates of the pyrophosphate, one water molecule, as well as the carboxylate groups of Asp107, Glu109, and Glu151. Mutagenesis experiments have shown that Asp107 is essential for catalysis (Quadri *et al.*, 1998). The metal ion in the complex is assumed to be Mg^{2+} as Lambalot *et al.* (1997) showed that PPTase activity strictly depends on Mg^{2+} . Reuter *et al.* (2006) speculate that this Mg^{2+} ion plays a role in catalysis instead of just cofactor binding. They speculate that the hydroxyl group of the side chain of the conserved serine residue of the substrate PCP replaces the water molecule involved in the Mg^{2+} complexation. The β -phosphate is accessible for a nucleophilic attack by the hydroxyl group of the serine residue due to the lateral gap in the CoA-binding pocket.

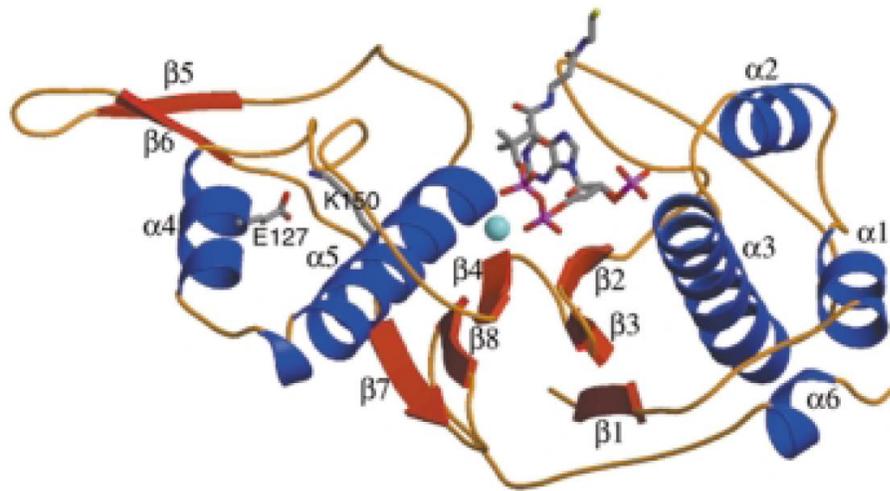


Figure 6 Ribbon representation of Sfp in the presence of Mg^{2+} and CoA.

α -helices are represented as blue and β -sheets as red. The light blue sphere represents Mg^{2+} . The pantetheinyl moiety of CoA is arbitrary since it was not well defined in the electron density map. The conserved residue Lys150 salt bridges with Glu127. Figure from Reuter *et al.* (1999). (PDB code: 1QR0)

1.10. Sequence Alignment of Phosphopantetheinyl Transferases

Figure 7 shows the sequence of Sfp aligned with four other Sfp-type PPTases. The highly conserved residues are shaded in orange. PPTases share a low sequence identity among themselves. For example, Sfp and EntD have a percent identity of 15%. The region between $\alpha 5$ and $\beta 7$ is proposed to be the substrate PCP binding region. However, this region in EntD is shortened which could explain the fact that EntD has a reduced protein substrate spectrum compared with Sfp. The region between $\alpha 5$ and $\beta 7$ could be responsible for the wide spectrum of protein substrates of Sfp in recognizing more structural features of the PPant acceptor instead of the specific amino acid side chains. The conserved residues are the residues that are involved in Mg^{2+} and CoA binding in Sfp. The CoA adenine base binding sequence motif Gly74-Lys75-Pro76 is conserved among the Sfp-type PPTases except in *E. coli* EntD. The glycine residue, which has no direct interactions with the adenine base is replaced with a cysteine. *E. coli* EntD also has a valine residue that replaces Lys75 from Sfp. Lys75 has stacking interactions with the adenine base which could be substituted by the side chain of valine. Trp147 in Sfp has no direct interaction with CoA and yet is still conserved among all PPTases including the AcpS-type PPTases (Figure 8). Trp147 could be replaced by phenylalanine as observed in *E. coli* EntD. The residue in this position shields the active center from hydrophobic residues by van der Waals interactions and therefore, it is required to have a bulky side chain. Figure 8 demonstrates the invariant and conserved residues between Sfp and the AcpS-type PPTases. Asp107 which is involved in Mg^{2+} binding is conserved among all AcpS-type PPTases and Glu 109 which is also involved in Mg^{2+} complexation is

only partially conserved. In these PPTases Glu109 is replaced by a hydrophobic residue such as valine or isoleucine. Gly105 and Lys150 are not involved in CoA binding yet they are still highly conserved among all PPTases. The residue in that position is confined to the small amino acid glycine due to spatial reasons. The side chain carboxylate of Glu127 salt bridges with the δ -amino group of Lys150. However, it is still unclear if Glu127 is a conserved residue due to the very low conservation of the amino acids around it.

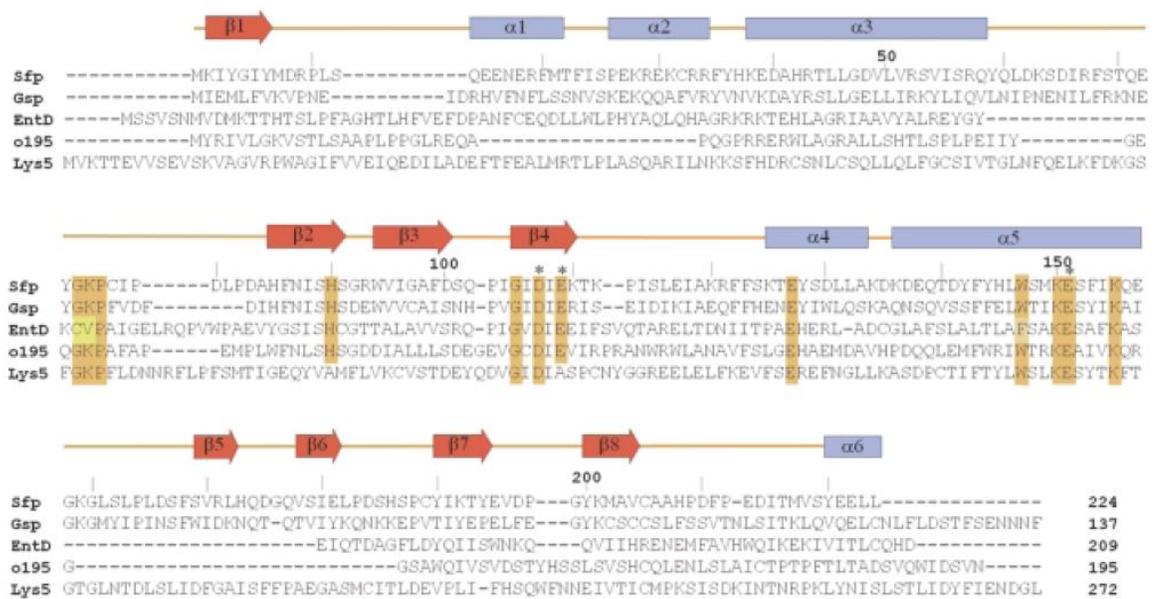


Figure 7 Sequence alignment of Sfp with EntD and 3 other PPTases.

Highly conserved and invariant residues are highlighted in orange. Residues occupying conserved residues but differ from the consensus amino acid are shaded yellow. The three acidic amino acids of Sfp involved in Mg^{2+} complexation are marked by an asterisk (*). Figure from Reuter *et al.* (1999).

Sfp	72	EYGKFCIP-----DLPDAHFNISHSGR---9aa--PIGIDIE--35aa--HLWSMKEAFIKQE	157
Lpa14	72	EYGKFIIP-----ALPDMHFNISHSGR---9aa--PIGIDIE--35aa--HLWSMKEAFIKQA	157
Psf1	75	GNGKPVVR-----QIPSFHFNLSHSGD---9aa--PVGIDIE--35aa--HLWSMKEAFIKLT	160
Gsp	73	EYGKFEVDF-----DIHFNLSHSD---9aa--PVGIDIE--35aa--ELWTKESYIKAI	156
HrpG	83	EYGKFFILNESKE-----SIYFNLSHSNN--10aa--SVGIDIE--37aa--KMWTLKEAYIKSR	173
ORF2	27	DKGKFRLSDAQFAQ-----TGVHENVSHSGD--25aa--EFGVDIE--37aa--DLWALKEAYIKAK	136
Lys5	94	SFGKFFLDNN-----28aa-----DVGIDIA--37aa--YLWNLKESYTKET	188
o195	55	EQGKFAFAP-----EMPLWFNLSHSGD--10aa--EVGCDIE--37aa--RINTRKEAIVKOR	143
HI0152	73	ESGRPFYFLDER-----IDFNISHSGD--14aa--AVGIDIE--34aa--RCWCLREAVLKSQ	161
Slr0495	72	PQKKEELVDR-----ERRSPWFNVASHGN--10aa--EIGVDLQ--37aa--QLNTAKEAFLKAT	162
c17c9.02c	77	KAGRPYCQSAHCPPII-----FDENVSHYGG--19aa--NIGVDIV--38aa--LLWTCKEAILKAL	179
HetI	89	SRGKFIIGDR-----FAESGLLFNLSHSGD--10aa--QIGIDLE--37aa--RYWTCKEAIVKAT	180
T28H10.1	77	ENGKPSLIQNKSDYS----RQNFYENVSHHGD---9aa--RIGVDVM--44aa--RIWCLKESILKAT	177
T04G9.4	84	ERKPFPLAVPAD-----TTFGLNVSHQGD--10aa--KVGVDVM--44aa--RYWCLKEAILKAT	181
T20F6.9	431	MYGKPEVDWQSYKN---CDSPPLHFNISHSDS--10aa--PVGIDLE--40aa--KLWTLKEAYWKAL	529
SC3F7.03c	115	QGRPFVVD-----APVEVGLTHSGD---9aa--AVGVDLE--35aa--LRWTCKEAVLKYY	198
SC5A7.23	58	KLGVPAQPVLTGERGAPRWPEGIAGSMTHCDG--13aa--SLGIDAE--36aa--LLFSAKESVYKAW	158
Nsh-OrfC	81	HAIVLLPGPRGE---PRWPDGVVGSILTHCAG--14aa--GVGIDAE--36aa--LLFCVKEAVFKAW	179
ORF3	46	ELGIDFVPLPPGRHRAPQWPTGVVGSMTHCSDG--13aa--SVGIDAE--38aa--LLFSCKEAVYKAW	146
EntD (E. c.)	75	YKCVPAIGELR----QFVWPAEYVGSISHCGT---9aa--PIGIDIE--36aa--LAFSAKESAFKAS	167
EntD (S. f.)	69	YKCVPAIGELR----QFVWPAEYVGSISHCGT---9aa--PIGIDIE--36aa--LAFSAKESAFKAS	161
EntD (S. t.)	65	ASGVPIGDKR----QPLNPDGLFGSISHCAT---9aa--RIGIDIE--35aa--LAFSAKESVYKAF	156
Rv2794c	65	QLGVPAIPILKGDKEPCNPDGMVGSILTHCAG--13aa--SVGIDAE--34aa--LLFCAKEATYKAW	163
ACPS (S. a.)	4	GIGVDLI--42aa--GRFATKEAFSKAL	65
ACPS (B. s.)	4	GIGLDIT--41aa--GRFAAKEAFSKAF	64
ACPS (M. t.)	5	GVGIDLV--40aa--ARWAAKEAVIKAW	64
Ppt1	9	TVGVDLV--52aa--GRWAAKEAFIKAW	80
ACPS (R. p.)	4	GVGTDIV--42aa--KRFSAKEAVSKAF	65
ACPS (E. c.)	5	GLGTDIV--40aa--KRFAVKEAAAKAF	64
ACPS (T. p.)	4	GVGIDIV--38aa--TRFAAKEAFKAL	61
ACPS (S. c.)	5	GVGIDVA--39aa--ARFAAKEALAKAL	63
ACPS (B. j.)	4	GIGSDLI--44aa--KRFAAKEACSKAL	67
ACPS (C. t.)	5	GVGIDII--40aa--ARFAAKEAVAKAL	62
ACPS (A. a.)	3	MIGVDIV--40aa--ARWAGKEAVLKAF	60
ACPS (H. p.)	3	MIGDIDV--37aa--GFFALKEACSKAL	57
ACPS (B. b.)	3	SIGCDII--39aa--GKFAAKESLIKAL	61
Ppt2	18	GVGVDIV--58aa--GVWALKECSLAKAL	95
FAS2 (S. p.)	1724	NVGVDVE--36aa--GRWSAKEAVFKSL	1779
FAS2 (S. c.)	1775	GVGVDVE--36aa--GTWSAKEAVFKSL	1833
FAS2 (P. g.)	1739	KIGVDVE--36aa--GRWSAKEAVFKSL	1794
FAS2 (C. a.)	1767	GVGVDVE--36aa--GTWSAKEAVFKAL	1822
FAS2 (E. n.)	1742	KIGVDVE--36aa--GRWSAKEAVFKSL	1797
StcJ	1428	TIGVDTV--39aa--SGWCAKEAVFKCL	1486

Figure 8 Partial sequence alignment of PPTases of the Sfp-type (top) and AcpS-type (bottom).

Invariant and conserved residues are indicated as in Figure 7. Figure from Reuter *et al.* (1999).

1.11. EntH: Prevention and Repair Protein

The gene for the hotdog-fold thioesterase EntH (YbdB) is found within the cotranscriptional unit which includes the genes for EntC and EntA of the upper pathway, and EntE and EntB of the lower pathway. Chen *et al.* recently demonstrated EntH

structure-function analysis (2009). They determined the substrate specificity of EntD and EntE to provide a better understanding on the biochemical function of EntH. EntD and EntE substrate promiscuity gives rise to misacylated *holo*-EntB. Acyl-CoA has been demonstrated by Chen *et al.* (2009) to be a substrate for EntD when CoA is scarce, thus resulting in a misacylated *holo*-EntB. Percent conversion was determined by ESI-MS by calculating the ratio between the acylated *holo*-EntB peak and the combination of the EntB peak and acylated *holo*-EntB peak. EntD is also able to completely convert *acyl-holo*-EntB with reactions of lauroyl-CoA, palmitoyl-CoA, and 4-hydroxybenzoyl-CoA. For the reactions with isobutyryl-CoA and hexanoyl-CoA, a conversion of 23% and 67% is observed, respectively (Chen *et al.*, 2009). The percent conversion of *holo*-EntB to *acyl-holo*-EntB from various aromatic compounds was also determined by Chen *et al.* A 100% conversion was obtained for the reactions of 2,3-DHB and 2,4-DHB while only 47, 25, 9, 14, and 0% conversion for reactions of 4-hydroxybenzoate, 3-hydroxybenzoate, 2-hydroxybenzoate, benzoate, and phenylacetate was observed. EntH can prevent EntD-catalyzed misacylation of EntB through hydrolysis of endogeneous acyl- or aroyl-CoAs. It has been suggested that EntH harvests CoA from endogeneous acyl- and aroyl-CoAs and therefore helps with the EntD-catalyzed formation of *holo*-EntB and prevents the formation of misacylated *holo*-EntB. EntH binds to *holo*-EntB and not to *apo*-EntB (Leduc, 2007). EntH also binds to the phosphopantetheine arm of the thiol unit and not to the adenine base of CoA. It is likely that EntH recognizes the acyl-pantetheinephosphate and aroyl-pantetheinephosphate of CoA or EntB thioester substrate. Also, EntH discriminates potential substrates based on their acyl or aroyl unit

since EntH has a preference for aromatic substrates against aliphatic substrates. The reaction catalyzed by EntH to remove 2,3-DHB-*holo*-EntB is much slower than the EntF-catalyzed reaction. The faster reaction catalyzed by EntF prevents hydrolysis of the correctly acylated *holo*-EntB by EntH. The slower rate of EntF with misacylated 2,4-DHB-*holo*-EntB gives enough time for EntH to hydrolyze the misacylated *holo*-EntB.

1.12. Protein-Protein Interactions in the NRPS Module

The NRPS module is responsible in converting 3 molecules of DHB and L-serine into enterobactin. Protein-protein interactions within the NRPS module are required for the channeling of the biosynthetic intermediates. EntB ArCP domain needs to interact with three proteins during the enterobactin biosynthetic reaction cascade. First, it needs to participate in protein-protein interaction with EntD during phosphopantetheinylation. Second, the EntB-ArCP domain interacts with EntE during activation of DHB and thiolation onto the phosphopantetheine arm of *holo*-EntB. And finally, EntB ArCP domain interacts with EntF during condensation of DHB with serine.

Indirect evidence for protein-protein interactions with EntB ArCP domain. Indirect evidence utilizing combinatorial mutagenesis and selection experiments have identified different protein interaction interfaces on ArCP domain of EntB that recognizes EntD and EntF. The N-terminal portion of the ArCP domain of EntB has been identified to be the interface for protein-protein interaction with EntD. Specifically, the EntB-ArCP residues G242 and D244, which immediately precede the phosphopantetheinylated S245, are essential for EntD to convert *apo*-EntB into *holo*-EntB. Mutations of these two

residues result in a disruption of EntD recognition and conversion (Lai *et al.*, 2006). The C-terminal portion of EntB-ArCP is responsible for the recognition of EntF. The three residues on EntB M249, F264, and A268 following S245 constitute an interface for interaction with EntF. These residues were highly conserved in mutagenesis and selection experiments (Lai *et al.*, 2006). EntB-ArCP has an initial rate of 36 min^{-1} for the production of enterobactin *in vitro*, whereas M249A, F264A, and A268Q mutations resulted in a decrease in the rate constant with a k_{obs} of 4, 9, and 0.4 min^{-1} , respectively (Lai *et al.*, 2006). These mutations also greatly decreased the kinetic parameters for the condensation of DHB and serine by EntF. M249A, F264A, and A268Q had a reduced $k_{\text{cat}}/K_{\text{m}}$ of 6, 72, and $2 \text{ min}^{-1} \mu\text{M}^{-1}$, respectively, compared to EntB-ArCP with a $k_{\text{cat}}/K_{\text{m}}$ of $410 \text{ min}^{-1} \mu\text{M}^{-1}$. However, M249 and A268 mutations did not affect the ability of the EntB mutants to recognize EntE upon acylation as the wild-type (WT) kinetic parameters were similar to those of the mutants. The acylation kinetics of F264A mutation was not studied since they determined that it only had a moderate deficiency in enterobactin production. The ability of the F264A and A268Q mutants to become phosphopantetheinylated by EntD was also carried by Lai *et al.* (2006) to determine the specificity of these residues to recognize EntF. EntB F264A was accepted as efficiently as WT EntB and A268Q was accepted at a 2 fold lower rate suggesting that F264 residue is not involved in EntD recognition and that A268 may be involved to a small degree. EntB M249A had an increased rate of phosphopantetheinylation by EntD. F264 was also identified by Drake *et al.* (2006) to have an elevated activity for the acylation of *holo*-EntB. Furthermore, Drake *et al.* (2006) mutated 9 residues in the ArCP domain of EntB

that are centrally located near Ser245 and positioned in such a way that they might introduce charge or hydrophobic contacts with EntE. They determined the efficiency of acylation of *holo*-EntB. Only two residues, Asp240 and Asp263, demonstrated a decrease in activity when compared to wild-type EntB. To confirm that these mutants did not affect the ability of EntD to phosphopantetheinylate EntB, *in vivo* pantetheinylation of EntB was achieved by producing protein in minimal media.

Direct Evidence for EntB-EntE interactions. EntB-EntE interaction has also been further analyzed by mutating three residues on EntE (Drake *et al.*, 2006). These mutations constituted of R437D, K473D, R494D, and a double mutant comprised of R437D and K473D. All four mutants were capable of catalyzing the exchange of radiolabeled pyrophosphate into ATP and therefore were competent to catalyze the adenylation reaction. However, the R437D mutant and the double mutant R437D/K473D had a greatly reduced activity in assays of the complete thioester-forming reaction. Direct evidence of EntB-EntE interactions have been reported by using spectrofluorometry techniques (Khalil *et al.*, 2009). Far-UV, and near-UV circular dichroism experiments demonstrated that the interaction is favored only in the presence of 2,3-DHB. This finding was further supported by pull-down assays. EntB as bait protein was able to interact more efficiently with chromosomally expressed EntE in the presence of exogenously added 2,3-DHB. Interestingly, 2,3-DHB can also bind to EntB, which is not a substrate to either of the two domains of EntB. This binding to EntB suggests that it is allosterically regulated by 2,3-DHB by feedback inhibition. Since the reaction catalyzed by EntC is reversible, the committed step in enterobactin

biosynthesis is believed to be the conversion of isochorismate to 2,3-dihydro-dihydroxybenzoate by the ICL domain of EntB. It is not unreasonable to suggest that EntB is regulated by excess amounts of intracellular DHB. Therefore, the DHB module will cease to produce DHB, allowing enough time for the NRPS module to utilize DHB and converting it into enterobactin by binding it to serine. Furthermore, a two-hybrid method has been able to detect an interaction *in vivo* between EntB and the repair protein EntH (Leduc *et al.*, 2007). Copurification of these two proteins further confirms their interaction. Two-hybrid and copurification experiments also demonstrated that interactions between the ArCP domain of EntB and EntH are dependent on the presence of the phosphopantetheinyl group on EntB.

Protein-protein interactions with EntF. Combinatorial mutagenesis and selection was also used to identify interdomain interaction between the thiolation and thioesterase domain of EntF (Zhou *et al.*, 2006). Two residues, G1027 and M1030, in the EntF PCP domain were identified to be responsible for PCP-TE interactions. These PCP domain mutants were recognized by the C domain in *cis* as suggested by the ability of the C domain to transfer DHB to Ser-S-EntF. However, the mutant EntF PCP domains were unable in transferring the DHB-S-Ser moiety to the TE domain. This led to the conclusion that these residues are responsible for interdomain interaction between the PCP and TE domain of EntF. Frueh *et al.* (2008) used NMR experiments to give a more structural understanding of the PCP-TE domain interactions. The PCP domain contacts the TE domain by the beginning of helix α 2PCP that contains S48, the site of phosphopantetheinylation, the end of loop L3PCP, and the helix α 1'PCP. The region

around S48 interacts with the core of the TE domain. The two serine residues, S48 of PCP and S180 of TE domain, are 17 Å apart, which allows the 20 Å long 4'-PPant arm tethered to S48 to reach S180. Two helices, α_4 and α_5 , on the TE domain form a mobile flap that can open to accommodate the PPant group. The mobile loop L3PCP of the PCP domain interacts with β_1 of the TE domain. Additionally, they were able to illustrate the interaction between EntD, and EntF C domain with the PCP and TE domains of EntF. As expected, when EntD and its homologue Sfp were titrated into a solution of PCP-TE, the binding site was identified to be around the active site Ser48. Also, PCP and TE domain interacted very little or not at all together suggesting that PPTases drive the equilibrium towards the open form of PCP-TE domain. This open form will allow the PCP domain to be primed with the PPant arm. Similarly, EntF C domain was titrated into a solution containing PCP-TE domains and chemical shift changes were observed. The shift changes corresponded to EntF C domain interacting with the PCP domain. Only minor effects were observed for the TE domain.

1.13. Fur Protein

High levels of intracellular iron can lead to the production of harmful levels of the hydroxyl radical under aerobic conditions. Therefore, the ferric uptake regulator (Fur) protein is required for iron homeostasis. Fur is a 17 kDa protein and it functions as a homodimer. The quaternary structure of Fur in solution is a dimer regardless of the presence or absence of Fe^{2+} (Escobar *et al.*, 1999). One Fe^{2+} ion binds to each monomer. In iron-rich conditions, Fur binds Fe^{2+} resulting in a conformational change allowing it to

bind to the target DNA sequence known as the iron box. Binding of the Fe^{2+} -Fur complex inhibits transcription of all *Ent* cluster genes. Additionally, Fur regulates the *fiu* and *cir* genes, which are the genes for the receptor proteins that recognize Ent-related siderophores. Under iron-limiting conditions, inadequate levels of Fe^{2+} prevent holorepressor formation. The iron box is 19 base pairs consisting of three hexamers. Each hexamer has a GATAAT sequence and the third hexamer is in reverse order and separated from the first two by a single nucleotide (Earhart., 2004).

1.14. Enterobactin Secretion and Ferric Enterobactin Uptake

The enterobactin secretion mechanism still remains uncharacterized. The charge and large size of enterobactin eliminate the possibility of passive diffusion across the membrane. Therefore, an export pump is required to release enterobactin from the cell. Indeed, previous studies (Furrer *et al.*, 2002; Bleuel *et al.*, 2005) demonstrated that cells with the *ybdA* gene (EntS) deleted released decreased amounts of enterobactin. EntS is a 43 kDa membrane protein homologous to the major facilitator superfamily (MFS) class of the proton motive force (PMS)-dependent membrane efflux pumps. EntS mutants also release enterobactin inefficiently, and they release higher amounts of the secondary siderophores: DBS_3 , DBS_2 , and DBS. The release of these secondary siderophores suggests that they are exported with a different and yet unknown mechanism. As illustrated in Figure 9, once enterobactin is translocated across the cytoplasmic membrane by EntS and other unknown mechanisms, it is then most likely accepted by an efflux complex comprising of the outer membrane factor TolC (Bleuel *et*

al., 2005). Similarly to *entS* gene deletions, *tolC* deleted cells were unable to efficiently export enterobactin. Unlike the enterobactin export mechanism, the ferric enterobactin import system has been extensively characterized. The 80 kDa outer membrane protein, FepA, binds to ferric enterobactin and transfers it across the outer membrane with energy derived from three proteins, TonB, ExbB, and ExbD (Bleuel *et al.*, 2005). The secondary siderophores are transported into the periplasm by the Fiu and Ciu outer membrane proteins (Hantke, 1990). Inside the periplasm, ferric enterobactin will be concentrated and presented to the cytoplasmic membrane proteins by binding to the periplasmic binding protein FepB. Transfer of ferric enterobactin into the cytoplasm then occurs by means of four cytoplasmic membrane proteins FepC, FepD, FepE and FepG (Figure 9).

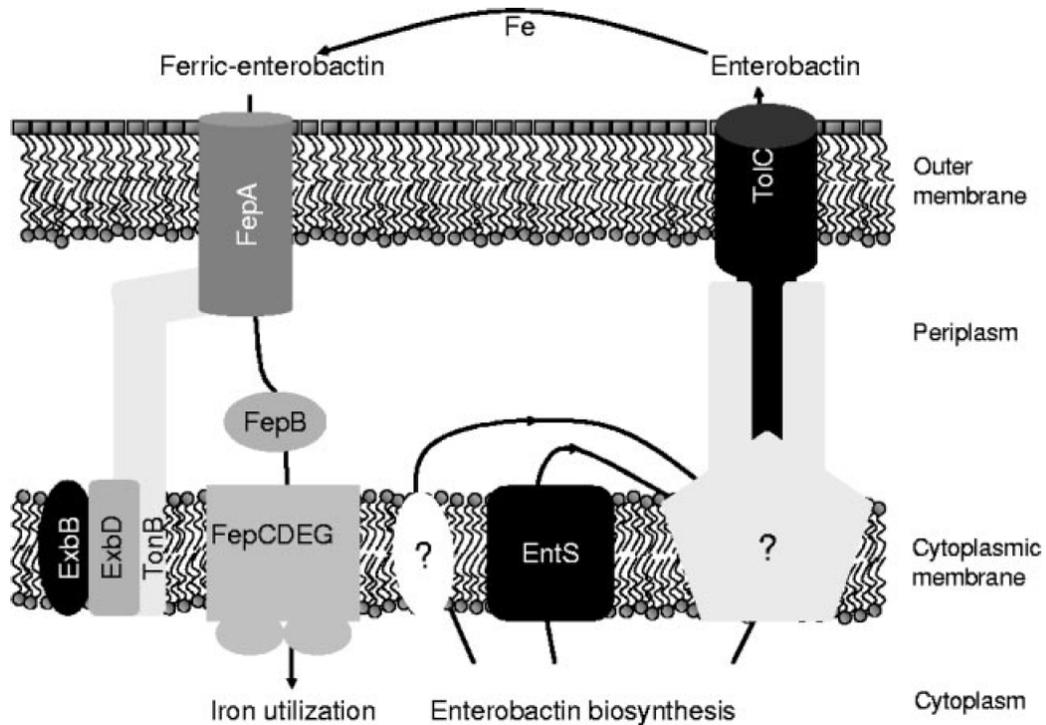


Figure 9 Proposed model of enterobactin transport in *E. coli*.

Export of enterobactin into the periplasm takes place via the cytoplasmic membrane protein EntS. Unknown protein complex may then transfer enterobactin to TolC. The outer membrane protein TolC then excretes enterobactin from the cell. Upon iron binding, ferric enterobactin binds to FepA. Through the energy provided by the proton motive force of TonB, ExbB and ExbD, ferric enterobactin is transferred to the periplasm. FepB then presents it to the cytoplasmic membrane proteins FepCDEG for transfer to the cytoplasm. Figure from Bleuel *et al.* (2005).

1.15. Mammalian Siderophores

Recent evidence of mammalian siderophores has emerged from studies of the lipocalin 24p3 (Devireddy *et al.*, 2010). Lipocalins are proteins that can bind small molecular weight ligands. 24p3 can bind iron and transport in and out of the cells. Iron-bound 24p3, *holo*-24p3, is recognized by the cell surface receptor 24p3R, which then internalizes *holo*-24p3 and then releases the bound iron. 24p3R also recognizes *apo*-24p3R and internalizes it where *apo*-24p3 will chelate intracellular iron and transfer it to

the extracellular medium. Interestingly, 24p3 does not interact with iron. Instead, 24p3 is associated with the small molecular weight iron-chelating compound 2,5-DHB which binds to the iron. Therefore, 2,5-DHB is the iron chelating compound of 24p3.

1.16. Research Objective

The primary goal of this research was to purify and characterize the stability, conformation, and activity of the phosphopantetheinyl transferase EntD from *E. coli* enterobactin synthetase in the presence and absence of its cofactors Mg^{2+} and CoA. PCR based site-directed mutagenesis altering two residues around the active site will help identify the ligand and substrate binding residues. The structure and stability of the variants was compared with wild-type EntD by circular dichroism, temperature denaturation, and fluorescence spectroscopy. The activity of wild-type EntD and variants on converting *apo*-EntB to *holo*-EntB was determined by ESI-MS and a gel-based assay. Furthermore, binding of CoA to all three EntD proteins was analyzed by conjugating CoA to a fluorescent dye and using fluorescence polarization. The secondary goal of this research was to overexpress, purify and characterize the stability of wild-type EntB and two variants. Two residues on EntB ArCP domain identified by phage-display to interact with EntE were changed into alanine. The stability of these variants was compared with wild-type by native polyacrylamide gel, and ESI-MS. Additionally, the ability of EntD to pantetheinylate EntB variants was studied by ESI-MS.

2. Material and Methods

2.1. Strains, Growth Media, and Buffers

All pCA24N expression vectors were acquired from the ASKA repository (Kitagawa *et al.*, 2005). Growth media and the most commonly utilized buffers used in this study are listed in Table 1. All chemicals and reagents were from Bioshop unless otherwise indicated. Chloramphenicol and IPTG (isopropyl- β -D-thiogalactopyranoside) were prepared as 30 mg/ml and 1M stock solutions respectively and filtered through 0.2 μ m sterilized syringe filter and stored at -20 °C in 1 ml aliquots.

Growth Media and Buffers	Recipe
LB	1% Bacto-tryptone 0.5 Bacto-yeast extract 1% NaCl (Plates: 1.5% agar)
5X PAGE-Loading Buffer	0.25 M Tris-HCl (pH 6.8), 0.25M β -mercaptoethanol, 0.25% bromophenol blue, 5% SDS, and 25% glycerol
SDS-PAGE Resolving Gel (12%)	0.375 M Tris-HCl (pH 8.8), 12% Acrylamide (29:1 acrylamide:bis acrylamide), 0.1% SDS, 0.1% APS, 0.1% TEMED
SDS-PAGE Stacking Gel (4%)	0.125 M Tris-HCl (pH 6.8), 4% acrylamide (29:1 acrylamide:bis acrylamide), 0.1% SDS, 0.1% APS, 0.1% TEMED
SDS-PAGE Running buffer	Tris 3.5 g/L, Glycine 14.4 g/L, 0.1% SDS
TAE	Tris 4.8 g/L, 0.02M glacial acetic acid, 1 mM EDTA
TBE	Tris 10.8g/L, Boric acid 5.5 g/L, EDTA 0.74 g/L
PBS	NaCl 8 g/L, KCl 0.2 g/L, KH ₂ PO ₄ 0.24 g/L, 1.44 Na ₂ HPO ₄

Table 1 Recipes for growth media and buffers.

2.2. Preparation of Competent Cells

Competent *E. coli* AG-1 cells were prepared based on a protocol described by Sambrook et al., (2001). *E.coli* AG-1 cells were streaked on fresh LB agar plate without antibiotic and incubated overnight at 37 °C. A single colony was selected and inoculated into 5 ml LB medium. This starter culture was grown overnight at 37 °C with shaking at 250 rpm. The overnight culture was then diluted 60 fold with fresh LB medium and incubated with shaking at 250 rpm until an OD₆₀₀ of 0.6 had been reached. The culture was transferred to a sterile 50 ml falcon tube and left on ice for 30 minutes. Afterwards, the cells were harvested by centrifugation in a JA-17 rotor at 3000 rpm for 15 minutes at 4 °C. The supernatant was discarded and the cell pellet was gently resuspended in 10 ml of freshly prepared ice-cold sterile 0.1 M CaCl₂. The resuspended cell pellet was incubated on ice for 15 minutes and centrifuged as before. The supernatant was discarded and the cell pellet was resuspended in 2 ml of ice-cold sterile 0.1M CaCl₂ with 15% glycerol. The competent AG-1 cells were transferred to pre-chilled sterile microcentrifuge tubes in 0.1 ml aliquots and frozen at -80 °C.

2.3. *E.coli* Transformation

50 µl of competent *E. coli* cells were incubated with 5 µl of plasmid DNA on ice for 30 minutes. The cells were heat shocked at 42 °C for 75 seconds. Next, the cells were placed on ice for 2 minutes and grown in LB medium shaking at 37 °C for 1 hour at 225 rpm. The cells were transferred to a sterile microcentrifuge tube and centrifuged at 3000 rpm for 3 minutes. 800 µl of the supernatant was discarded and the cells were

resuspended with the remaining 200 µl of supernatant. The resuspended cultures were then plated on LB agar + chloramphenicol (30 µg/ml) plate and grown overnight at 37 °C. Transformation protocol modified from Sambrook et al., (2001).

2.4. Agarose Gel

1% agarose gels were prepared by dissolving agarose in 1X TBE buffer. Ethidium bromide was added to a final concentration of 1 µg/ml before the gel solidified. Samples were mixed with Blue-Orange 6X loading dye (Promega) and separated on the gel at 100 volts for 40 minutes. 1 kb DNA Ladder (Promega) was used as standard.

2.5. Plasmid DNA Extraction

Plasmid DNA was extracted using the Qiaprep Spin Miniprep kit by Qiagen. The protocol used was based on the manual provided along with the system. *E. coli* AG-1 cells carrying the plasmid of interest were grown overnight at 37 °C with shaking in 5 ml LB medium containing 30 µg/ml chloramphenicol. Aliquots of 1.5 ml were collected in microcentrifuge tubes by centrifugation for 3 minutes at 5000 rpm and the supernatant was discarded. The pellet was resuspended in 250 µl buffer P1. Next, 250 µl of lysis buffer P2 was added and the samples were inverted several times gently. Afterwards, 350 µl of neutralization buffer N3 was added and mixed thoroughly by inverting the tubes. After centrifugation for 10 minutes at 17,900 x g, the supernatant was applied to the QIAprep spin column by pipetting. After a brief one minute centrifugation, the flow-through was discarded. The QIAprep spin column was then washed by adding 0.75 ml buffer PE and centrifuged for one minute. The flow-through was discarded and the

QIAprep spin column was centrifuged for an additional 1 minute to remove residual wash buffer. The QIAprep column was placed in a clean 1.5 ml microcentrifuge tube. The DNA was eluted by adding 50 μ l of water to the center of the QIAprep spin column and centrifuging for another minute after incubating it for one minute. DNA concentrations were determined by UV absorbance assuming an A_{260} of 1 equals 50 μ g/ml of dsDNA.

2.6. Site-Directed Mutagenesis

Site-directed mutagenesis was conducted using the protocol outlined in the Quickchange™ XL Site-Directed Mutagenesis Kit. PCR based site-directed mutagenesis was used to mutate two residues each from EntB and EntD. The oligonucleotides were designed based on the desired mutations and purchased from Biocorp Inc. (Montreal). They were suspended in distilled water to obtain a concentration of approximately 100 μ M. The concentrations were determined based on A_{260} . Each incubation contained 10 ng of double stranded wild-type recombinant EntB or EntD DNA, 0.5 μ M of forward and reverse oligonucleotides, 1 μ l of 1 mM dNTP solution (Promega), 5 μ l of 5X *pfu* buffer, and 1 μ l of *pfu* polymerase (2.5u/ μ l from Fermentas), in a final volume of 50 μ l. After denaturing the template DNA for 1 minute at 95 °C, 18 PCR cycles were performed using the following parameters:

- 1) Denaturation: 95 °C, 1 minute
- 2) Annealing: 52 °C, 1 minute
- 3) Extension: 72 °C, 15 minutes

After completion of the PCR cycles, the reaction underwent a final extension for 15 minutes at 72 °C. Afterwards the reaction mixture was removed from the PCR tubes and kept at 4 °C. The reaction mixture was incubated with 1 µl of Dpn I restriction enzyme (10 U/µl) at 37 °C for two hours to digest the parental supercoiled dsDNA. The amplified PCR product was confirmed by electrophoresis on a 1% agarose gel containing ethidium bromide, where the DNA was visualized under ultraviolet light. The PCR product was purified and concentrated using the Wizard® SV Gel and PCR clean-Up System from Promega as described in section 2.7. The purified and concentrated DNA was transformed into competent DH5α *E. coli* cells using the same transformation procedure as described in section 2.3. Three colonies were selected and inoculated into 5 ml LB media containing chloramphenicol (30 µg/ml). The plasmid DNA was extracted as described in section 2.5 and the sequence was confirmed by sending the plasmid DNA to the McGill University and Genome Quebec Innovation Center for DNA sequencing.

2.7. PCR Clean-up

The PCR product was purified using the Wizard® SV Gel and PCR clean-Up System from Promega. Briefly, 40 µl of PCR product was mixed with 40 µl of Membrane Binding Solution (4.5 M guanidine isothiocyanate, and 0.5 M potassium acetate (pH 5.0)). The mixture was transferred to the SV Minicolumn assembly and incubated for 1 minute at room temperature. After centrifuging at 14,000 rpm for 1 minute and discarding the flow-through, the column was washed with Membrane Washing Solution (10 mM potassium acetate (pH 5.0), 80% ethanol, and 16.7 µM EDTA (pH 8.0)). Finally, the

column was transferred to a clean 1.5 ml microcentrifuge tube and 10 μ l of nuclease-free water was added and centrifuged as before to elute the DNA.

2.8. Sodium dodecyl sulfate – polyacrylamide gel electrophoresis (SDS-PAGE)

Denaturing SDS-PAGE was used to estimate the molecular weight of proteins. The gels consisted of various resolving gel concentrations (8-12%) and 4% stacking gel. The gels were made according to Sambrook and Russell (2001). 30% acrylamide/bisacrylamide solution 29:1, Precision Plus Protein Kaleidoscope Standards, and Coomassie Brilliant Blue G-250 dye were acquired from Bio-Rad Laboratories. 10% ammonium persulfate was prepared and stored at 4 °C. 10% SDS was prepared in dH₂O and kept at room temperature. The gels were cast between 7 cm by 10 cm gel plates separated by 0.75 mm spacers. Protein samples were diluted with 5X loading buffer and boiled for 10 minutes at 90 °C to ensure complete denaturation. The samples were centrifuged for 10 seconds and placed on ice before loading on the gel. Electrophoresis started at 100 V to allow passage of the sample through the stacking gel. The voltage was increased to 150 V once the sample migrated from the stacking to the resolving gel. Electrophoresis was stopped once the bromophenol blue tracking dye reached the bottom of the gel. The gel was washed a total of three times with dH₂O to remove any residual detergent. The proteins were visualized by staining with the Blue Silver stain for 30-60 minutes. Excess dye was removed by washing with dH₂O for 60 minutes.

2.9. Native PAGE

Native proteins were analyzed using non-denaturing polyacrylamide gel electrophoresis. These gels were prepared in the same manner as SDS-PAGE (Section 2.8) but with the omission of any denaturing chemicals or steps. SDS was omitted for the preparation of the resolving gel and the running buffer and the protein samples were not heated. The stacking gel was also omitted and the resolving gel and running buffer were made with TAE buffer to obtain a higher resolution. β -mercaptoethanol was also omitted from the 5X loading dye. The voltage was kept at 80V to help prevent any protein denaturation.

2.10. Protein Expression

E. coli AG-1 cells carrying the plasmid of interest were streaked onto LB + chloramphenicol (30 μ g/ml) plates and incubated overnight at 37 °C. A single colony was selected and inoculated into 10 ml LB + chloramphenicol and incubated overnight at 37 °C with shaking at 225 rpm. The cells were 100-fold diluted and grown for 3-4 hours until an OD₆₀₀ of 0.6 was reached. Initially, the cells were induced with 1 mM IPTG and incubated at 37 °C for four hours. Due to expression of GroEL as a protein contaminant during EntD purification, the induction conditions were altered. In the revised protocol, cells were induced with 0.5 mM IPTG and incubated overnight at 25 °C.

2.11. Cell Lysis

Initially, Bugbuster Protein Extraction Reagent from Novagen was used for cell lysis. Harvested cells were resuspended in Bugbuster reagent at room temperature, using 5 ml of Bugbuster reagent per gram of wet cell paste. At this point, protease inhibitor cocktail from Bio-Rad Laboratories and PMSF (1 mM) were added to the cell suspension. The cell suspension was then incubated at room temperature for 40 minutes. Soluble proteins were separated from the insoluble cell debris by centrifugation at 43,000 x g in a JA-20 rotor at 10 °C for 60 minutes. The soluble extract was then filtered with a 0.2 micron syringe filter and then it was loaded directly onto various protein purification resins. Afterwards, once the active ingredient in Bugbuster was determined to be β -Octyl-thiolglucopyranoside (OTG), a 1% OTG cell lysis buffer was made instead of using Bugbuster Protein Extraction Reagent. The cell lysis buffer also included DNase, RNase I, and lysozyme at a final concentration of 3 μ g/ml for DNase and RNase I, and 30 μ g/ml for lysozyme. Cell lysis was then carried on as described above.

2.12. Protein Purification Under Denaturing Conditions on Ni-NTA Spin Columns

Small-scale protein purification under denaturing conditions was based on the protocol outlined in the Ni-NTA spin kit handbook provided by the supplier (QIAGEN). Cell pellets derived from 50 ml *E. coli* cell culture grown to OD₆₀₀ of 2.0 were resuspended in Buffer B (7M urea, 0.1 M NaH₂PO₄, 0.01 m Tris-Cl, pH 8.0) and incubated for 15 minutes at room temperature. The cellular debris was pelleted by centrifuging

the lysate at 12,000 x g for 30 minutes at room temperature. The supernatant was loaded on Ni-NTA spin columns equilibrated with Buffer B and centrifuged for 5 minutes at 270 x g. The Ni-NTA spin columns were washed twice with Buffer C (8 M urea, 0.1 M NaH₂PO₄, 0.01 M Tris-Cl, pH 6.3) and the protein was eluted twice with Buffer E (8 M urea, 0.1 M NaH₂PO₄, 0.01 M Tris-Cl, pH 4.5). The flow-through, washes and eluates were analyzed on a 12% SDS-PAGE.

2.13. Purification of H6-EntB, H6-EntD, and H6-EntE by IMAC Chromatography

Fresh affinity matrix was packed by use of a pump in a 1.2 cm x 10.0 cm column. The column was equilibrated with 3 column volumes of distilled water followed by 3 column volumes of Buffer A (50 mM Hepes (pH 8.0), 300 mM KCl, 0.5 mM Tris(2-carboxyethyl)phosphine (TCEP), 20 mM imidazole, (10 mM imidazole and 10% glycerol for H6-EntD)). Clear lysates containing soluble H6-EntB, H6-EntD, or H6-EntE, hereinafter called EntB, EntD, EntE, were loaded onto 10 ml Profinity Imac columns (Bio-Rad Laboratories) connected to a BioLogic DuoFlow FPLC system (Bio-Rad Laboratories). The columns were washed with 10 column volumes of buffer A. Fractions of the proteins that eluted from the linear imidazole gradient (20-250 mM) were collected. Cell lysate, flow-through, eluted fractions and pooled fractions were collected and analyzed by SDS-PAGE. H6-EntD and H6-EntE were purified to near homogeneity (>95% purity) and they were dialyzed against 50 mM Hepes (pH 8.0), 150 mM KCl, 0.5 mM TCEP, and 15% glycerol, and stored at -20 °C. After several uses the resin was cleaned with 3 column volumes of 1 M NaOH. The resin was then washed with distilled water until a pH 7-8 was

obtained. Three column volumes of 0.1 M NiSO₄/1 M Acetic acid (pH 4.0) was used to regenerate the resin. H6-EntB required an additional purification by anion exchange chromatography.

2.14. EntB Anion-Exchange

Purification of EntB on Profinity Imac was not sufficient to yield homogeneous protein. Therefore, anion-exchange was used as a second purification step to yield a higher purity protein. After Profinity Imac purification, EntB was dialyzed against Buffer A-anion (50 mM Hepes (pH 8.0), 100 mM KCl, and 0.5 mM TCEP) to lower the salt concentration and to remove imidazole. The dialysate was filtered with a 0.45 micron filter and loaded onto a 10 ml Uno-Q anion exchange column (Bio-Rad Laboratories) connected to a BioLogic DuoFlow FPLC system. The column was washed with 10 column volumes of Buffer A-anion and the protein was eluted with a linear KCl gradient (100-500 mM). Fractions containing eluted protein were analyzed by SDS-PAGE, pooled and dialyzed against 50 mM Hepes (pH 8.0), 150 mM KCl, 0.5 mM TCEP, and 15% glycerol. Aliquots of the protein were stored at -20 °C.

2.15. Measurement of Protein Concentration

During protein purification and gel analysis, protein concentrations were determined following the procedure accompanying the Bio-Rad protein assay kit. To 800 µl of protein sample and water, 200 µl of dye reagent concentrate was added to give a final volume of 1 ml. The reaction mixture was incubated for 5 minutes before measuring the absorbance at 595 nm on a Thermo Scientific GenesysTM 10 UV

spectrophotometer. A standard curve was produced by measuring absorbance of different amounts of bovine serum albumin (2, 4, 6, 8 μg) at 595 nm. The concentration of proteins was determined by linear regression analysis from the standard curve. For all other purposes, concentrations of the hexahistidine-tagged proteins were determined based on the calculated extinction coefficients at 280nm: EntB, $51910 \text{ M}^{-1} \text{ cm}^{-1}$; EntD, $30940 \text{ M}^{-1} \text{ cm}^{-1}$; EntE, $57300 \text{ M}^{-1} \text{ cm}^{-1}$; all extinction coefficients were predicted from the Expasy site (<http://au.expasy.org/tools/protparam>) using the ProtParam application.

2.16. Dialysis

Buffer exchange of the purified and partially purified proteins was accomplished by dialysis. Dialysis tubing with a molecular weight cut-off of 12,000-14,000 Daltons was cut into a desired length and soaked in dH_2O for 30 minutes. The dialysis tube was then soaked in the dialysis buffer. One side of the tubing was clipped and loaded with dialysis buffer to verify for leaks and further equilibrate the tube. The buffer was removed and replaced with the protein solution. The other end was clipped making sure to leave some space in the tube to allow for expansion. The tube was placed in a beaker with 1L of buffer and dialyzed for 4-16 hours at 4°C while slowly being stirred with a stir bar to increase diffusion rate. The dialysis buffer was changed for a total of three times.

2.17. Buffer Exchange on NAP-5 Columns

Small volumes of protein samples were buffer exchanged with Illustra™ NAP™-5 columns from GE healthcare according to the instruction manual. The NAP-5 column was equilibrated with 10 ml of elution buffer. Afterwards, 0.5 ml of protein sample was

added to the column and allowed to enter the gel bed completely. The purified protein sample was then eluted into a 1.5 ml microcentrifuge tube with 1 ml of elution buffer and the protein concentration was determined by measuring the absorbance at 280.

2.18. Synthesis of Alexa Fluor 488-CoA and BODIPY-TMR-CoA

Alexa Fluor 488 C₅-maleimide (1 mg) and BODIPY-TMR maleimide (1 mg), both from Invitrogen, were dissolved with 0.25 ml of DMSO or DMF, respectively. 0.5 mg (0.89 μ mol) of Alexa Fluor 488 C₅-maleimide in DMSO was mixed with 0.4 mg (0.61 μ mol) of CoA disodium salt in 100 mM ammonium phosphate pH 7.0. For BODIPY-TMR maleimide, 0.2 mg (0.36 μ mol) in DMF were mixed with 0.2 mg (0.24 μ mol) of CoA disodium salt in 50 mM Tris (pH 7.5) in 100 μ l of 9:1 DMF/50 mM Tris pH (7.5). The reactions were stirred for 3.5 hours at room temperature in the dark, after which 5 ml of water was added to the reaction. The diluted reactions were then washed for a total of 7 times with ethyl acetate to remove any unreacted fluorescent dye. The samples were stored in the dark at 4 °C.

2.19. Thin-Layer Chromatography

The purity of conjugated CoA and hence, the efficiency of the conjugation reaction, was assessed by using thin-layer chromatography. A silica gel based plate of 10 cm X 15 cm was used with a baseline starting from 1 cm from the edge. Approximately 8-12 μ l drops of the Alexa Fluor 488-CoA and Bodipy-TMR-CoA and standards (CoA, Bodipy-TMR or Alexa Fluor 488 C₅-maleimide, and wash 1) were spotted on the TLC plates. The TLC plates were then placed in a shallow pool of 60 % acetonitrile until the

mobile phase reached to a height of 10 cm. The plates were removed and air-dried. Afterwards, the plates were placed in an Ultraviolet Fluorescence Analysis Cabinet (Spectroline Model CL-150) and observed with a 365 nm UV lamp.

2.20. Biophysical Experiments

2.20.1. Far-UV Circular Dichroism

Far-UV circular dichroism studies were performed on a Jasco J-815 Spectropolarimeter to determine the overall secondary structure of the proteins in different conditions. The spectropolarimeter was connected to a Peltier heating and cooling unit. Protein samples of a concentration of 0.2-0.3 mg/ml (5-9 μ M for EntB monomer; 7-11 μ M for EntD) were prepared in PBS buffer (pH 7.4) with 10% glycerol. Spectra were collected from an average of 5 scans at a scanning rate of 20 nm/min, 0.2 nm step resolution, and 0.25 second response time, at a bandwidth of 1 mm under a constant nitrogen flow. The spectra were performed from 200-260 nm at 20 °C in a 1 or 2 mm pathlength quartz cuvette. The spectra were corrected for any contribution from the buffer and quartz cuvette. Buffers were scanned under the same conditions and the baseline was subtracted from the protein samples using the provided JASCO software tool.

2.20.2. Near-UV Circular Dichroism

Near-UV circular dichroism scans were performed on a Jasco J-710 spectropolarimeter to assess the tertiary structure of EntD in different conditions (i.e. apo-EntD, + 10 mM Mg^{2+} , + 100 μ M CoA, + 10 mM Mg^{2+} + 100 μ M CoA). The spectra

were collected in a 10 mm pathlength quartz cuvette with a protein concentration of 0.6 mg/ml (23 μ M). The scans were carried under the same conditions as the far-UV spectra except they were performed from 250-320 nm. The data were corrected as described in section 2.20.1.

2.20.3. Temperature Denaturation Monitored by Circular Dichroism

The thermal stability of the secondary structures of proteins under different conditions was determined by monitoring the loss of CD signal at 222 nm. The temperature denaturation curves were collected with the same parameters as in section 2.20.1. For EntD and its variants, thermal denaturation was carried from 20-60 °C at a rate of 20 °C/hour. For EntB and its variants, the thermal denaturation experiments were carried from 25-95 °C at a rate of 30 °C/hour. In order to determine the actual melting temperature, the first order derivative of the temperature denaturation curves was obtained with the JASCO analysis software provided with the JASCO J-815 spectropolarimeter.

2.20.4. Fluorescence Spectroscopy

Fluorescence spectroscopy scans were carried out to determine any structural changes for EntD wildtype and mutants. The scans were carried on a Varian Cary Eclipse Spectrofluorometer at a temperature of 20 °C. EntD and variants were diluted to a final concentration of 2 μ M in Hepes buffer (50 mM Hepes pH 8.0, 150 mM KCl, 0.5 mM TCEP and 10% glycerol) in the presence or absence of Mg^{2+} (10 mM) and/or CoA (100 μ M). For an excitation wavelength of 280 nm, an accumulation of 10 emission scans were

taken between 300-400 nm with a scan speed of 2400 nm/min, 700 volts, and excitation and emission slits of 5 nm. Same parameters were used for the excitation wavelength of 295 nm.

2.20.5. Fluorescence Anisotropy

Fluorescence anisotropy experiments were used to determine the binding affinity of BODIPY-TMR-CoA to EntD wild-type and mutants. Fluorescence anisotropy values were obtained directly from the Varian Cary Eclipse Spectrofluorometer at a temperature of 20 °C. 50 nM of Bodipy-TMR-CoA was incubated with increasing amounts of EntD WT or variants (0, 500, 1000, 1500, 2500, 5000, 10000, 15000 nM) in the fluorescence polarization buffer (50 mM Hepes pH 7.5, 10 mM MgCl₂ (+/-), 1.0 mM TCEP, 0.0025% Triton X-100, 1% DMSO, and 5% glycerol) for 5 minutes in the dark. Triplicate values were then used to determine the apparent dissociation constant by fitting the observed anisotropies to equation 1 by nonlinear regression analysis where y denotes the fluorescence anisotropy value and B_{max} is the maximal number of binding sites.

$$y = (B_{max} + x) / (K_d + x) \quad (1)$$

2.20.6. Analytical Ultracentrifugation

Conformational changes of EntD WT upon binding to Mg²⁺ and CoA was observed using sedimentation velocity analytical ultracentrifugation (AUC). Experiments were carried out on a Beckman Coulter XL-1 AUC with a four-place An-60Ti rotor. Samples

(390 μ l) containing 10 μ M of EntD in AUC buffer (Hepes (pH 8.0), 150 mM KCl, and 0.5 mM TCEP, 10 mM Mg^{2+} , 100 μ M CoA) and the reference (400 μ l) of AUC buffer minus protein were loaded into the double sector centerpiece separately. Each AUC run consisted of 250 absorbance scans that were recorded at 20 °C with a rotor speed of 40,000 rpm. Protein samples were monitored by UV absorbance at 280 nm in a continuous mode with a time interval of 10 minutes. Scans at different time points were fitted to a continuous size distribution model by the program SEDFIT (Schuck, 2000) in order to determine sedimentation coefficients and frictional ratios.

2.21. Electrospray Ionization Mass Spectrometry

2.21.1. Determination of Protein Mass

Protein mass determination of EntB and its variants was performed by a Micromass Q-ToF 2 triple-quadrupole mass spectrometer. Protein sample preparation was as reported by Weinglass et al. (2003). Approximately 100 μ l of protein sample corresponding to 40 μ g was diluted with 3 volumes of methanol and briefly vortexed. Then, 1 volume of chloroform was added and vortexed briefly. Precipitation was achieved by adding 2 volumes of MilliQ water and vortexing briefly. The sample was then centrifuged at 12,000 xg for 5 minutes. The precipitated protein was present at the interface of the two phases. The two phases were removed without disturbing the precipitated protein. The sample was washed twice by resuspending with 300 μ l of methanol and centrifuging for 5 minutes at 12000 xg. Protected from light the protein sample was then air-dried at room temperature. Afterwards, the protein sample was

dissolved in 30:70:0.1 ACN:MilliQ water:TFA (v/v/v) to a final concentration of 6 μ M. The sample was applied directly at a flow rate of 1 μ l/min. Samples were analyzed in the positive ion mode. The signal was deconvoluted by using the Micromass MassLynx v4.0 software. The instrument was calibrated using [Glu]-fibrinopeptide B as a standard.

2.21.2. Activity Assay

Activity of EntD, and EntE was determined by observing changes in the masses of *apo*-EntB and *holo*-EntB upon pantetheinylation and acylation, respectively. *holo*-EntB was prepared with 6 μ M of *apo*-EntB incubating with the EntD reaction mixture consisting of 50 mM Hepes (pH 8.0), 150 mM KCl, 0.5 mM TCEP, 10 mM MgCl₂, 200 μ M CoA, and varying concentrations of EntD (50, 100, and 200 nM) in a volume of 100 μ l for 60 minutes at 37 °C. The reaction was then stopped by methanol/chloroform precipitation as described in the previous section (2.21.1) and analyzed by ESI-MS. *acyl-holo*-EntB was prepared by reaction of 6 μ M of *apo*-EntB with first 50 mM Hepes (pH 8.0), 150 mM KCl, 0.5 mM TCEP, 10 mM MgCl₂, 400 μ M CoA, and 100 nM EntD in a volume of 50 μ l for 60 minutes at 37 °C. An EntE reaction mixture of a volume of 50 μ l consisting of 50 mM Hepes (pH 8.0), 150 mM KCl, 0.5 mM TCEP, 10 mM MgCl₂, 200 μ M CoA, 10 mM ATP, 1 mM 2,3-DHB, and varying concentrations of EntE (200, 400, 800 nM) were then added to give a final volume of 100 μ l. The reaction was incubated for 20 minutes at 37 °C and then stopped by methanol/chloroform protein precipitation. The sample was then analyzed by ESI-MS.

2.21.3. Time-Course Assay

The initial velocity for the rate of conversion of *apo*-EntB to *holo*-EntB catalyzed by EntD was determined by a time-course assay analyzed by ESI-MS. A reaction mixture containing 6 μ M of *apo*-EntB was incubated with the EntD reaction mixture consisting of 50 nM of EntD (wild-type or H44A variant) as described above (Section 2.21.2). An aliquot of the reaction was taken at each time interval (0, 1, 3, 5, 10, 15, 30, 60 min) and stopped as described previously. The duplicate values were then analyzed by ESI-MS. The percent *holo*-EntB formed was calculated by the peak intensities determined by the MassLynx software and using equation 2:

$$\% \text{ } \textit{holo}\text{-EntB Formed} = \frac{\textit{holo}\text{-EntB}}{\textit{apo}\text{-EntB} + \textit{holo}\text{-EntB}} \quad (2)$$

2.22. Gel-based Activity Assay

The activity of EntD was also determined by SDS-PAGE. A reaction containing 6 μ M of *apo*-EntB was mixed with EntD reaction mixture consisting of 50 mM Hepes (pH 8.0), 150 mM KCl, 1 mM TCEP, 10 mM MgCl₂, 10 μ M Alexa-CoA or BODIPY-TMR-CoA, and 50 nM of EntD (wild-type and variants). The reaction mixture was incubated for 60 minutes at 37°C and then stopped by loading 20 μ l of the sample (4.8 μ M of EntB) into 5X SDS-loading buffer. The samples (4.8 μ M EntB) were then loaded on a 10% SDS-PAGE and run for 75 minutes until the dye reached 2 mm from the bottom. The gel was then removed and placed under a 365 nm UV light to observe the fluorescently labeled *holo*-EntB. The proteins were then visualized with Blue-Silver staining.

2.23. Gel-based Time Course Assay

The initial rate for the conversion of *apo*-EntB to *holo*-EntB was also determined by a gel-based time course assay. A reaction containing 6 μM of *apo*-EntB was incubated with the EntD reaction mixture containing 10 μM of Alexa-CoA or BODIPY-TMR-CoA. A 20 μl aliquot of the reaction was taken at each time interval (0, 1, 3, 5, 10, 15, 30, and 60 minutes) and placed into a 5X SDS-loading buffer. After running the samples on a 10% SDS-PAGE, the fluorescently labeled *holo*-EntB bands were visualized using the Fluorchem FC2 imaging instrument. The fraction of *holo*-EntB formed was calculated by using the band analysis tool on the Fluorchem FC2 software. An identical area was selected around each band in order to determine band intensities. The band intensity at 60 minutes was assumed to be 100 % *holo*-EntB. The fraction of *holo*-EntB formed was then calculated by using equation 2. The initial rate was determined from the slope of reaction trace over the first 15 minutes.

3. Results

3.1. Characterization of EntD Cofactor-Binding Residues

The first part of this project involved characterization of EntD and variants upon cofactor binding and determining their activity. The conformational change of EntD upon CoA and Mg^{2+} binding was studied using various biophysical techniques. The conversion of *apo*-EntB to *holo*-EntB by the catalytic activity of EntD and variants was determined by ESI-MS and gel based assay.

3.1.1. Homology Model

In order to have a better understanding of the nature of cofactor (CoA, Mg^{2+}) binding to *E. coli* EntD, a homology-based model of EntD was generated by using the crystal structure of the *Bacillus subtilis* surfactin biosynthetic enzyme Sfp (PDB code 1QR0) as template. The structure-based alignment of the two primary amino acids sequences obtained from ClustalW2 (www.ebi.ac.uk/Tools/clustalw2) was used as input for Modeller 9v7 (Figure 10). The two proteins have a 15% sequence identity. The generated three-dimensional EntD model was aligned with Sfp by LSQMAN with an RMSD of 0.81 Å as determined by PDBeFold (www.ebi.ac.uk/msd-srv/ssm/). The model and the corresponding Sfp crystal structure were viewed by using PyMol (DeLano Scientific LLC) to determine structural similarities and differences between the two proteins (Figure 11).

```

EntD/1-209   1  MVDKMTTHTSLPFAGHTLHFVVEFDPANFCBQ.....DLWLPHYAQLQHAG.RKRKTSHFACRTAAVYAL
Sfp/1-228   1  .....MKIYGIYMDRPLSQBNERFMTFLSPEKREKCRRFYHKEDAHRTLLCDVLRVSI

EntD/1-209   65  REY...GYKCVFAIGELRQVWVPA..EYVGSISHCGTTALA VVSROPICIDIEEIPSVQTARSLTDNII
Sfp/1-228   56  SRQYQLDKSDIRFSTQBYGRPCIPDLPAHFNISSGRWVIGAFDSOPICIDIEKTKPI..SLIAKRFF

EntD/1-209   129  TPAEHERLADCGLAFSLALTLAFAKESAFKASEIQTDAGFLDYQIISWNNKQVVIHRENEMFAVH..W
Sfp/1-228   124  SKTEYSDLAKDKDEQTDYFYHLWSMKESFIKQEGKGLSLPLDSFSVRLHQDQVSTELPDSEHSPCYIKT

EntD/1-209   196  QIKEKIVITLCQHD.....
Sfp/1-228   194  YEVDPGYKMAVCAAHPDFPEDITMVSYEELLRAAA

```

Figure 10 Sequence alignment of EntD with Sfp performed using ClustalW2.
The two proteins have a 15% sequence identity. Highly conserved residues are shaded in red and residues with similar properties are colored red.

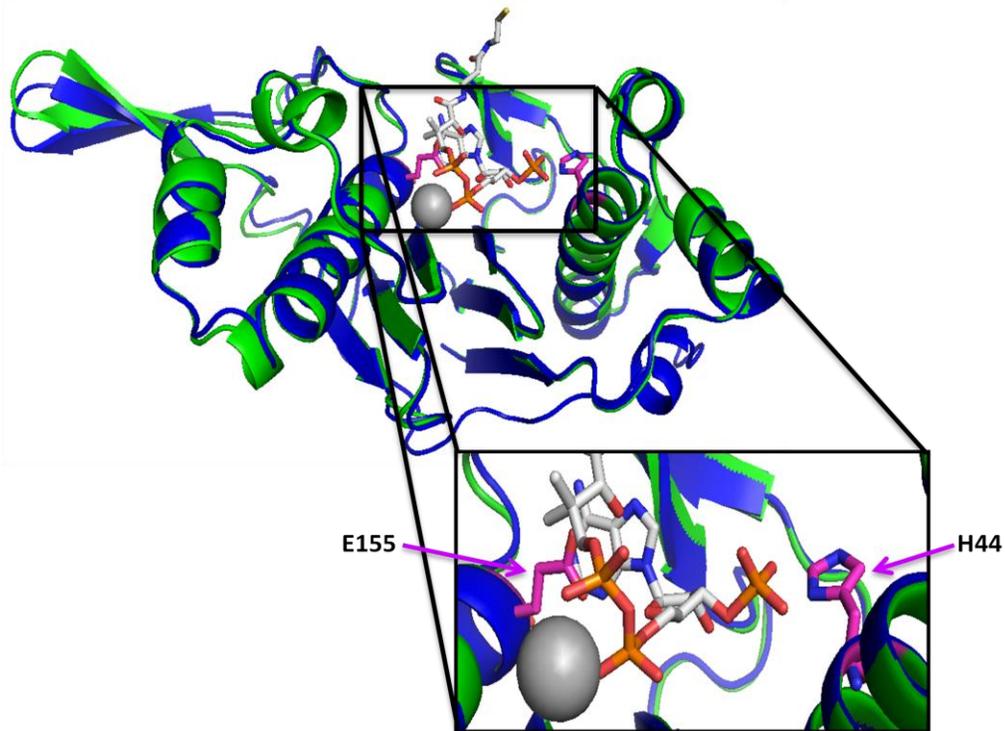


Figure 11 Homology-based model of *E. coli* EntD.

Structural superposition of EntD (green) homology model to *B. subtilis* Sfp (blue: PDB code 1QR0) in the presence of CoA (white stick: atoms colored according to elements) and Mg^{2+} (grey sphere). Possible cofactor residues are highlighted in magenta.

3.1.2. Site-Directed Mutagenesis

E. coli K-12 AG-1 cells (*recA1 endA1 gyrA96 thi-1 hsdR17(r_K⁻ m_K⁺) supE44 relA1*) harboring pCA24N plasmids containing the genes encoding EntD, EntB, and EntE with in-frame N-terminal hexahistidine tags were obtained from the ASKA repository (Kitagawa *et al.*, 2005). The plasmid adds seven spacer amino acids at the N-terminal end, and five spacer amino acids at the C-terminal end. Additionally, the C-terminal end may include an in-frame fusion with GFP (i.e. pCA24N-EntB-GFP, and pCA24N-EntE-GFP) (Figure 12). The ORF clones pCA24N can be removed by *Sfi*I restriction sites at both the N- and C-

terminal boundaries of the ORF. Cutting with SfiI allows for easy transfer of the ORF to other vector systems. Expression of the ORF clone is directed by the IPTG-inducible promoter P_{T5-lac} which is suppressed by $lacI^q$. The pCA24N plasmids also carry the chloramphenicol resistance gene (*cat*). This high-copy number plasmid can be used for many applications such as: construction of DNA microarray (Rodolfo *et al.*, 2008), production and purification of proteins (Khalil *et al.*, 2009; Campbell *et al.*, 2008), analysis of protein localization by monitoring GFP fluorescence (Venkatesh *et al.*, 2005), and analysis of protein-protein interaction (Khalil *et al.*, 2009). According to the EntD homology model, two residues were identified to be potential ligand binding residues (Figure 11). The first residue identified was His44 where the δ -nitrogen of the imidazole group salt bridges to the 3'phosphate of CoA. This residue was mutated to an alanine to disrupt any interaction of CoA with His44 (H44A). The second residue identified was the highly conserved Glu155 which, according to the homology model, interacts with the adenine base of CoA. This residue was also changed to alanine (E155A). The introduction of the mutations was confirmed by DNA sequencing of the resulting plasmids at the McGill University and Genome Quebec Innovation Center.

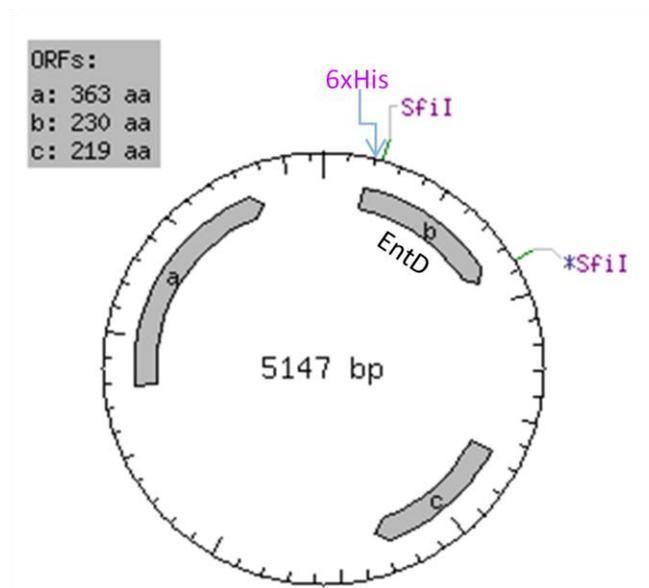


Figure 12 pCA24N plasmid with *entD* gene (b).

The 3 lettered arrows indicate the size of the ORFs. (a) lac repressor $lacI^q$ (b) phosphopantetheinyl transferase EntD, and (c) chloramphenicol acetyltransferase (CAT). SfiI restriction sites flank the ORF of EntD.

3.1.3. Protein Purification Under Denaturing Conditions

pCA24N-EntD expression conditions were first monitored using Ni-NTA spin columns. AG-1 cells harboring pCA24N-EntD plasmids were grown under three different induction temperatures (18, 25, and 37 °C) and IPTG concentrations (0, 0.25, 0.50, 1.0 mM). Purification of the crude extract was attempted using the Qiagen Ni-NTA spin columns as described in section 2.12. Nitrilotriacetic acid (NTA) occupies four of the six available ligand binding sites on the nickel ion. This leaves two sites available to interact with the 6xHis tag. The spin columns were washed for a total of three times with a low imidazole concentration (20 mM). High imidazole concentration (300 mM) competes for the two binding sites on the nickel ion occupied by the histidine residues, thereby

eluting the protein. However, the protein did not bind tightly to the Ni-NTA resin and therefore eluted with the flow-through and the wash steps. Figure 13 shows the amount of EntD expressed under the different conditions. The presence of a strong protein band at 0 mM IPTG indicates that EntD expression is leaky. However, large scale protein expression at 0 mM IPTG failed to produce any significant amount of EntD. At 37 °C, the protein expression is lower than at 25 °C. In the end, protein expression was achieved at 25 °C overnight with 0.5 mM IPTG since it produced the highest amount of EntD.

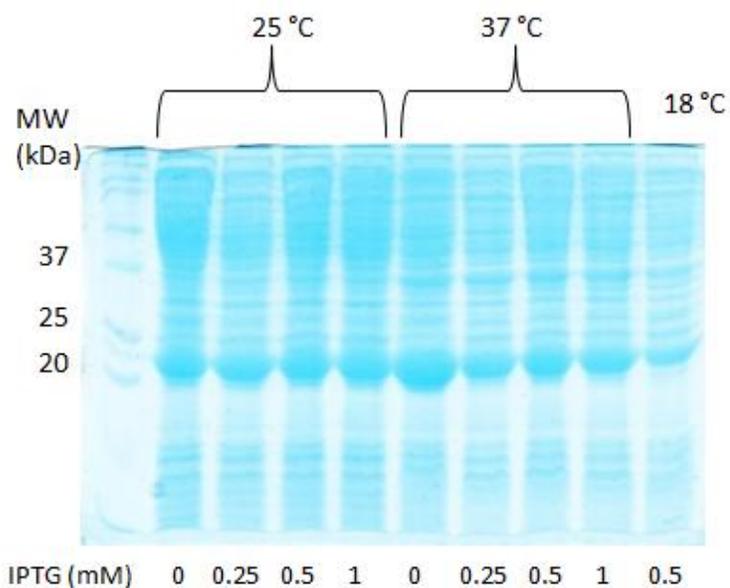


Figure 13 12% SDS-PAGE of EntD protein expression under different conditions. 2 μ l of the crude extract were loaded in each lane. Proteins were visualized using Blue Silver Stain.

3.1.4. Large-Scale Protein Purification of EntD and Variants

EntD and its variants were purified from 1L cultures on a 10 ml Profinity Imac column connected to a Bio-Rad Biologic Duoflow system. Unlike NTA, iminodiacetic acid (IDA) resin binds to three of the binding sites on the nickel ion leaving another three

sites free to bind to the hexahistidine tag. The crude extract was loaded onto the column and then washed with a 10 mM imidazole to elute any non-specific binding proteins. Higher imidazole concentrations in the wash buffer resulted in the early elution of EntD. EntD and the variants were eluted with a linear imidazole gradient from 10-250mM. Figure 14 shows the elution profile for EntD WT, where the left y-axis represents the absorbance at 280 nm (blue) and the right represents the conductivity in mS/cm (red). The proteins of interest eluted at approximately 30% of the gradient. The elution profiles of WT EntD and the two variants (H44A and E55A) were similar. The fractions with the highest absorbance were pooled (fractions 11-15) and dialyzed against the dialysis buffer. Initially, a major contaminant was observed at around 70 kDa. This was later determined to be GroEL by LC-MS/MS (McGill Mass Spectrometry Core Facility) (Figure 15). Growing the cells at 25 °C overnight decreased the amount of GroEL contamination (Figure 16). Approximately 15, 7, and 10 mg of EntD WT, H44A, and E155A respectively, were obtained from 1 L culture.

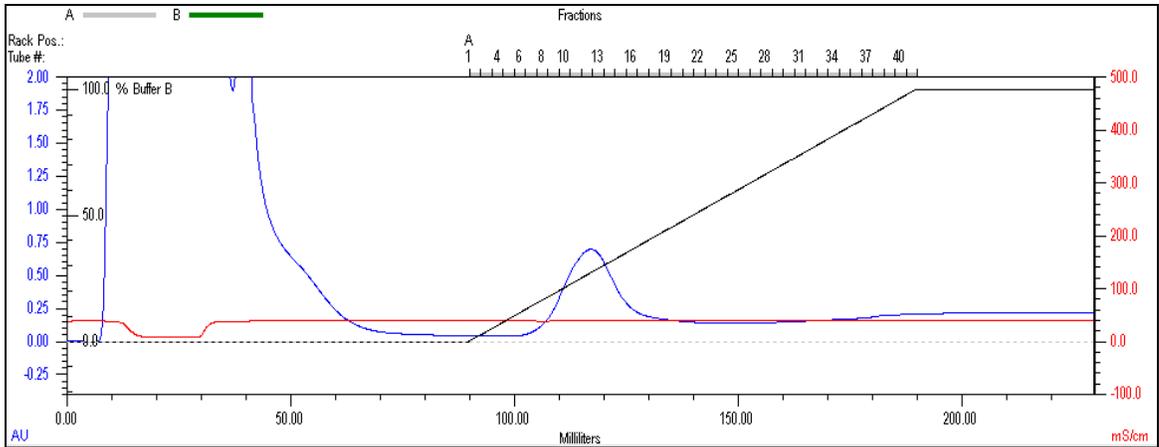


Figure 14 Elution profile of wild-type EntD from the Imac Column.

EntD eluted at approximately 30% of the linear imidazole gradient 10-250 mM imidazole (black slope). Elution profile monitors the absorbance at 280 nm (blue) and the conductivity (red).

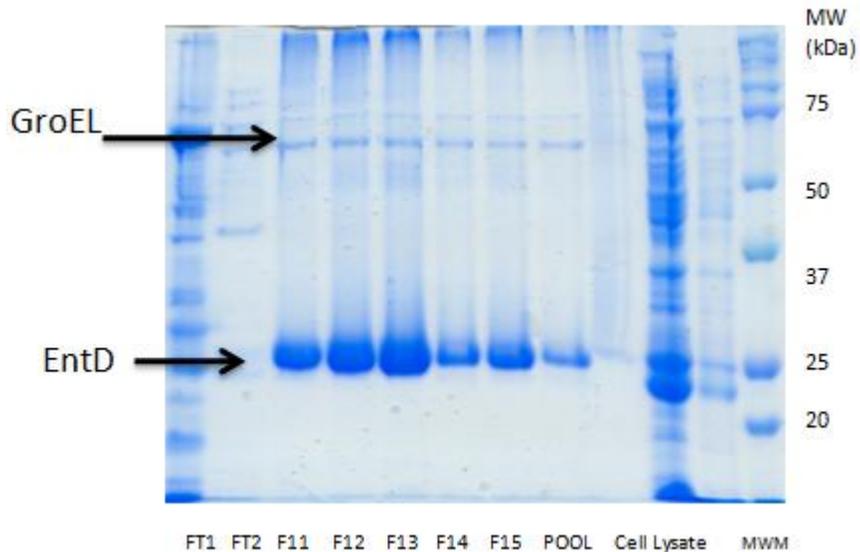


Figure 15 SDS-PAGE of purification fractions of EntD wild-type.

1 L cell culture was harvested after being induced for 4 hours at 37 °C. 20 µl of sample was loaded in each lane. Fractions 11-15 were pooled. GroEL contamination (70 kDa) was confirmed by LC-MS/MS. Samples were visualized using Blue Silver Stain.

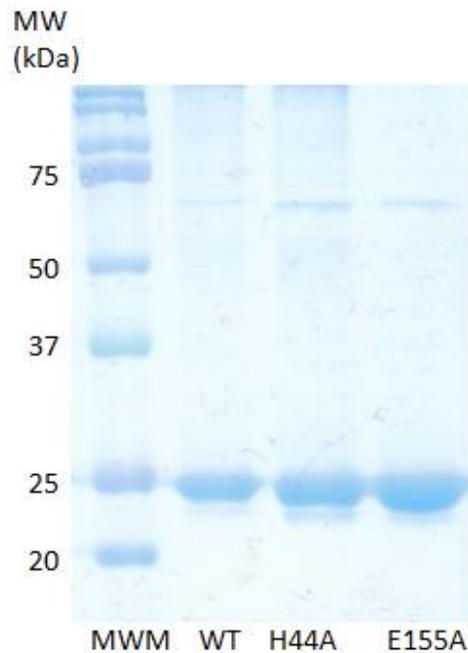


Figure 16 SDS-PAGE of EntD wild-type, H44A, and E155A.

Cells were harvested after being induced overnight at 25 °C. Approximately 10 µg of each protein sample was loaded on each lane. The GroEL contamination decreased due to the change in growth conditions. Proteins were visualized using Blue Silver stain.

3.1.5. Secondary Structure of EntD and Variants Assessed by Far-UV Circular Dichroism

Far-UV CD is a useful technique in determining the protein secondary structure. Different protein peptide bonds of secondary structures such as α -helices, β -sheets, and random coils exhibit differential absorbance of right versus left circularly polarized light (Johnson, 1999). The secondary structures of EntD wild-type, H44A, E155A were compared by far-UV circular dichroism (200-260 nm). Figure 17 demonstrates that EntD wild-type and variants have a double minima at 208 nm and 222 nm, typical of protein with a high α -helical content. A lower 222 nm signal indicates the presence of β -sheets.

Therefore, the double minimum with a stronger 208 nm signal indicates that we have a mixture of α -helices and β -sheets. Wild-type EntD, and H44A variant have overlapping spectra indicating an identical secondary structure (Figure 17). E155A exhibits a similar spectra but with a slightly stronger signal. This variation could be due to a difference in protein concentration. Table 2 tabulates the secondary structure percentages determined by the K2D server (www.embl.de/~andrade/k2d). Globular secondary structure of all three proteins are nearly identical thus demonstrating that the secondary structure of the variants is unaffected by the mutation. The secondary structure estimation by K2D was confirmed by predicting the secondary structure from the EntD wild-type primary sequence using the PSIPRED Protein Structure Prediction Server (bioinf.cs.ucl.ac.uk/psipred/) (Figure 18). PSIPRED predicted an α -helical content of 27%, and a β -strand content of 21%, thereby agreeing with secondary content estimations by K2D (Table 2, row 5). The homology-based model of wild-type EntD exhibits secondary structure content of 38, 28 and 34% for α -helices, β -strands, and random coil, respectively, which is in moderate agreement with the K2D estimations. However, it should be noted that the EntD model does not include the additional 21 amino acids added by the plasmid design.

Protein	α -helices (%)	β -strands (%)	Random coil (%)
EntD WT	31	17	52
EntD H44A	30	19	51
EntD E155A	30	15	55
PSIPRED (WT)	27	21	52
Model (WT)	38	28	34

Table 2 Secondary structure percentages determined by K2D server, PSIPRED and homology model.

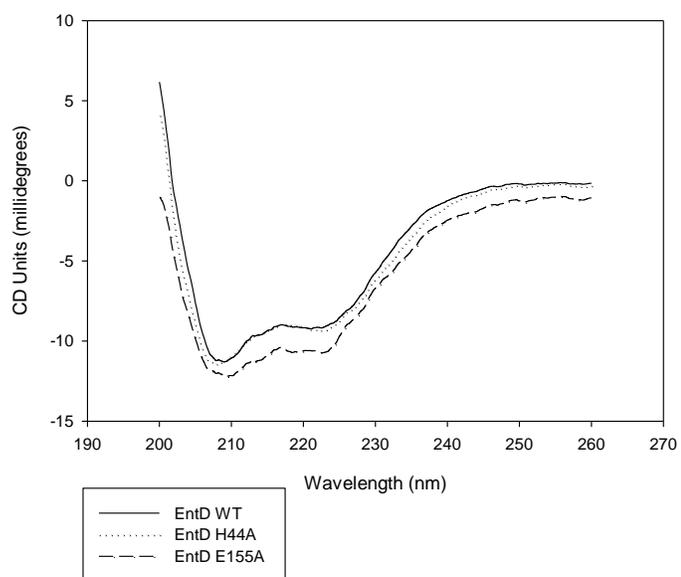


Figure 17 Overlay of the far-UV circular dichroism spectra of EntD, H44A, and E155A. CD spectra were obtained from 300 μ l of sample at a concentration of 0.2 mg/ml.

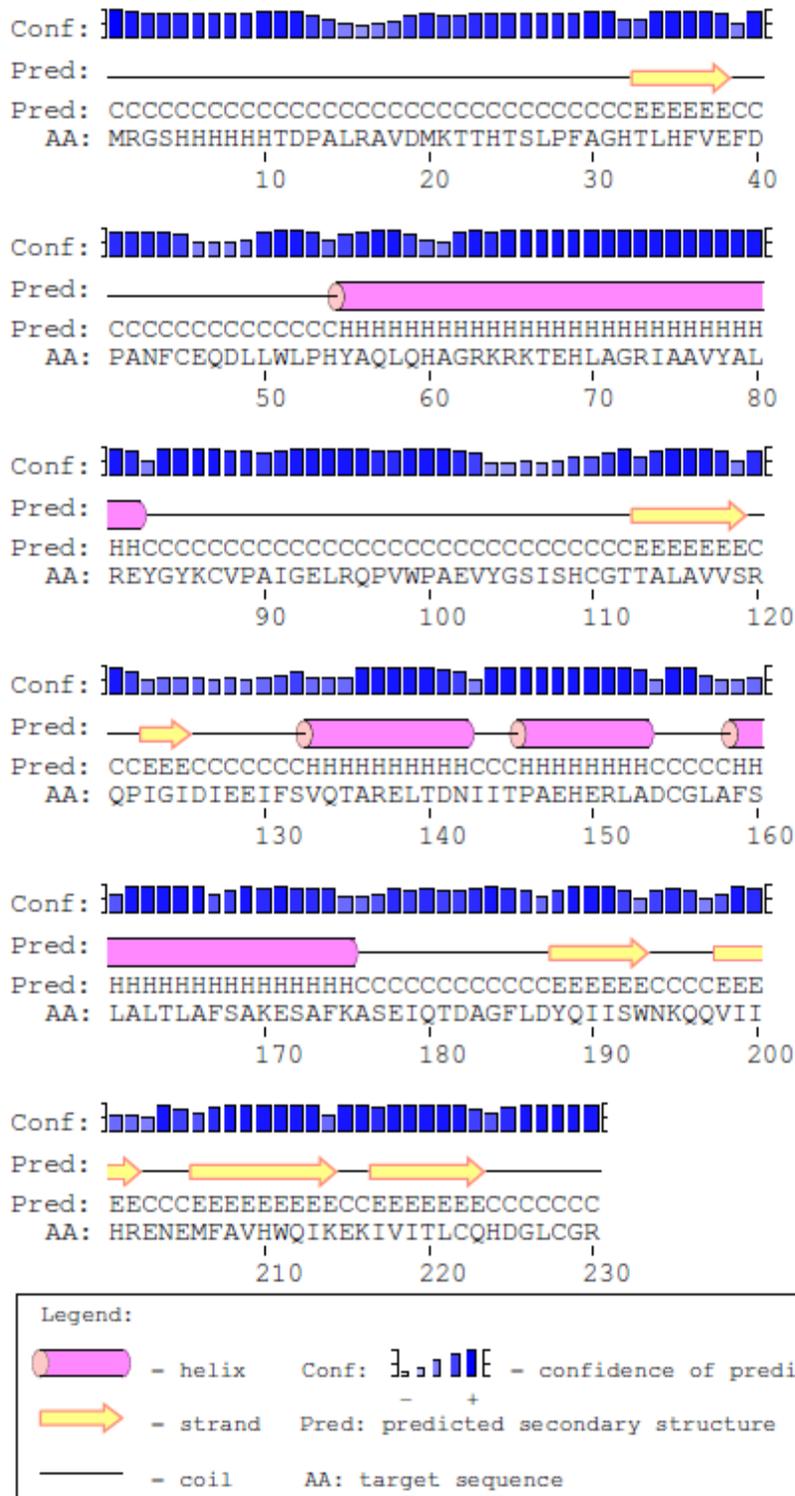


Figure 18 EntD E155A secondary structure prediction by PSIPRED.

The primary amino acid sequence of 6xHIS-tagged EntD wild-type was submitted to the PSIPRED server.

3.1.6. Tertiary Structure of EntD and Variants Assessed by Fluorescence Spectroscopy

EntD has 4 tryptophan, 6 tyrosine, and 10 phenylalanine residues which are aromatic side chains that contribute to fluorescence. An excitation wavelength of 280 nm excites tyrosine and tryptophan residues whereas an excitation wavelength of 295 nm mainly excites tryptophan residues. Fluorescence spectra of EntD and variants excited at 280 and 295 nm were compared to assess their tertiary structures. Emission scans were recorded between 300-400 nm. At an excitation wavelength of 280 nm, the two variants, H44A and E155A, have similar emission spectra with slight blue shifts in emission maxima compared to WT. The emission spectrum of EntD wild-type is slightly less intense than the variants (Figure 19A). The emission spectra of EntD and the variants when excited at 295 nm were slightly different from one another (Figure 19B). Figure 19b demonstrates that H44A had the lowest maximum peak.

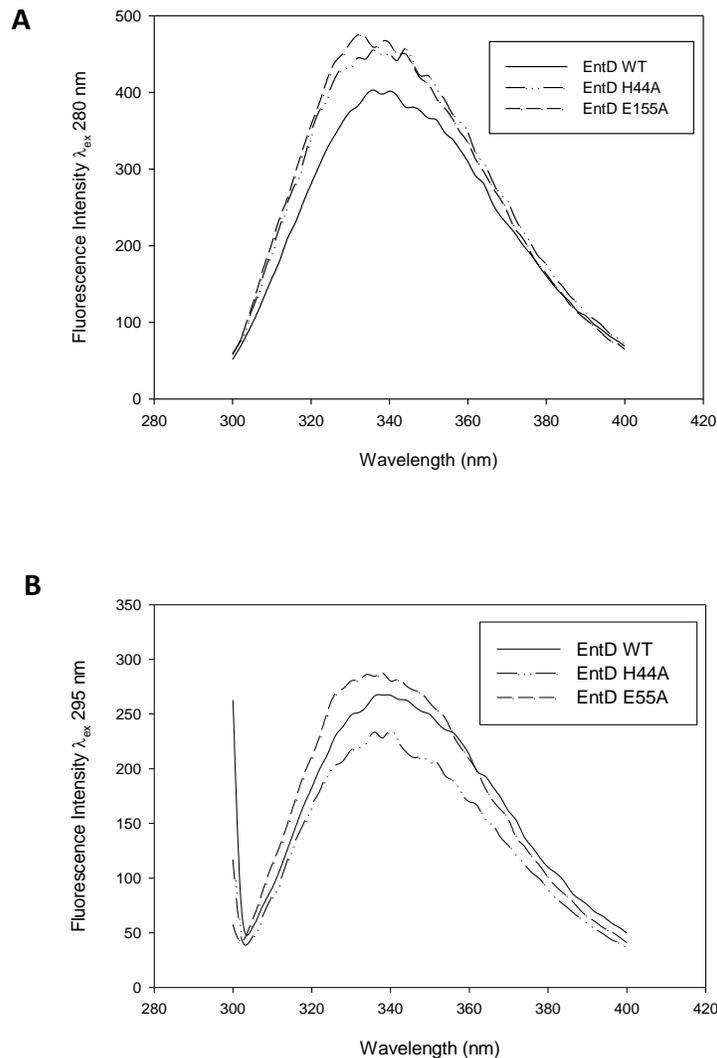


Figure 19 Fluorescence emission scans of EntD WT and variants.

Fluorescence spectra of 0.3 mg/ml protein samples excited at 280 nm (A) and 295 nm (B) were collected from 300-400 nm.

3.1.7. Secondary Structure Change upon Cofactor Binding Determined by Far-UV CD

The effect of cofactor (Mg^{2+} , CoA) on the secondary structure of EntD WT and variants was analyzed using far-UV CD. 10 mM of Mg^{2+} and 100 μ M of CoA were added to the protein samples and the far-UV CD spectra were monitored in the range of 200-

260 nm. Figure 20 demonstrates that the presence of Mg^{2+} and CoA does not alter the secondary structure of wild-type EntD. Similarly, the global secondary structure of the variants is not significantly altered in the presence of the cofactors (Figure 21A-B).

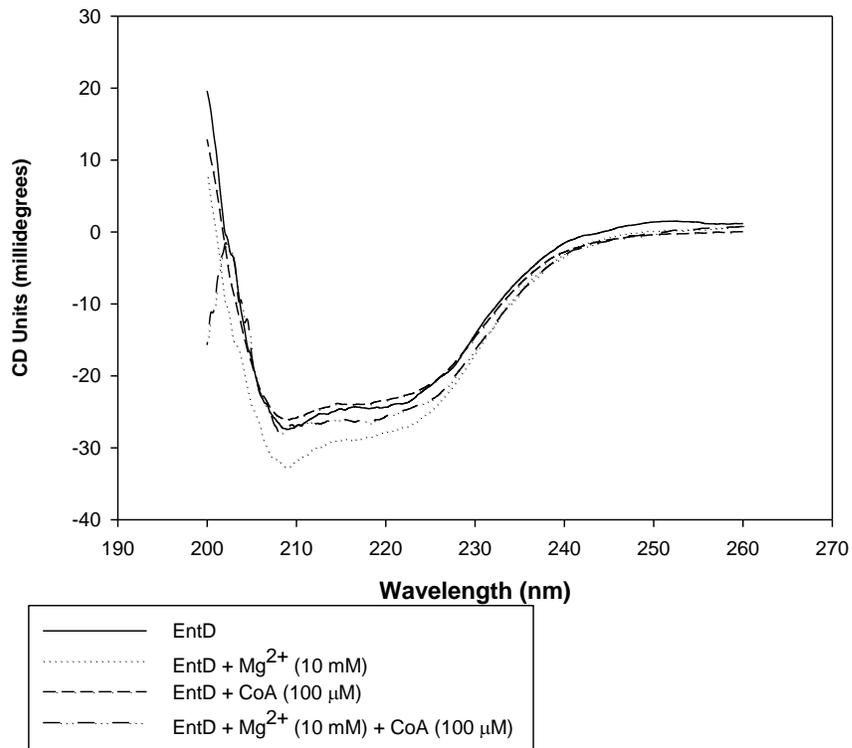


Figure 20 Far-UV circular dichroism of EntD WT in the presence of cofactors. Secondary structure of EntD monitored from 200-260nm in the presence of Mg^{2+} and CoA.

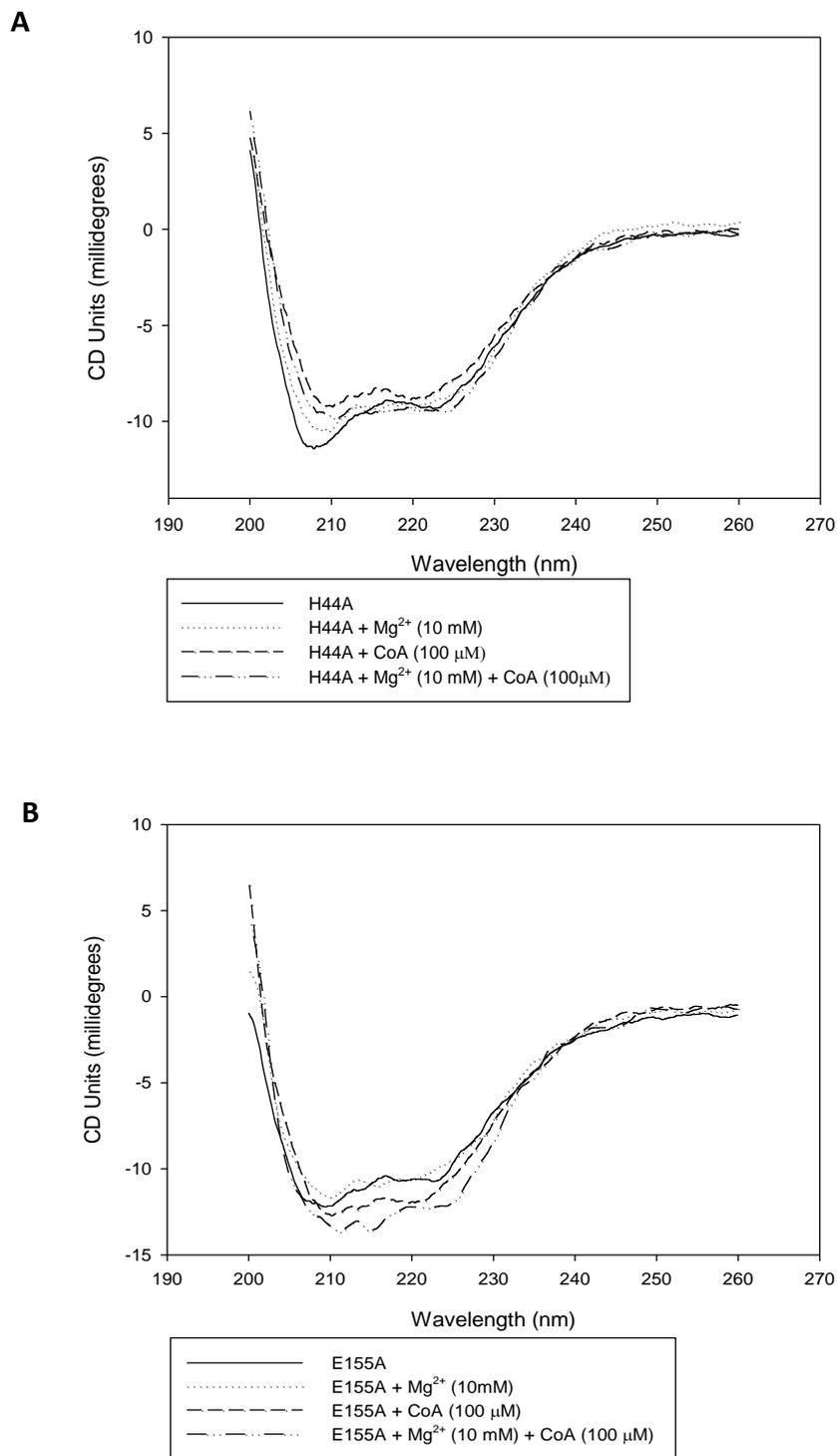


Figure 21 Far-UV CD spectra of EntD variants in the presence of Mg²⁺ and CoA. The presence of the cofactors did not significantly affect the secondary structure of H44A (A) and E155A (B).

3.1.8. Thermal Stability of EntD and Variants in the Presence of Cofactors Monitored by Circular Dichroism

The effect of cofactor binding on the stability of EntD and variants was also determined by monitoring the thermal stability by CD. All three proteins were subjected to thermal denaturation from 20 °C to 60 °C at a rate of 20 °C/hour. Loss of α -helical content was monitored at 222 nm. Figure 22 shows sigmoidal shaped curves for EntD wild-type indicating a cooperative unfolding in the absence and presence of cofactors. Similar melting temperature curves were obtained for both variants (data not shown). The stability of EntD and variants was increased in the presence of both cofactors. In the absence of cofactors, the melting temperature of EntD WT, H44A, and E155A were 37.8, 35.4, and 38.2 °C respectively. The melting temperature did not significantly change upon addition of one of the cofactors. However, in the presence of both cofactors, the melting temperature of all three enzymes increased to approximately 44, 45.4, and 44.6 °C, respectively. This indicates that the protein undergoes a conformational change upon ligand binding thus, increasing its thermal stability.

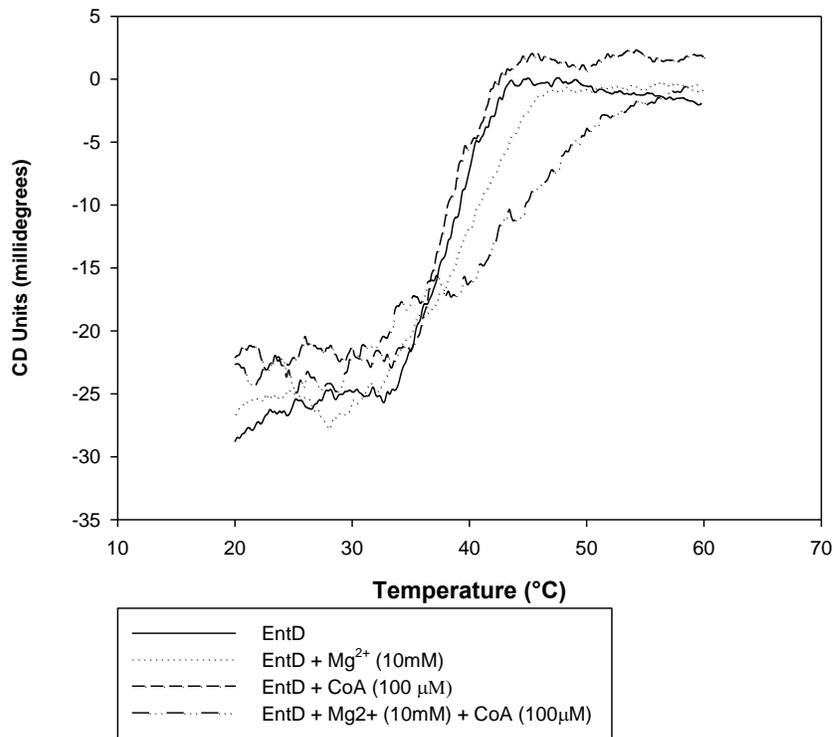


Figure 22 Thermal denaturation curves for EntD WT in the absence and presence of cofactors.

The thermal stability of EntD increased in the presence of both cofactors.

Sample	EntD WT (Tm °C)	H44A (Tm °C)	E155A (Tm °C)
<i>apo</i>	37.8	35.4	38.2
+ Mg ²⁺ (10 mM)	39.9	39.0	41.6
+ CoA (100 μM)	36.5	39.0	40.2
+ Mg ²⁺ (10 mM) + CoA (100 μM)	44	45.4	44.6

Table 3 Melting temperature of EntD and variants determined by circular dichroism

3.1.9. Cofactor Binding Assessed by Near-UV Circular Dichroism

The tertiary structure of wild-type EntD and variants in the presence of the cofactors was assessed by near-UV CD. Similar to fluorescence spectroscopy, near-UV CD monitors the environment of the aromatic amino acids as well as cysteine residues. Near-UV CD spectra of EntD and variants were monitored from a range of 250-320 nm. Phenylalanine residues have a high degree of symmetry and therefore are weakly absorbing from 250-270 nm with the strongest bands expected between 267-269 nm and 260-262 nm. The hydroxyl group of tyrosine greatly reduces the symmetry of the ring. As a result, its near-UV CD signal is eight times more intense than phenylalanine. The two major bands of tyrosine are expected at 277 and 283 nm. The two bands of tryptophan residues are between 288-293 nm. The intensity of the signal of all three EntD proteins, slightly increases upon addition of Mg^{2+} or CoA (Figure 23A-C). However, a more intense signal for all three enzymes is observed in the presence of both cofactors. E155A exhibits a lower signal change between 250-290 nm compared to WT and H44A. The more intense signal signifies that the aromatic residues are being held tightly and losing mobility upon cofactor binding. The increase in signal is observed from 250-290 nm corresponding to phenylalanine, tyrosine, and tryptophan residues. The signal increase could also arise from aromatic residues approaching another aromatic side chain.

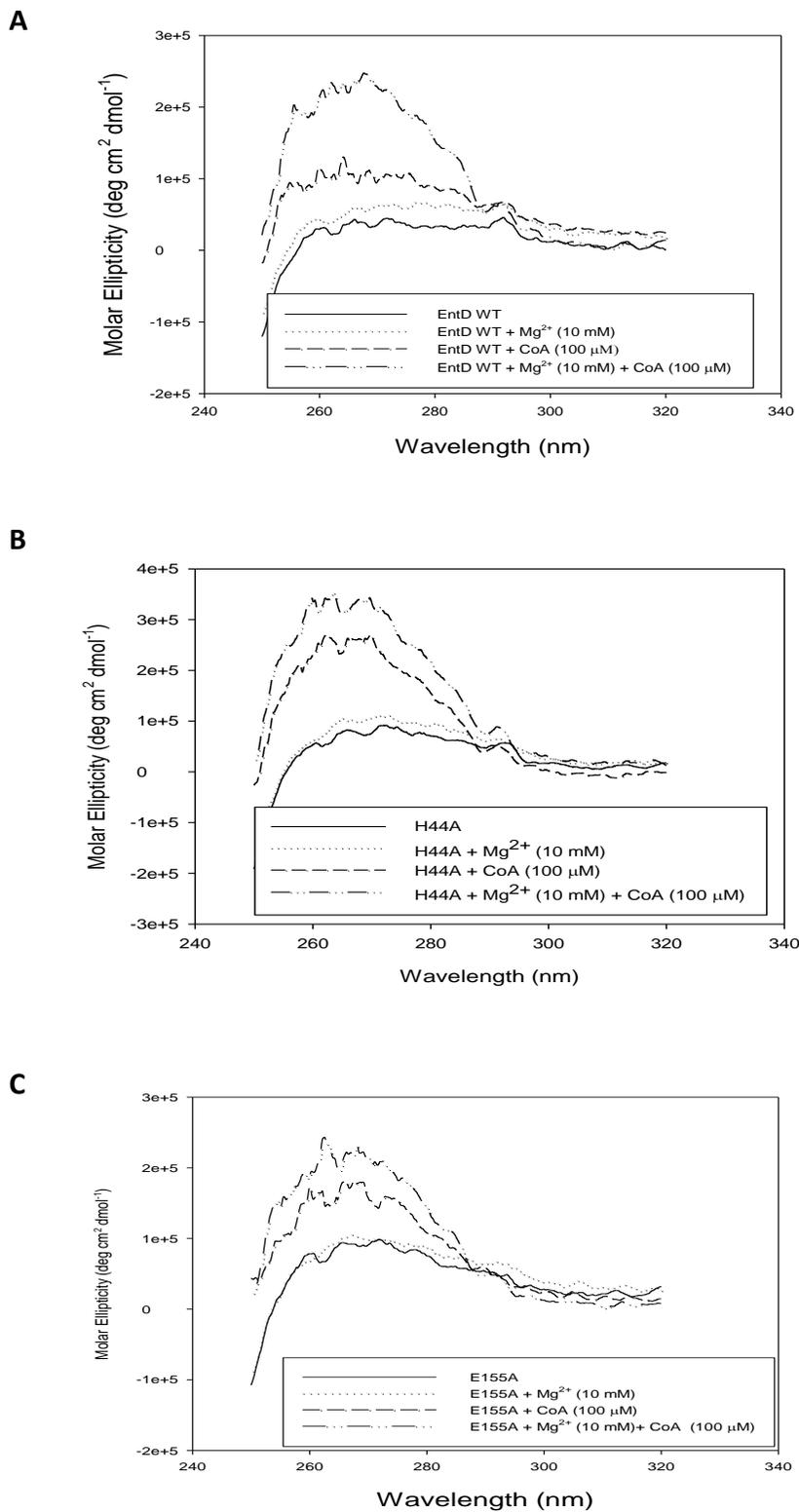


Figure 23 Near-UV CD spectra of EntD and variants in the presence of cofactors. Increase in near-UV CD signal upon addition of CoA and Mg²⁺ confirms conformational change of EntD WT (A), H44A(B), and E155A (C).

3.1.10. Sedimentation Velocity Analytical Ultracentrifugation

Sedimentation velocity experiments on wild-type EntD were carried in the presence of Mg^{2+} (1 mM) and CoA (100 μ M). The Mg^{2+} concentration had to be reduced from 10 mM to 1 mM. EntD precipitated in the presence of 10 mM Mg^{2+} in the absence of glycerol. The lower cation concentration enabled the protein to stay in solution. The removal of glycerol resulted in the precipitation of both EntD variants, H44A and E155A. Therefore, only wild-type EntD in the presence of the cofactors was examined by sedimentation velocity AUC (Figure 24). Moreover, removal of glycerol from wild-type EntD also prevented us from replicating this experiment due to precipitation of the protein. Table 4 lists the sedimentation coefficient and frictional ratio values of EntD in the presence of both cofactors. In the presence of both cofactors the sedimentation coefficient is increased from 2.32 in the absence of both cofactors to 2.46. Minor increases in the sedimentation coefficient were observed in the presence of either cofactor. The frictional ratio of EntD, which is a measure of the shape of the protein, also decreases in the presence of both Mg^{2+} and CoA going from 1.68 to 1.37. The presence of one cofactor has a minor effect on the frictional ratio. Therefore, binding of both cofactors changes the conformation of the protein. The increase in sedimentation coefficient and decrease in frictional ratio indicate that the protein is going from an elongated to a more globular shape.

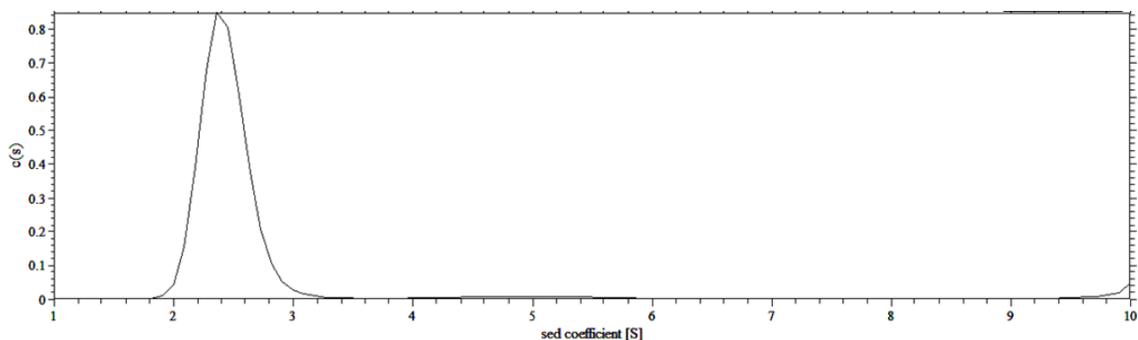


Figure 24 Sedimentation Velocity AUC of wild-type EntD in the presence of cofactors. The sedimentation coefficient of EntD was determined to be 2.46 with a frictional ratio of 1.37.

Protein	$S_{20,w}$	f/f_0
EntD	2.32	1.68
EntD + Mg^{2+}	2.34	1.61
EntD + CoA	2.37	1.58
EntD + Mg^{2+} + CoA	2.46	1.37

Table 4 Sedimentation coefficient and frictional ratio of wild-type EntD in the presence of cofactors as determined by SEDFIT analysis

3.1.11. Bioconjugation of Alexa Fluor 488 and BODIPY-TMR with CoA

Two different dyes, Alexa Fluor 488 C₅-maleimide (Alexa) and BODIPY-TMR-maleimide (BODIPY-TMR), were each conjugated to CoA. The reaction mixture, containing an excess of dye, was diluted with water after 3.5 hours of incubation in the dark at room temperature. Unreacted BODIPY-TMR was removed by ethyl acetate extraction. The first two washes removed the majority of the unreacted dye. Five

additional washes only removed small amounts of unreacted dye (Figure 25). Ethyl acetate washing failed to remove unconjugated Alexa dye since it was more soluble in a polar solution.

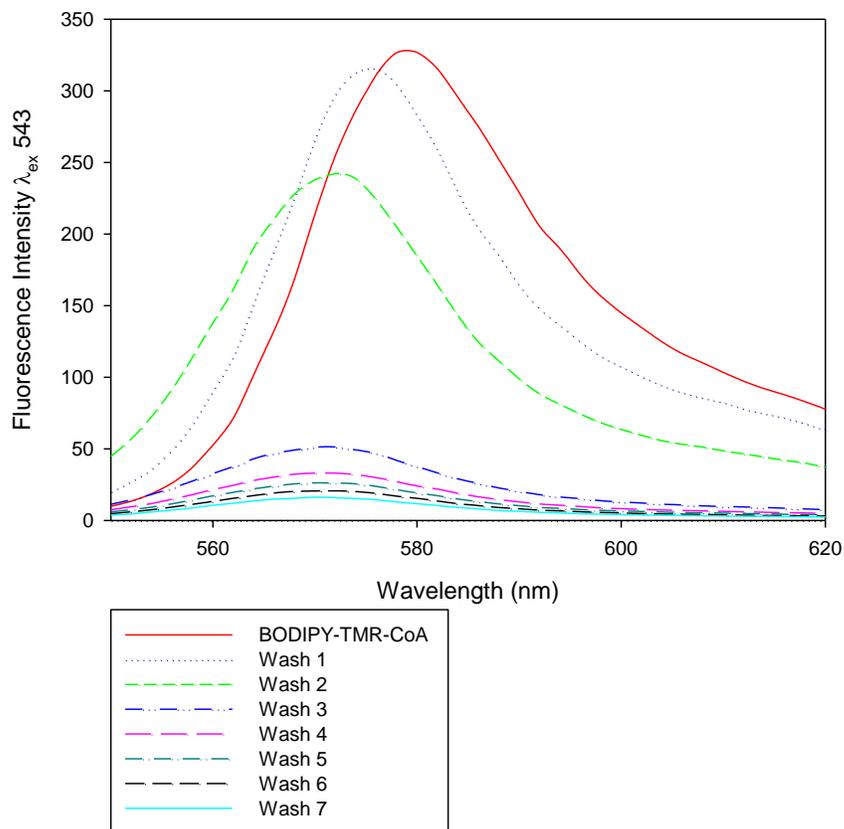


Figure 25 Fluorescence emission spectra of BODIPY-TMR-CoA and 7 ethyl acetate washes.

The fluorescence emission spectra of Bodipy-TMR-CoA (3 μ M) after washing was compared with the wash samples to evaluate removal of unreacted dye from the reaction. All samples were excited at 543 nm.

After 3.5 hours of reaction, the percent yield of BODIPY-TMR-CoA was determined to be 56%. Therefore, the reaction mixture contains 44% of unconjugated CoA. The amount of

CoA and BODIPY-TMR were calculated based on the extinction coefficients determined by their standard curve. The calculated extinction coefficients for CoA at 262 nm and BODIPY-TMR at 543 nm were 11391 and 46527 M⁻¹ cm⁻¹, respectively. The percent yield of Alexa-CoA was not determined since unconjugated dye still remained in the mixture after dilution. Figure 26 and Figure 27 demonstrate that thin-layer chromatography (TLC) can be used to assess the purity of a product. CoA is highly polar and thus will not move on the polar stationary phase plate. Alexa and BODIPY-TMR are both non-polar and therefore will move higher on the plate with the 60% acetonitrile mobile phase. Figure 26D demonstrates equal intensity for unconjugated Alexa and Alexa-CoA was observed in the product lane. Therefore, it is assumed that approximately 50% of Alexa dye is conjugated to CoA. Figure 27D demonstrates that the reaction mixture for BODIPY-TMR-CoA contains very small amounts of unconjugated dye not removed by the ethyl acetate wash. The amount of unconjugated BODIPY-TMR is minimal since the intensity of the BODIPY-TMR-CoA band is larger and more intense than unconjugated BODIPY-TMR (Figure 27).

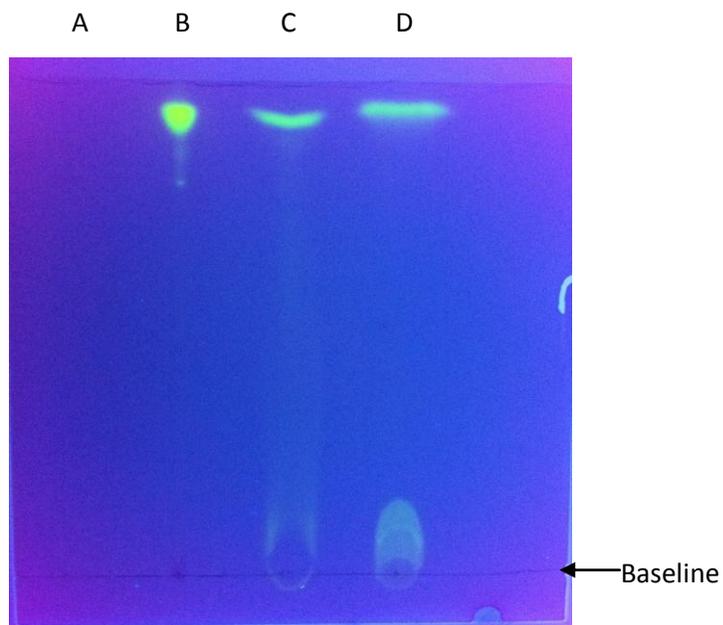


Figure 26 Thin-layer chromatography for conjugation reaction of Alexa Fluor 488 with CoA.

2 μ l of CoA (A), 0.5 μ l of Alexa (B), 6 and 12 μ l of reaction mixture (C and D, respectively) were visualized on the TLC plate with a 365 nm UV light.

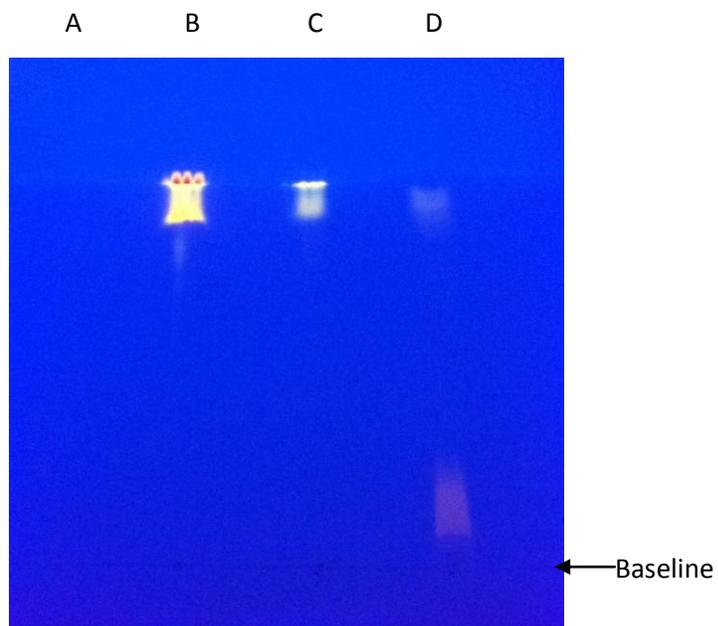


Figure 27 Thin-layer chromatography of conjugation reaction of BODIPY-TMR with CoA.

2 μ l of CoA (A), 0.5 μ l of BODIPY-TMR (B), 12 μ l of wash 1 (C), and 8 μ l reaction mixture (D) were visualized on the TLC plate with a 365 nm UV light.

3.1.12. Determination of BODIPY-TMR-CoA K_d by Fluorescence Anisotropy

Change in the steady-state fluorescence anisotropy of BODIPY-TMR-CoA was monitored with increasing concentrations of all wild-type and variant EntD proteins in the presence or absence of Mg^{2+} . The apparent binding affinity of Alexa-CoA to the proteins was not determined as the fluorescence anisotropy signals were too low due to a high presence of unreacted dye. The fluorescence anisotropy signal was expected to increase upon titration of the protein due to lower rotational mobility of the bound dye. Saturation was achieved at high protein concentrations (Figure 28 A,B). In the absence of Mg^{2+} , an apparent K_d of 1.33, 1.47, and 3.50 μM was calculated for wild-type EntD, H44A, and E155A, respectively (Figure 28A; Table 5, column 2). The apparent K_d remained in the same order of magnitude in the presence of Mg^{2+} . The apparent K_d for wild-type EntD, H44A, and E155A was calculated to be 2.34, 2.86, and 1.11, respectively (Figure 28B; Table 5, column 3). The data indicate that all three proteins are able to bind CoA with similar affinity in the presence or absence of Mg^{2+} .

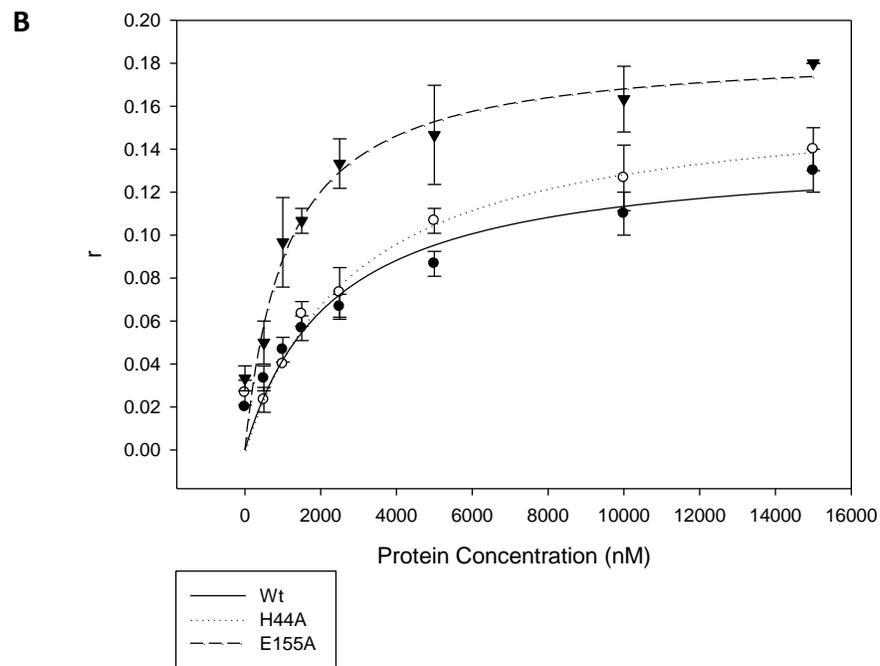
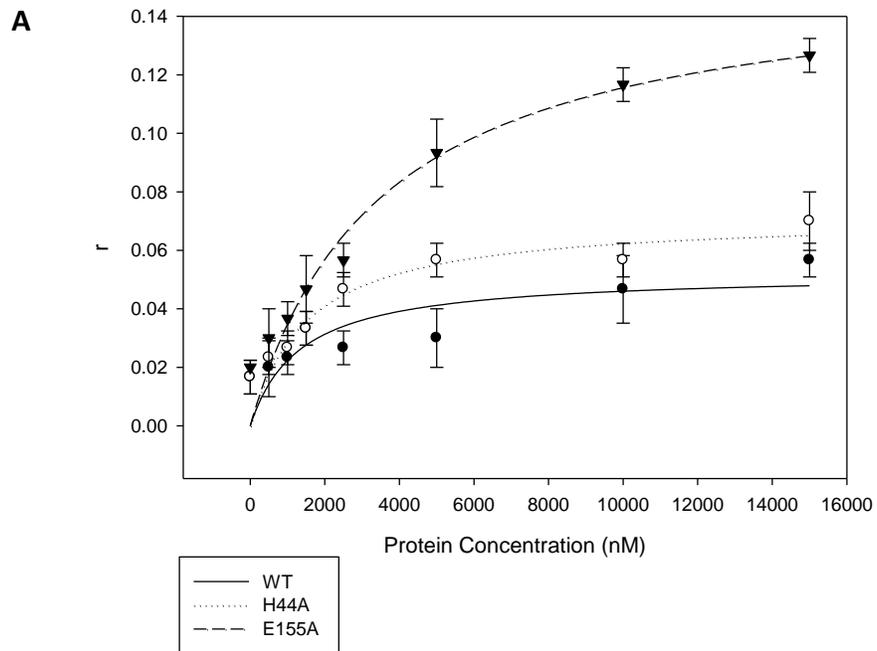


Figure 28 Fluorescence anisotropy of BODIPY-TMR-CoA with EntD proteins. 50 nM of BODIPY-TMR-CoA was titrated with EntD proteins in the absence (A) and presence of Mg^{2+} (10 mM) (B). Data shown are an average of three separate reactions.

EntD Protein	apparent K_d (-Mg ²⁺)	apparent K_d (+Mg ²⁺)
WT	1.33	2.34
H44A	1.47	2.86
E155A	3.50	1.11

Table 5 Apparent K_d of BODIPY-TMR-CoA with EntD proteins determined by fluorescence anisotropy.

3.1.13. Determination of Phosphopantetheinyl Transferase Activity by ESI-MS

The enzymatic activity of wild-type and variant EntD proteins were determined by electrospray ionization mass spectrometry (ESI-MS). EntD catalyzes the covalent attachment of the PPant group of CoA onto the conserved ser245 residue on *apo*-EntB. This covalent modification increases the mass by 339 Da. This mass increase can be detected by ESI-MS. The ProtParam tool (ca.expasy.org/tools/protparam) was used to calculate the theoretical mass of *apo*-EntB. A mass of 34888.9 Da was predicted when including the hexahistidine tag, the seven spacer amino acids at the N-terminal end, and the five amino acids at the C-terminal end added by the plasmid. The conversion to *holo*-EntB would increase the mass to 35227.9 Da. Figure 29 and Figure 30 demonstrate that wild-type EntD in the presence of CoA is capable of converting *apo*-EntB to *holo*-EntB. Table 6 summarizes the activity of wild-type EntD and variants. H44A was also able to pantetheinylate EntB. However, E155A was inactive since it failed to convert *apo*-EntB to *holo*-EntB. The addition of 1 mM EDTA or the omission of CoA prevented the formation of *holo*-EntB by wild-type EntD, therefore proving that this a PPTase specific reaction.

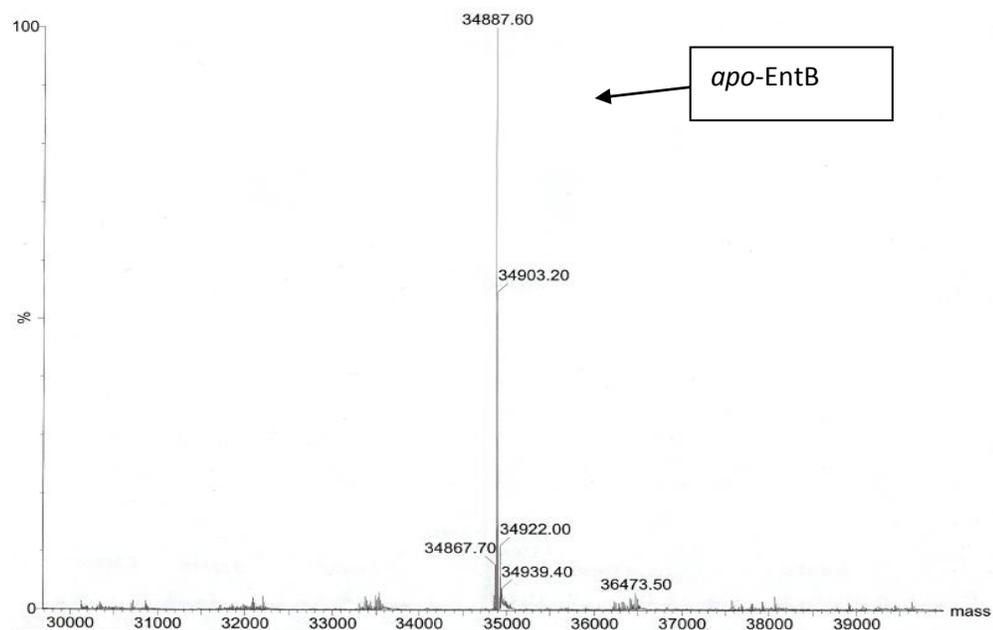


Figure 29 Mass spectrum of *apo*-EntB.

Signal from ESI-Q-TOF mass spectrometry instrument deconvoluted by MassLynx software determines a molecular weight of 34887.6 Da for *apo*-EntB.

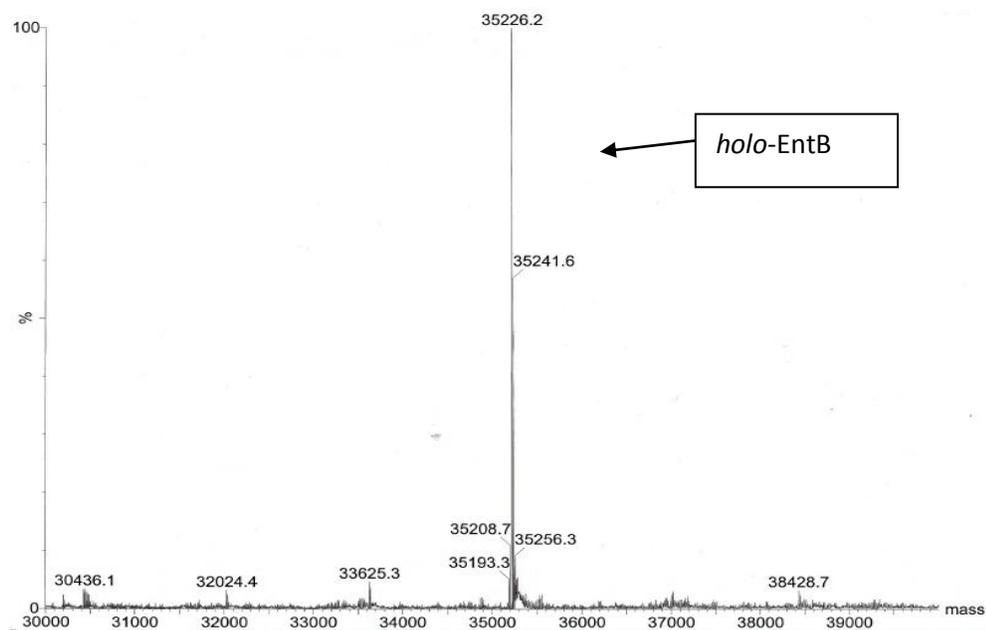


Figure 30 Mass spectrum of *holo*-EntB.

6 μM of *apo*-EntB was reacted with 100 nM of EntD, 200 μM CoA, and 10 mM Mg^{2+} . An increase of 338.6 Da corresponds to the covalent attachment of PPant group from CoA.

EntD	EntB	Expected Mass (Da)	Observed Mass (Da)	Δ Observed Mass (Da)
WT	<i>apo</i> -EntB	34888.9	34887.6	338.6
	<i>holo</i> -EntB	35227.9	35226.2	
H44A	<i>apo</i> -EntB	34888.9	34903.5	340
	<i>holo</i> -EntB	35227.9	35243.5	
E155A	<i>apo</i> -EntB	34888.9	34903.5	0.5
	<i>holo</i> -EntB	35227.9	34903.0	

Table 6 Summary of end-point activity for wild-type EntD and variants

3.1.14. Determination of EntE Activity on EntB WT by ESI-MS

The activity of EntE was also determined by ESI-MS. EntE activates 2,3-DHB and attaches it to the newly introduced PPant group of *holo*-EntB. The attachment of DHB increases the mass of *holo*-EntB by 136 Da. Table 7 demonstrates that EntE was active as we observed a mass increase of *holo*-EntB from 35226.2 Da to 35363.4 Da. This change in mass corresponds to 137.2 Da which corresponds to the mass of 2,3-DHB.

Protein	Expected Mass (Da)	Observed Mass (Da)	Δ Observed Mass (Da)
<i>holo</i> -EntB	35227.9	35226.2	137.2
<i>acyl-holo</i> -EntB	35363.9	35363.4	

Table 7 Summary of EntE activity on *holo*-EntB determined by ESI-MS.

3.1.15. Time-Course Assay with ESI-MS

Time-course assays of wild-type EntD and H44A were also performed using ESI-MS. A time-course assay with E155A was not completed since it failed to convert *apo*-EntB

to *holo*-EntB after 1 hour of incubation. The percent abundance of *holo*-EntB formed at each time point was calculated by the following equation:

$$\% \text{ BPI of holo-EntB} / (\% \text{ BPI of holo-EntB} + \% \text{ BPI of apo-EntB}) \times 100$$

where BPI is the base peak intensity determined from the deconvoluted spectra. Duplicate values show that wild-type EntD and H44A have a linear initial rate until complete conversion of *apo*-EntB occurs (Figure 31). The slope of the first 15 minutes or 10 minutes was used to determine the initial velocity of wild-type EntD or H44A, respectively. Wild-type EntD had an initial velocity of 0.66 $\mu\text{M}/\text{min}$ and H44A had a higher rate of 1.03 $\mu\text{M}/\text{min}$.

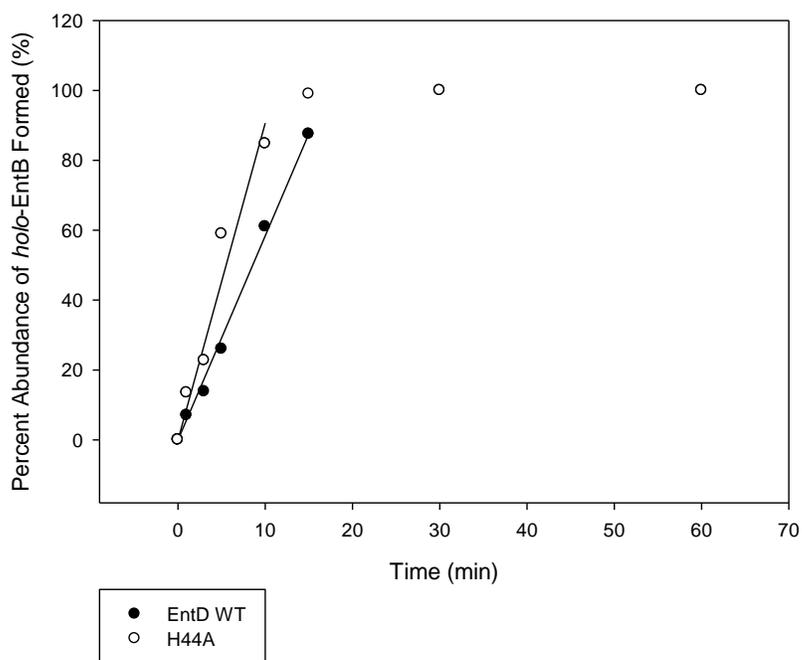


Figure 31 Time-course assay of EntD WT, and H44A determined by ESI-MS.

Both wild-type EntD (dark circles) and H44A (white circles) have a linear initial rate for the conversion of *apo*-EntB to *holo*-EntB. Data shown are an average of two separate reactions. The initial velocity was determined from the slope of the first 15 minutes.

3.1.16. Determination of Activity by SDS-PAGE Using Fluorescent CoA Conjugates

The activity of EntD and variants was also determined by transferring the fluorescently labeled CoA to EntB, thus converting *apo*-EntB into fluorescently labeled *holo*-EntB. The fluorescently labeled *holo*-EntB was then run on a 10% SDS-PAGE and visualized with a 365 nm UV light (Figure 32A). The fluorescent bands on the gel correspond to *holo*-EntB fluorescently labeled with either BODIPY-TMR (red) or Alexa Fluor 488 (green). Wild-type EntD and H44A were able to fluorescently label EntB whereas E155A had no activity. The major fluorescent bands at 37 kDa correspond to full-length fluorescently labeled *holo*-EntB. The fluorescent band at 30 kDa consists of

fragmented *holo*-EntB with a portion of the ArCP domain cleaved. However, Ser245 is still intact as evidenced by the fluorescent label. In contrast, the band at 25 kDa, consisting of the ICL domain of EntB, is not fluorescently labeled due to absence of Ser245. The protein bands were visualized by blue silver stain showing an equal amount of protein in each lane (Figure 32B).

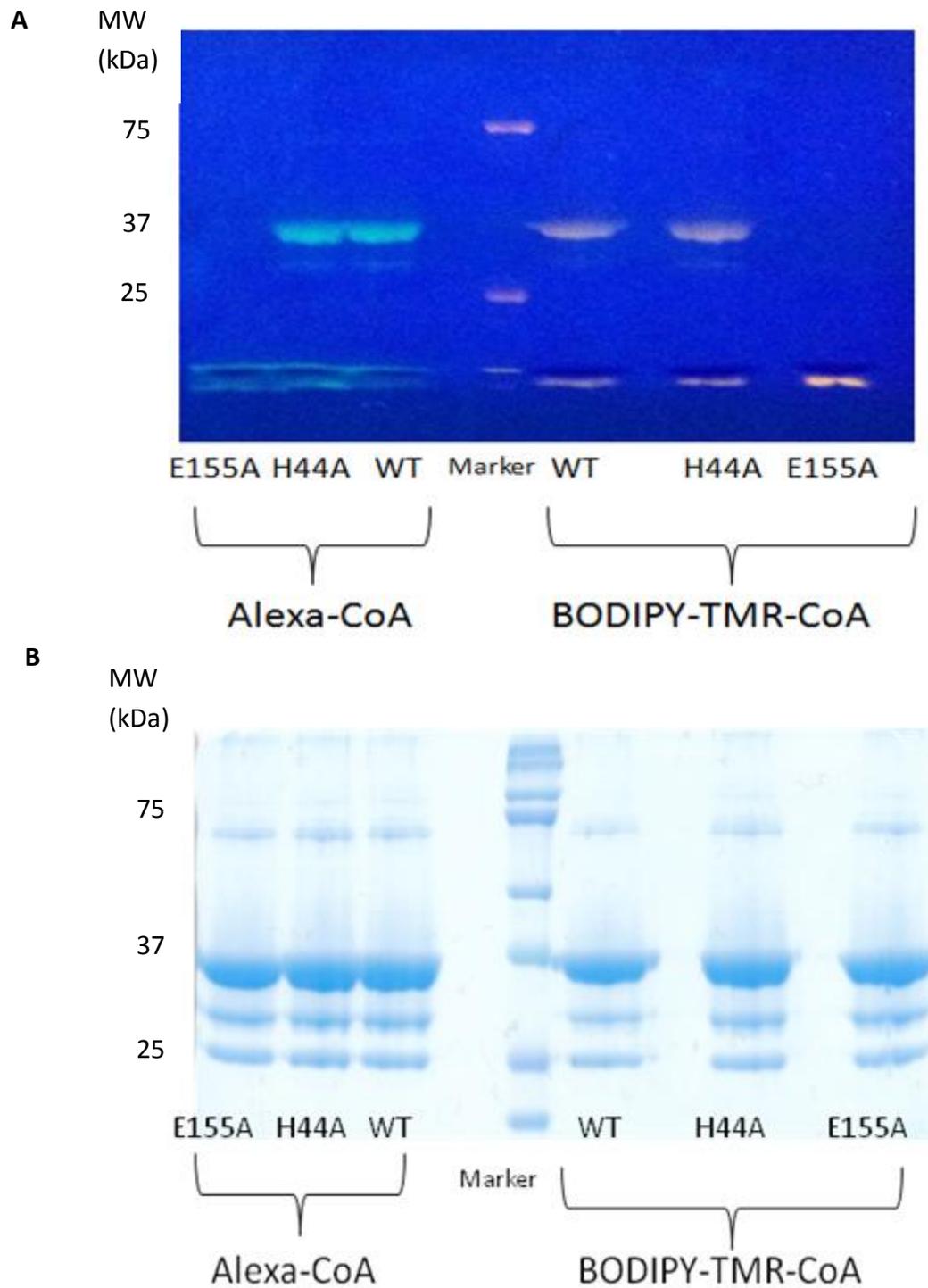


Figure 32 10% SDS-PAGE demonstrating incorporation of fluorescent into EntB substrates.

(A) SDS-PAGE visualized with 365 nm UV light. (B) Coomassie-stained gel of this reaction. Wild-type EntD and H44A were able to transfer the BODIPY-TMR-PPant (red) and Alexa-PPant (green) onto EntB. *apo*-EntB incubated with E155A failed to become fluorescently labeled.

3.1.17. Gel-based Time-Course Assay of EntD and Variants with EntB

Time course assay for EntD WT and H44A was also determined by gel-based assay. *apo*-EntB was reacted with EntD or H44A with either BODIPY-TMR-CoA or Alexa-CoA over a time of 60 minutes. At each time interval, an aliquot of the reaction corresponding to 6.5 µg of total EntB was taken and placed in 5X SDS-LB to stop the reaction. After running the gel, the fluorescently labeled *holo*-EntB bands were visualized with a UV transilluminator (Figure 33A). Figure 33A demonstrates an increase in the fluorescent band over time. Figure 33B demonstrates that each lane contained the same amount of protein at different time points. The quantity of fluorescently labeled *holo*-EntB at each time point was determined by densitometry. It was assumed that the signal intensity at 60 minutes corresponds to 100% conversion of all *apo*-EntB to *holo*-EntB. The relative abundance of *holo*-EntB formed was then calculated by the equation:

$$(\text{S.I. at each time point})/(\text{S.I. at 60 minute}) * 100$$

where S.I. is the signal intensity determined by densitometry. A linear initial rate was observed for the reaction catalyzed by EntD WT and H44A with BODIPY-TMR-CoA or Alexa-CoA. However, Alexa-PPant was less efficiently transferred to EntB as observed from the lower initial velocity (Figure 34) with both enzymes. EntD WT had a velocity of 0.44 µM/min with BODIPY-TMR-CoA and 0.35 µM/min with Alexa-CoA. H44A had a higher velocity with both conjugates consistent with the velocities determined by ESI-

MS. H44A had a velocity of 0.60 $\mu\text{M}/\text{min}$ with BODIPY-TMR-CoA and 0.38 $\mu\text{M}/\text{min}$ with Alexa-CoA.

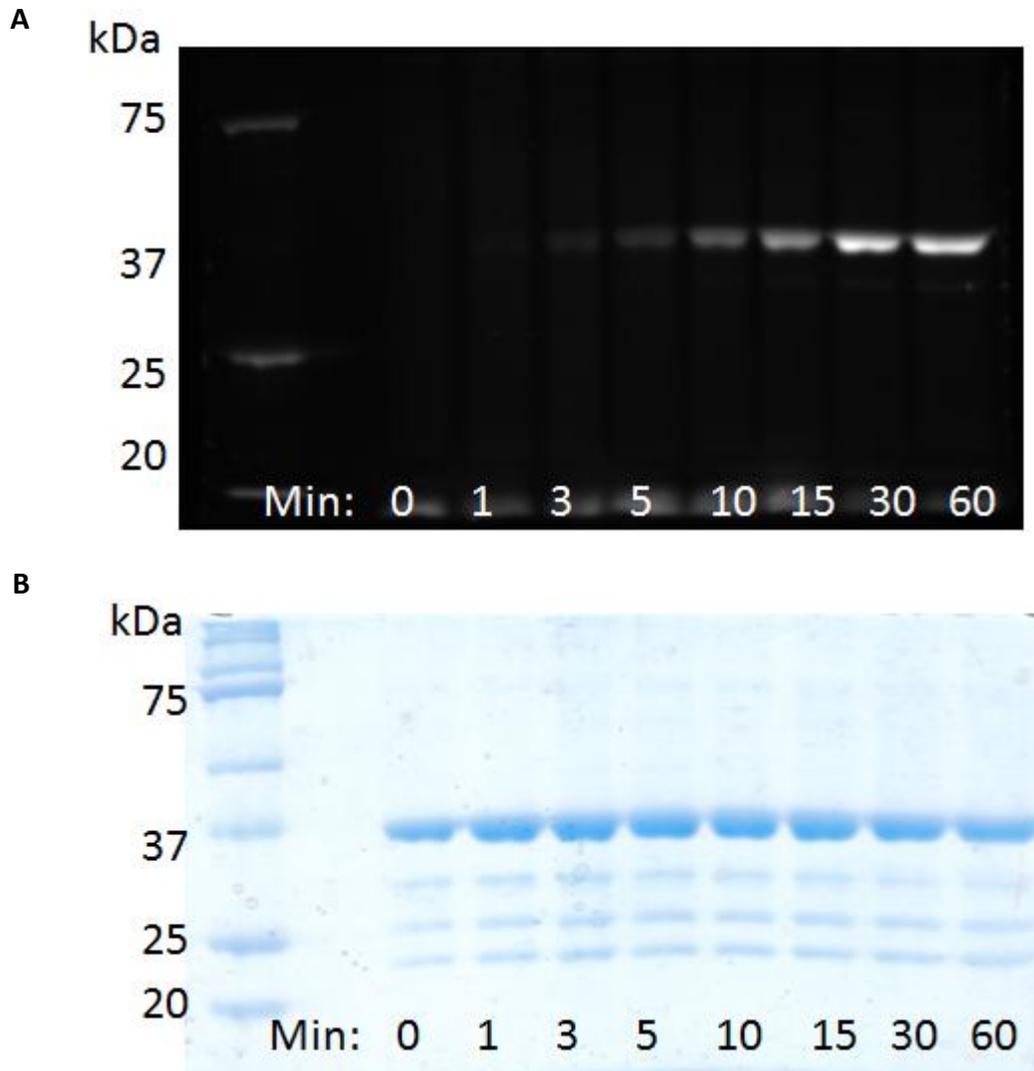


Figure 33 10% SDS-PAGE of time course assay of EntD WT with BODIPY-TMR-CoA.

(A) Fluorescently-labelled holo-EntB visualized by UV light. (B) Coomassie-stain of the gel. *holo*-EntB labeled with BODIPY-TMR-PPant band intensity increase with reaction time. 6.5 μg of EntB at each time point was loaded in each lane. 50 nM EntD WT, 10 μM BODIPY-TMR-CoA, 10 mM MgCl_2 , and 340 μg were incubated for 60 minutes.

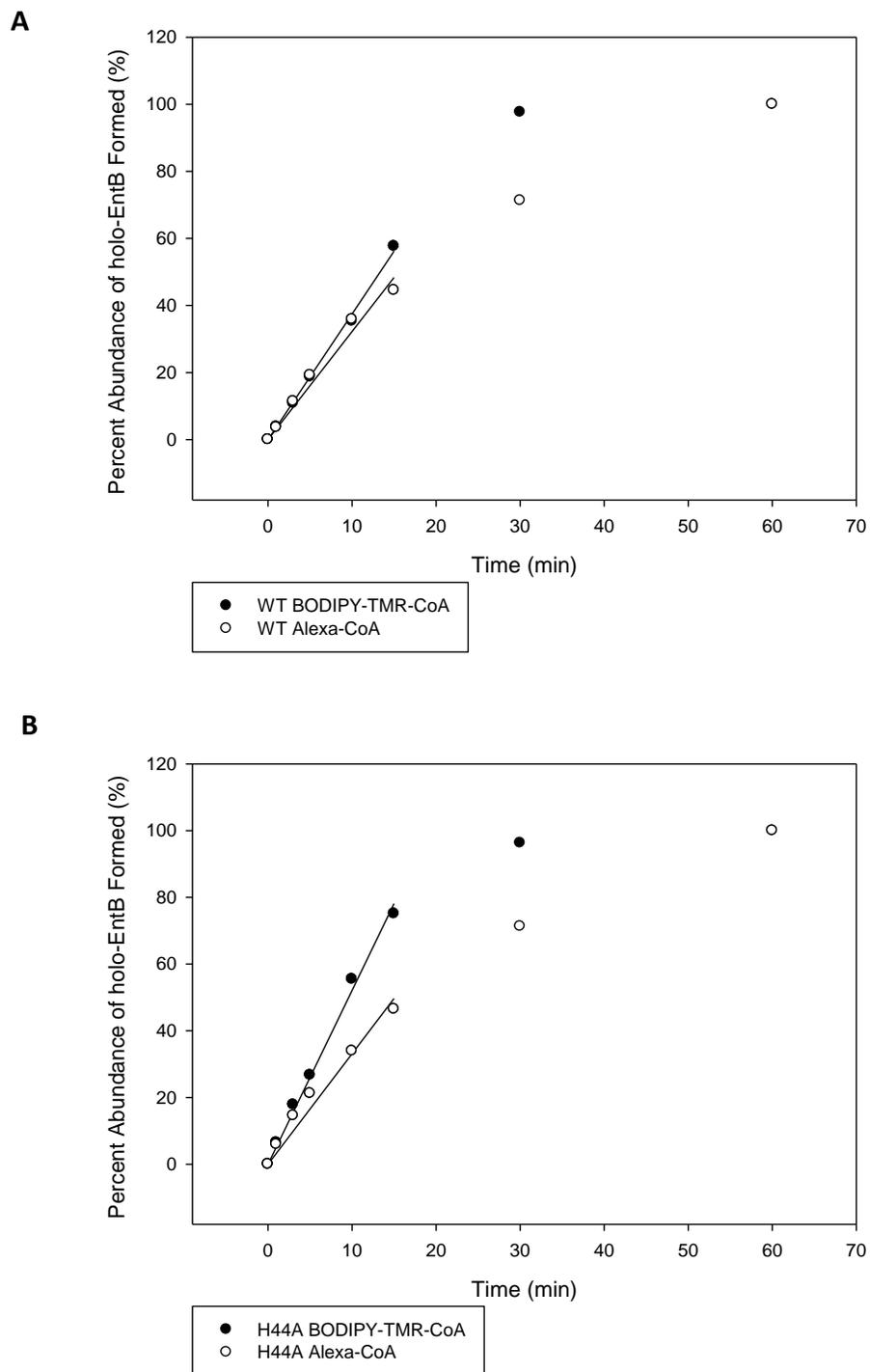


Figure 34 Time course of the reaction of *apo*-EntB with increasing time. *apo*-EntB was reacted with EntD WT (A) and H44A (B) with BODIPY-TMR-CoA or Alexa-CoA. Data shown are an average of two separate reactions. Initial velocity was determined from the first 15 minutes.

3.2. Characterization of EntB Variants

The second part of my project involved the characterization of EntB variants in which residues predicted to interact with EntE were mutated. Site-directed mutagenesis was used to change two of these residues (W277, R282) into alanine. The stability of the all three purified proteins was compared with native gel, and mass spectrometry. Furthermore, the ability of EntD to convert *apo*-EntB into *holo*-EntB was determined.

3.2.1. Site-Directed Mutagenesis

Phage display experiments by Sofia Khalil identified two potential EntB-EntE interface residues on the ArCP domain of EntB. The two residues identified were W277 and R282. W277 is located in the middle of $\alpha 4$ of the ArCP domain of EntB. R282 is positioned C-terminally of $\alpha 4$ (Figure 35 and Figure 36). The two residues were mutated to alanine (W77A, and R282A). Figure 36 does not include the 6XHIS tag, the seven spacer amino acids at the N-terminal end, nor the five spacer amino acids at the C-terminal end added by the plasmid. The identification of the mutants was confirmed by DNA sequencing of the resulting plasmids at the McGill University and Genome Quebec Innovation Center.

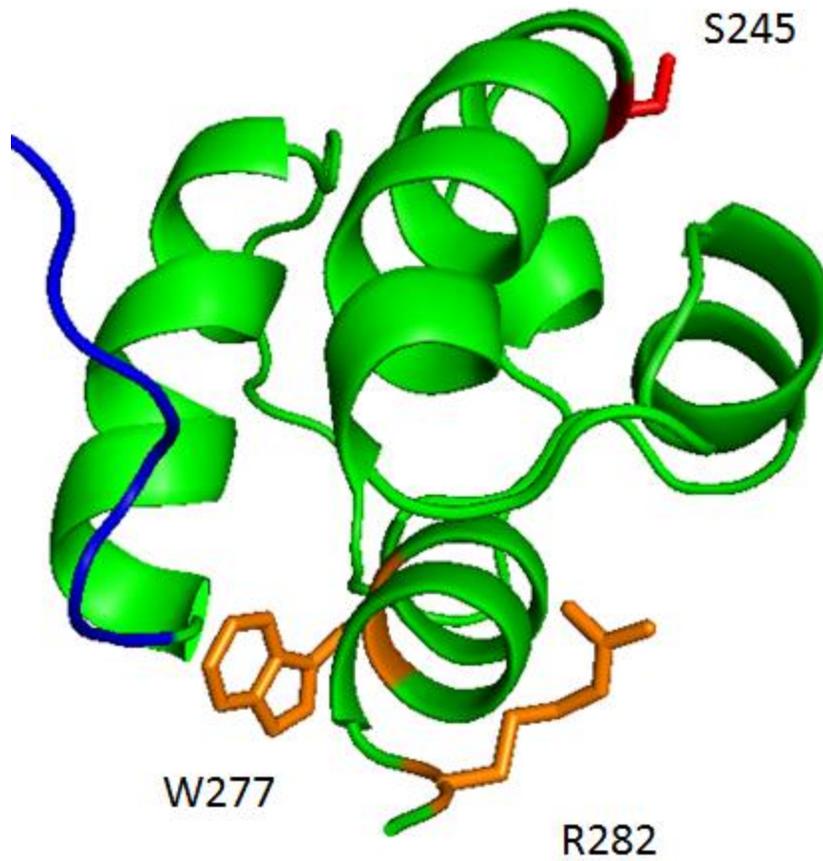


Figure 35 Ribbon Representation of EntB ArCP domain.

The conserved serine residue is highlighted in red and the two EntB-EntE interface residues (W277,R282) on ArCP domain (green) identified by phage display are highlighted in orange. (PDB code: 2FQ1)

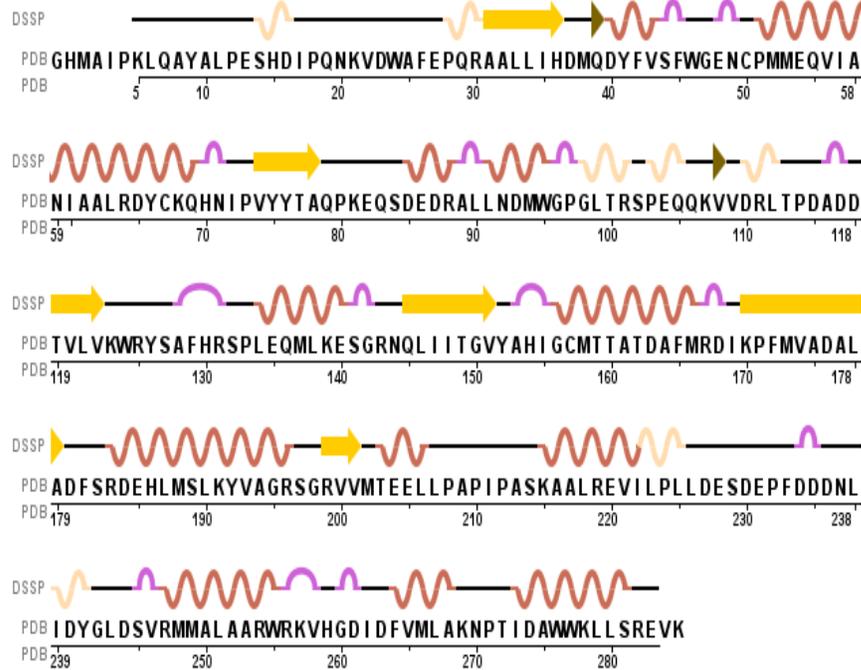


Figure 36 Amino acid sequence of EntB with secondary structure assignment. W277 and R282 are C-terminally positioned. Figure from www.pdb.org.

3.2.2. Purification of EntB WT and Variants

Purification of EntB WT and variants consisted of two steps. The first step was similar to the purification of EntD. Partial purification of EntB WT and variants was achieved with a 10 ml Profinity IMAC column connected to a Bio-Rad FPLC Biologic Duoflow system (Figure 37). A linear imidazole gradient of 20-250mM eluted EntB at approximately 30% of the gradient, which corresponds to 75 mM imidazole. The eluted protein was then dialyzed to remove imidazole and lower the salt concentration from 300 mM to 100 mM. The dialyzed protein was then purified on a 10 ml UNO Q anion exchange column. EntB WT and variants eluted at 55% of a 500 mM KCl gradient. Full-length EntB corresponds to the band nearest to 37 kDa. The two minor bands below the

full-length EntB band were analyzed by LC-MS/MS (McGill Mass Spectrometry Core Facility) and found to correspond to proteolytic cleaved EntB fragments. The band at 30 kDa contains the ICL domain and the N-terminal part of the ArCP domain containing the serine 245 residue. The band at 25 kDa corresponds to the ICL domain of EntB. Figure 38 illustrates that further purification of EntB WT and variants was achieved by anion exchange chromatography. Unlike wild-type EntB, both variants show a major band at 30 kDa and minor bands at 37 and 25 kDa.

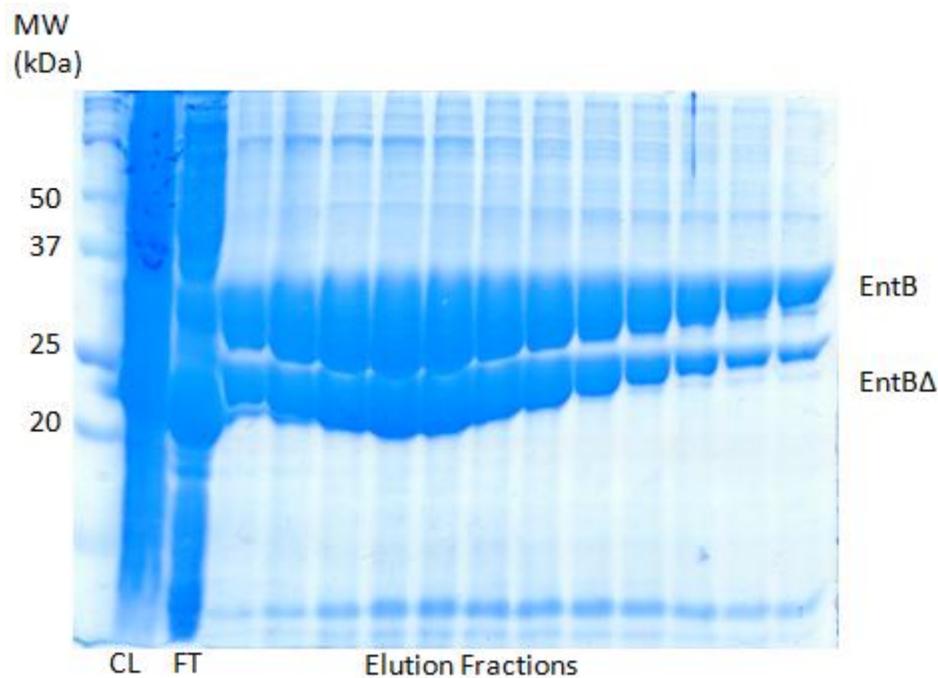


Figure 37 12% SDS-PAGE of EntB WT purification after Profinity Imac Column.

2 μ l cell lysate (CL), 5 μ l of flow-through (FT), and 20 μ l of each fraction were loaded in each lane and protein bands were visualized using Blue Silver Stain.

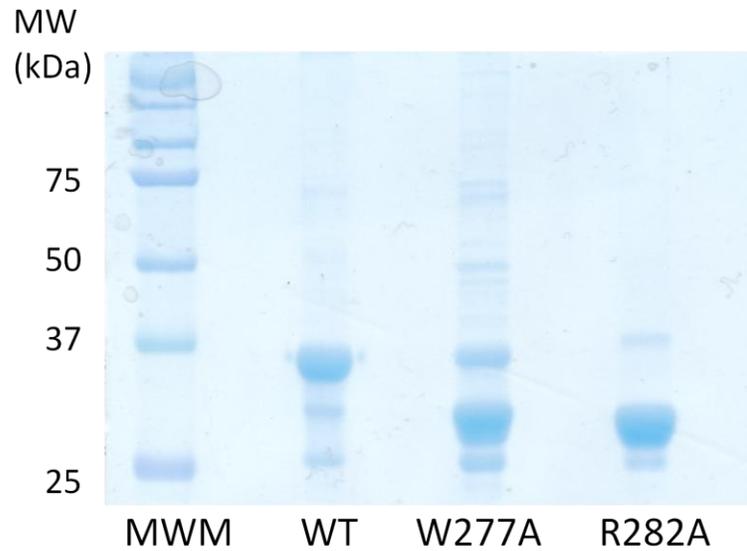


Figure 38 12% SDS-PAGE of EntB WT, W277A, and R282A after anion exchange chromatography.

10 μ g of each protein was visualized by Blue Silver stain.

3.2.3. Native Gel of EntB WT and Variants

The size difference between wild-type EntB and variants was also illustrated by native polyacrylamide gel electrophoresis (Figure 39). The native gel consisted of only a resolving gel. The stacking gel was omitted as were any denaturants such as SDS, and β -mercaptoethanol from the running buffer and loading buffer, respectively. The molecular weight marker was not well resolved and therefore it is difficult to assign a molecular weight to the protein bands. In any case, the native gel clearly indicates that the variants are of smaller size than wild-type EntB.

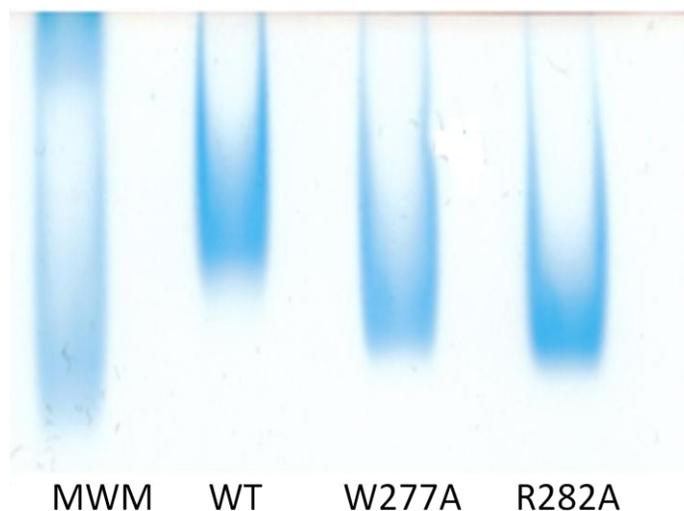


Figure 39 6% Native polyacrylamide gel of EntB WT and variants.

The molecular weight marker (MWM) was not well resolved and therefore protein sizes were not estimated. Proteins were visualized by Blue Silver stain.

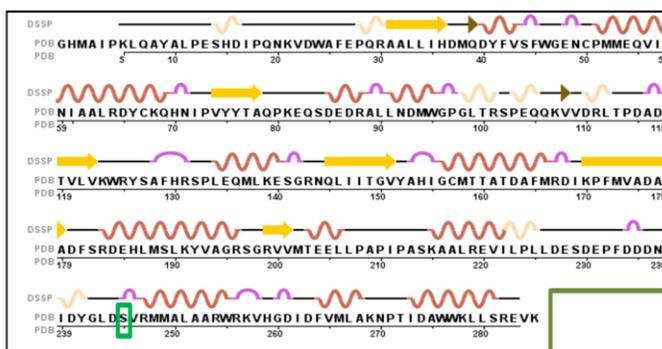
3.2.4. Molecular Weight Determination of EntB WT and Variants

The exact molecular weight of the variants was determined by ESI-MS similarly to EntB WT. The mass spectra of the variants showed multiple minor protein peaks between 30-40 kDa. The major protein peak for W277A and R282A corresponded to a mass of 30437.7 and 29919.17 Da, respectively. As Table 8 summarizes, W277A and R282A had a mass difference of 4452.2 and 4969.2 Da compared to EntB WT. Since the proteins are N-terminally hexahistidine-tagged, it is assumed that the cleavage occurs from the C-terminal end of the protein. These cleavages correspond to a majority of ArCP domain being cleaved. However, the variant proteins still contain Ser245, the site of phosphopantetheinylation (Figure 40).

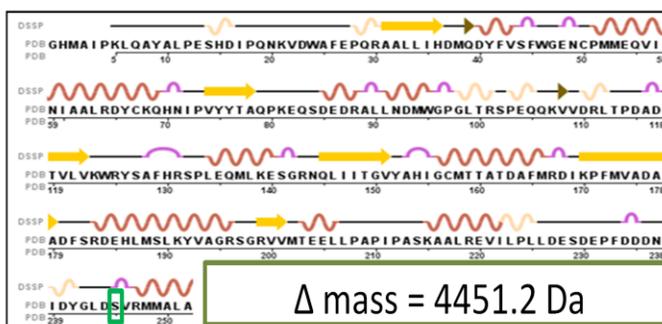
Protein	Observed Mass (Da) from ESI-MS	Δ Mass (Da)
EntB WT	34888.9	-
EntB W277A	30437.7	4451.2
EntB R282A	29919.17	4969.2

Table 8 Molecular weight determination of EntB WT and variants by ESI-MS.

EntB WT



EntB W277A



EntB R282A

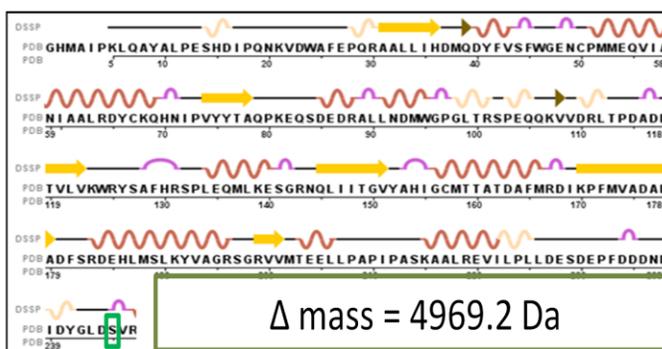


Figure 40 Schematic representation of cleaved residues of EntB variants.

ESI-MS data indicate that the C-terminally cleaved EntB variants still contain Ser245, the site of phosphopantetheinylation. Single point mutations at EntB positions 277 and 282 result in large C-terminal truncations of the protein.

3.2.5. Determination of EntD Activity on Variants

Since Ser245 was intact in the variants, the activity of EntD was then determined using these variants as substrates in our ESI-MS activity assay. Both *apo*-EntB variants, W277A and R282A, were reacted with EntD for 60 minutes in the presence of 200 μ M CoA and 10 mM Mg²⁺. ESI-MS determined that although these variants contain Ser245, they were not able to become pantetheinylated. A mass increase of 339 Da was not observed for W277A and R282A (Table 9). This suggests that the C-terminal portion of the EntB ArCP domain is required for pantetheinylation by EntD.

Protein	Observed Mass (Da) Before Reaction	Expected Mass (Da) After Reaction	Observed Mass (Da) After Reaction	Δ Observed Mass (Da)
EntB WT	34887.6	35227.9	35226.2	338.6
W277A	30437.7	30776.7	30437.7	0
R282A	29919.7	30258.1	29919.0	0.7

Table 9 Summary of EntB (WT, variants) mass determinations after reaction with EntD as measured by ESI-MS.

4. Discussion

4.1. Research Objective Revisited

Phosphopantetheinyl transferases are required for the transfer of 4'-phosphopantetheine from CoA to a conserved serine residue on acyl carrier protein domains or subunits. This posttranslational modification is essential for the functional activation of all fatty acid synthases, polyketide synthases, and non-ribosomal peptide synthetase complexes. The newly added thiol on the carrier protein serves as the key anchor point for tethering activated thioesters during chain elongation. The specific function of EntD was first determined by Lambalot *et al.* (1996) when they identified phosphopantetheinyl transferases as an enzyme superfamily. Sfp-type PPTases have a weak sequence homology with Sfp (Figure 8). However, sequence similarities in the C-terminal half of the PPTases with Sfp give rise to potential ligand binding residues. Additionally, the homology-based model of EntD identified two potential ligand binding residues. The first residue identified from the homology model of EntD was H44. Although, H44 is not a conserved residue among PPTases, the homology model of EntD identified it as interacting with the 3'phosphate of CoA. E155 was the second residue to be identified. It is highly conserved among all PPTases and in Sfp it has been identified to take part in metal complexation with Mg^{2+} . However, E155 in the generated homology model of EntD interacts with the adenine base of CoA. The difference between the E155 residue of EntD and Sfp is approximately 3.6 residues, which corresponds to one α -helical turn. This discrepancy most likely resulted from generating a homology model of

EntD using a protein with a low sequence similarity. The focus of this thesis was on the mutations changing these two residues into alanine. EntD H44A and E155A variants were produced in order to study the role of H44 and E155 and the effects of H44A and E155A mutations on the protein structure, stability, and activity.

4.2. Structural Characterization of Wild-Type EntD and Variants

H44A and E155A PPant transferases were prepared by site-directed mutagenesis of the *E. coli* gene encoding wild-type EntD. The introduction of the desired substitutions was confirmed by DNA sequence analysis. DNA sequence analysis also showed that unwanted mutations had not been incorporated. Wild-type EntD and both variants were purified on a Profinity Imaq column. Initial attempts to purify EntD from cells induced with 1 mM IPTG and grown for 4 hours at 37 °C using Ni²⁺-column affinity chromatography resulted in the coexpression and purification of the *E. coli* chaperonin GroEL (Figure 15). GroEL is a 60 kDa protein that functions to help fold or sequester misfolded proteins (Xu *et al.*, 1997). This contamination indicates that some EntD in the preparation likely associate with GroEL in order to remain soluble. This could explain that high concentrations of EntD wild-type and variants (> 1 mg/ml) in the absence of glycerol result in protein precipitation. Gill *et al.* (2006) showed that lowering the growth temperatures after induction reduces GroEL contamination from purification of the cytoplasmic N-terminal domain of the Na/HCO₃ cotransporter NCB1-A. Similarly, GroEL contamination from the EntD purification was considerably reduced by lowering the growth temperature to 25 °C after induction, increasing the induction time to 16

hours, and reducing IPTG concentration to 0.5 mM (Figure 16). In general, approximately 15 mg of wild-type EntD was obtained per 1 L culture whereas 7 mg of H44A and 10 mg of E155A were obtained per 1 L culture. Long term storage of the proteins in 15% glycerol prevented protein precipitation and protein degradation.

The wild-type and variant proteins were compared by SDS-PAGE, far-UV CD, and fluorescence spectroscopy. SDS-PAGE demonstrates that all three proteins are of similar size and that the variants do not undergo proteolysis. Although an exact molecular weight determination is not achieved by SDS-PAGE, it does reveal that the proteins are stable and not sensitive to proteolysis. The secondary structure of wild-type and variant proteins was determined by far-UV circular dichroism. CD spectra of all three proteins exhibit the double minima at 208 and a reduced signal at 222 nm, corresponding to a mixture of α -helices and β -strands. The CD spectra for WT and H44A overlap indicating a near identical secondary structure (Figure 17). However, the spectrum of E155A is more pronounced which might be a result of a higher protein concentration. CD spectra is sensitive to protein concentration, therefore, an increase in E155A concentration might explain the higher signal. K2D was used to analyze the CD spectra obtained. K2D uses an algorithm for the estimation of the percentages of protein secondary structure from UV circular dichroism spectra using a Kohonen neural network with a 2-dimensional output layer. The secondary structure features are determined from the circular dichroism spectra ranging from 200-240 nm. This analysis showed that the global secondary structure of all three proteins were similar. An α -helical content of 31% was calculated by K2D for EntD WT, and 30% for both variants. This value is in strong agreement with

the α -helical content of 27% determined by PSIPPRED from the primary sequence of EntD. The percent of β -strands between all three proteins were also similar with a percentage of 17% for EntD WT, and 19% and 15% for H44A and E155A, respectively, which is in agreement with PSIPRED which determined a β -strand content of 21%. However, the homology-based model of EntD exhibits a slight difference in the secondary structure with an α -helical content of 38% and a β -strand content of 28%. The difference between the EntD model and PSIPRED is due to the fact that the former does not include the 21 additional residues added by the plasmid. PSIPRED predicts these residues as random coil, therefore decreasing the α -helical and β -strand content. The thermal stabilities of all three proteins were determined by monitoring the loss of α -helical content at 222 nm. Table 3 demonstrates that the thermal stability of all three protein were similar with a melting temperature of 37.2, 35.4, and 38.2 °C for wild-type EntD, H44A and E155A, respectively. Fluorescence spectroscopy was used to assess the tertiary structure of all three proteins. Glutamate and histidine residues do not directly contribute to the intrinsic fluorescence of proteins. However, tryptophan fluorescence can be quenched by the polar functional groups of proximal glutamate and histidine residues (Eftink, 2000). Therefore, mutating H44 and E155 into alanine could have an effect on the emission spectra, where we would expect an increase in the fluorescence intensity. However, it should be noted that the homology-based model of EntD does not show E155 to be proximal to any of the four tryptophan residues while the side chain of H44 is 5.3 Å away from a tryptophan residue. Indeed, the emission spectra of the proteins excited at 280 nm demonstrate that H44A and E155A have overlapping spectra

and have a higher fluorescence intensity signal than EntD wild-type. Additionally, both variants are slightly blue-shifted which corresponds to an environment change around the aromatic amino acids, tyrosine and tryptophan. This is complemented with the emission spectra of all three proteins when excited at 295 nm. The emission spectra of both variants are blue-shifted. Interestingly, the fluorescence intensities of all three proteins slightly differ. E155A has a slightly higher intensity than EntD WT and H44A has lower signal intensity. The increase in fluorescence intensity for E155A might be due to the removal of a glutamate residue which can quench the tryptophan fluorescence signal. The slight variations in intensity could also be caused by concentration differences. However, the observed blue-shifts indicate that the tyrosine and tryptophan residues are in a more hydrophobic environment suggesting a possible minor change in the tertiary structure of the variants.

4.3. Overall Structural Effect of Cofactor Binding

Biophysical experiments were used to determine ligand binding to EntD and to monitor the structural effect upon ligand binding. Far-UV CD experiments revealed that the secondary structure of all three proteins did not change in the presence of both cofactors. We observe that in the presence of both cofactors the secondary structure of all three proteins remained similar (Figure 20 and Figure 21). Monitoring the thermal stability of the protein provided the first evidence for cofactor binding. Binding of both ligands increased the thermal stability by approximately 6.2 °C for EntD WT and E155A and as much as 10 °C for H44A. In contrast, the presence of only one cofactor

marginally changed the melting temperature with an average difference of 2.6 °C. An increase in hydrogen bonds results in an elevated thermal stability of proteins (Vogt *et al.*, 1997). Binding of CoA and Mg²⁺ in the active site changes the conformation of the protein, bringing the two halves of the protein together and increasing the amount of hydrogen bonds. This increase in hydrogen bond number results in a more rigid protein, thus increasing its thermal stability. Electrostatic interactions and hydrogen bonding between the cofactors and protein might also increase the thermal stability of the EntD proteins.

The results from cofactor binding using near-UV CD complement the data obtained from far-UV CD. Figure 23 demonstrates that a maximum signal for all three proteins was observed between 250-290 nm in the presence of both cofactors. This corresponds to the aromatic amino acids, phenylalanine, tyrosine, and tryptophan being held more rigidly and losing mobility due to cofactor binding. Therefore, near-UV CD data shows that all three proteins are able to bind to Mg²⁺ and CoA. However, the difference in signal between the *apo* and *holo* (+Mg²⁺, CoA) forms of the EntD proteins is less pronounced in the E155A variant. This indicates a minor change in the tertiary structure of E155A in the absence of cofactors. This finding is in good agreement with our fluorescence spectroscopy experiments.

Sedimentation velocity experiments of wild-type EntD also demonstrated that the protein underwent a conformational change upon cofactor binding. The conformational change was monitored by the change in the frictional ratio and sedimentation

coefficient values. Frictional ratio indicates the overall shape of a macromolecule. The frictional ratio is defined as the ratio the frictional coefficient (f) experienced by the molecule when sedimenting to the theoretical frictional coefficient (f_0) for an ideal sphere of corresponding molecular weight (Smith, 1988). Therefore, an ideal sphere will have a frictional ratio of 1.0, while many globular proteins have values of f/f_0 of 1.05-1.35 and elongated proteins have a ratio above 1.4 (Smith, 1988). In the absence of cofactors, EntD WT had a frictional ratio of 1.68 which corresponds to an elongated protein. However, a 20% decrease in the frictional ratio (1.37) was observed in the presence of both cofactors (Table 4). This further suggests that the protein is undergoing a conformational change upon ligand binding, going from an elongated to a more globular shape. The sedimentation coefficient indicates the rate at which the sedimentation boundary moves. It depends upon a protein's concentration, shape, size, and degree of hydration. Therefore, our reported sedimentation coefficient values were corrected to 20 °C and for the density of water. The presence of both cofactors increased the sedimentation coefficient value from 2.32 to 2.46. The elongated EntD protein will experience more hydrodynamic friction and thus will have a smaller sedimentation coefficient value than the globular protein in the presence of both cofactors.

The apparent dissociation constant, K_d , for the binding of conjugated CoA to all three proteins was determined by fluorescence anisotropy experiments. The apparent K_d was determined by titrating the BODIPY-TMR-CoA ligand with protein in the absence or presence of Mg^{2+} . The apparent K_d for Alexa-CoA was not determined since the non-

extracted unreacted dye contributed to a low signal to noise ratio. The processing of the data by non-linear regression analysis revealed that all three proteins bind BODIPY-TMR-CoA with approximately the same affinity ($K_d \sim 2 \mu\text{M}$) (Table 5). The apparent dissociation constants determined are not in agreement with the reported apparent K_d of $0.18 \mu\text{M}$ determined by fluorescence polarization and $0.69 \mu\text{M}$ determined by ITC for BODIPY-TMR-CoA binding to Sfp (Duckworth *et al.*, 2010). This decrease in K_d indicates that Sfp has a higher affinity for CoA than EntD. In fact, a K_m of 700 nM for Sfp with CoA was determined (Duckworth *et al.*, 2010) as opposed to a K_m of $17 \mu\text{M}$ for EntD (Gehring *et al.*, 1997). Although an increase in fluorescence anisotropy was observed with the addition of 10 mM Mg^{2+} , the apparent K_d was not affected. The increase in signal can be explained by a conformational change of the protein. This ligand-induced conformational change causes the protein to fold inward, thereby restricting the tumbling motion of the fluorescent labeled portion of CoA, and which in turn, results in a fluorescence anisotropy increase. This strongly correlates with our CD and AUC data which showed that a major conformational change was observed in the presence of both Mg^{2+} and CoA. We also observe an increase in the fluorescence anisotropy signal when BODIPY-TMR-CoA was titrated with E155A variant compared to wild-type EntD and H44A titrations. The crystal structure of Sfp, and to some extent the homology-based model of EntD illustrate that the adenine base of CoA is held tightly in the active site. However, the pantetheine group of CoA rises from the CoA-binding pocket into the solvent and makes no interactions with the protein (Reuter *et al.*, 1999). This allows the fluorescently-labeled pantetheine group to have a certain degree of mobility. A tertiary

structure difference of E155A determined by our fluorescence spectroscopy and near-UV CD data could explain the increase in fluorescence anisotropy observed. The slightly different structure could hinder the conjugated fluorescent label to rotate freely, which results in a higher fluorescence anisotropy value.

4.4. Measurement of EntD Enzyme Activity

The PPTase activity of all three proteins was determined by ESI-MS and gel-based assay. The former technique consists of observing a change in the mass of *apo*-EntB corresponding to the addition of the 4'-phosphopantetheine while the latter technique consists of observing the transfer of the fluorescently-labeled 4'-phosphopantetheine from CoA to EntB with a UV light. ESI-MS experiments demonstrated that both EntD WT and H44A were active and able to post-translationally modify *apo*-EntB while E155A was inactive. Time-course assay plots for EntD WT and H44A catalyzed reactions revealed an initial velocity of 0.66 and 1.03 $\mu\text{M}/\text{min}$, respectively. Surprisingly, H44A had a 56% faster rate than wild-type EntD. It is difficult to explain this rate increase without knowing the exact position of H44 relative to the CoA-binding pocket since the CoA-binding site residues in the homology-based model of EntD are misaligned. Therefore, a more refined homology-based model of EntD needs to be generated in order to have a better knowledge of the CoA-binding residues. Post-translational fluorescent labeling of carrier proteins by Sfp have previously been reported (Mercer *et al.*, 2005; Yin *et al.*, 2005; Foley *et al.*, 2009; Duckworth *et al.*, 2010). This concept was adapted to determine the activity of all three proteins by transferring the BODIPY-TMR-PPant or

Alexa-PPant group from CoA to *apo*-EntB. Figure 32A illustrates that EntD WT and H44A were both active while E155A was unable to transfer both fluorescently labeled PPant groups from CoA, therefore supporting our ESI-MS data. The inactivity of E155A with the fluorescently-labelled CoA was not due to the label since E155A was also inactive with unlabelled CoA determined by ESI-MS. We also observe two minor fluorescent bands at 30 kDa and 70 kDa (Figure 32A). LC-MS/MS data revealed that Ser245 remained intact in the 30 kDa EntB fragment and that it was absent from the 25 kDa EntB fragment. The fluorescent band at 30 kDa confirmed that Ser245 remained intact. The minor band at 70 kDa on the SDS-PAGE indicates the presence of EntB dimer. The conversion of E to A at position 155 in EntD resulted in an inactive enzyme. One of the explanations for this observation is that the minor structural change of E155A, as shown by far-UV CD, near-UV CD and fluorescence spectroscopy experiments, might prevent docking of *apo*-EntB. Circular dichroism and fluorescence anisotropy experiments revealed that cofactor binding occurs with E155A. Therefore, another possible explanation is that E155 is involved in catalysis rather than just cofactor binding. However, further experiments would need to be conducted to characterize the role of E155 in catalysis. Quadri *et al.* (1998) reported that mutation of the corresponding glutamate residue (E151) in *B. subtilis* Sfp to alanine also resulted in a loss of PPTase activity. Interestingly, they also concluded that the mutant is either defective in stability (e.g., folding) or catalysis. The time-course assay plots determined an initial velocity for wild-type EntD of 0.44 $\mu\text{M}/\text{min}$ and 0.35 $\mu\text{M}/\text{min}$ for BODIPY-TMR-CoA and Alexa-CoA, respectively. Similarly to our MS data, H44A had a higher initial rate with 0.60 $\mu\text{M}/\text{min}$ and 0.38 $\mu\text{M}/\text{min}$ for BODIPY-

TMR-CoA and Alexa-CoA, respectively. A more substantial difference is observed between the initial velocities determined by the two techniques. This significant difference in the initial velocities determined between the two techniques can be caused by many factors. First, the sensitivity between the two techniques is different. MS is a highly sensitive technique capable of accurately determining the amount of *holo*-EntB formed at each time point. In contrast, the sensitivity of the gel-based assay is low since it depends on the exposure time and fluorescence intensity determination. Second, CoA concentrations in both experiments differ. In the MS assay, 200 μ M CoA was utilized, corresponding to tenfold K_m (17 μ M) (Gehring *et al.*, 1997). However, the gel-based assay only consisted of 10 μ M of fluorescently-labeled CoA. Third, the conjugation state of CoA differs between the two experiments. The presence of a label on CoA in the gel-based experiments could have affected the binding affinity of CoA to EntD. Furthermore, labeled-CoA could also affect binding of *apo*-EntB to EntD. In any case, this study demonstrates that the gel-based assay can be used to compare the initial velocities of wild-type EntD and variants. Table 7 demonstrates the EntE-catalyzed acylation of *holo*-EntB with 2,3-DHB. These findings prove that our hexahistidine-tagged EntD, EntB, and EntE proteins are active.

4.5. Structural Characterization of EntB and Variants

The mutation of two residues, W277A and R282A, in the EntB ArCP domain resulted in the proteolytic fragmentation of EntB. These fragmentations prohibited the study of EntB and EntE interactions. It appears that EntB is sensitive to fragmentation

since Gehring *et al* (1997) have also reported proteolytic fragmentation of their EntB Δ C construct. SDS-PAGE was the first technique to identify a discrepancy in the size between wild-type EntD and variants (Figure 38). The major band at 37 kDa observed for WT is only weakly present for both variants. The variants both have a major band at 30 kDa. Similar results were also obtained from native gel. Molecular weight measurements using MS confirmed that both variants contained multiple fragments with a major peak at 30437.7 and 29919.17 Da for H44A and E155A, respectively. Although we assumed that Ser245 is still intact in both variants, EntD was unable to activate these carrier proteins. This suggests that the C-terminal 100 amino acids are crucial for pantetheinylation if we assume that the rest of the protein is properly folded. Therefore, the C-terminal two-thirds of the ArCP domain might play a role in recognition and binding to EntD. This extends the findings from Lai *et al.* (2006) which determined that the amino acids N-terminally positioned to Ser245 were only required for EntD-EntB interactions. PeptideCutter (ca.expasy.org/tools/peptidecutter) was used to analyze the potential cleavage sites cleaved by proteases or chemicals in the EntB protein sequences. EntB variants were not prone to any additional proteases than EntB WT. The addition of protease inhibitor cocktail following cell lysis and the fact that PeptideCutter did not identify any additional cleavage sites for the EntB variants strongly suggesting that the EntB variants are cleaved inside the cell.

5. Summary and Future Work

The goal of this study was to express, purify, and characterize *E. coli* EntD, and perform extensive biophysical and biochemical characterization of wild-type EntD and two variants, H44A and E155A.

Sequence alignment of *E.coli* EntD with Sfp from *Bacillus subtilis* has identified conserved residues in the active site (Reuter *et al.*, 1999). The homology-based model of EntD identified two potential cofactor binding residues. E155 was identified as interacting with the adenine base of CoA. In contrast, the corresponding residue in Sfp is involved in Mg^{2+} complexation. The generated homology-based model of EntD also identified H44 as interacting with the 3'-phosphate of CoA. However, this residue is not among the conserved residues of PPTases. The EntD model also failed to identify the remaining cofactor binding residues. Therefore, more refined homology-based models of EntD need to be regenerated by Modeller. It would also be of interest to mutate the conserved residues in EntD (i.e. D110, E112, K159) to characterize active site residues. E155 and H44 were converted to alanine to characterize their role in cofactor binding and activity. Site-directed mutagenesis was performed on wild-type EntD in order to obtain the desired mutations. Overexpression and purification of both variants were carried out following the same processes as those of the wild-type. A concentration above 1 mg/ml resulted in the precipitation of the EntD proteins. Therefore, it would be useful to study EntD stability with different buffers and stabilizing agents (i.e. amphiphiles, viscogens).

Biophysical characterization using circular dichroism and sedimentation velocity experiments of the EntD proteins revealed a conformational change upon cofactor binding. AUC experiments were not conducted on EntD variants due to protein precipitation. Therefore, increasing the protein stability with stabilizing agents would be advantageous for running future AUC experiments. Fluorescence anisotropy experiments also confirmed that EntD WT and variants bind conjugated CoA with an average apparent K_d of 2 μ M. The study presented in this thesis represents the first reported binding affinity of EntD with conjugated CoA. Competitive displacement assay could be used to further investigate the binding of CoA to wild-type EntD and variants.

The results from mass spectrometry and gel-based assay determined that EntD WT and H44A were active while E155A was unable to activate *apo*-EntB. The initial velocities for EntD WT and H44A determined by gel-based assay differed by at least 50% from the MS experiment. It will therefore be essential to use conjugated CoA in our MS-based assay to see if consistent initial velocity differences would be observed.

The second part of this project involved converting two potential EntB-EntE interacting residues on EntB. However, the proteolytic fragmentation of these two EntB variants prevented the study. SDS-PAGE, native polyacrylamide gel, and ESI-MS demonstrated that the EntB variants were unstable and resulted in their cleavage. Since the 6xHis-tag is positioned N-terminally, we assumed that the C-terminal amino acids are prone to cleavage. Peptide mass fingerprinting would be useful to accurately identify EntB variant fragments. Also, it would be interesting to determine if expression

of wild-type EntD and variants in iron-restricted *E. coli* cells can prevent proteolysis due to metabolon protection.

6. References

- Bleuel, C., Grosse, C., Taudte, N., Scherer, J., Wesenberg, D., Krauss, G. J., *et al.* (2005). TolC is involved in enterobactin efflux across the outer membrane of escherichia coli. *Journal of Bacteriology*, *187*(19), 6701-6707.
- Campbell, T. L., & Brown, E. D. (2008). Genetic interaction screens with ordered overexpression and deletion clone sets implicate the escherichia coli GTPase YjeQ in late ribosome biogenesis. *Journal of Bacteriology*, *190*(7), 2537-2545.
- Carrasco, N., Kretschmar, R., Xu, J., & Kraemer, S. M. (2009). Adsorption of hydroxamate siderophores and EDTA on goethite in the presence of the surfactant sodium dodecyl sulfate. *Geochemical Transactions*, *10*, 5.
- Chen, D., Wu, R., Bryan, T. L., & Dunaway-Mariano, D. (2009). In vitro kinetic analysis of substrate specificity in enterobactin biosynthetic lower pathway enzymes provides insight into the biochemical function of the hot dog-fold thioesterase EntH. *Biochemistry*, *48*(3), 511-513.
- Dave, B. P., Anshuman, K., & Hajela, P. (2006). Siderophores of halophilic archaea and their chemical characterization. *Indian Journal of Experimental Biology*, *44*(4), 340-344.
- Devireddy, L. R., Hart, D. O., Goetz, D. H., & Green, M. R. (2010). A mammalian siderophore synthesized by an enzyme with a bacterial homolog involved in enterobactin production. *Cell*, *141*(6), 1006-1017.
- Drake, E. J., Nicolai, D. A., & Gulick, A. M. (2006). Structure of the EntB multidomain nonribosomal peptide synthetase and functional analysis of its interaction with the EntE adenylation domain. *Chemistry & Biology*, *13*(4), 409-419.
- Duckworth, B. P., & Aldrich, C. C. (2010). Development of a high-throughput fluorescence polarization assay for the discovery of phosphopantetheinyl transferase inhibitors. *Analytical Biochemistry*, *403*(1-2), 13-19.
- Duckworth, B. P., & Aldrich, C. C. (2010). Development of a high-throughput fluorescence polarization assay for the discovery of phosphopantetheinyl transferase inhibitors. *Analytical Biochemistry*, *403*(1-2), 13-19.
- Earhart, C.F. (2004). Iron Uptake via the enterobactin system. *In Iron transport in bacteria*. ASM Press, NY. 133-146.
- Eftink, M.R. (2000). Intrinsic fluorescence of proteins: Topics in fluorescence spectroscopy, Volume 6. Kluwer academic Plenum Publishers, NY, 1-15.

- Escolar, L., Perez-Martin, J., & de Lorenzo, V. (1999). Opening the iron box: Transcriptional metalloregulation by the fur protein. *Journal of Bacteriology*, *181*(20), 6223-6229.
- Foley, T. L., & Burkart, M. D. (2009). A homogeneous resonance energy transfer assay for phosphopantetheinyl transferase. *Analytical Biochemistry*, *394*(1), 39-47.
- Frueh, D. P., Arthanari, H., Koglin, A., Vosburg, D. A., Bennett, A. E., Walsh, C. T., *et al.* (2008). Dynamic thiolation-thioesterase structure of a non-ribosomal peptide synthetase. *Nature*, *454*(7206), 903-906.
- Furrer, J. L., Sanders, D. N., Hook-Barnard, I. G., & McIntosh, M. A. (2002). Export of the siderophore enterobactin in escherichia coli: Involvement of a 43 kDa membrane exporter. *Molecular Microbiology*, *44*(5), 1225-1234.
- Garcia-Contreras, R., Zhang, X. S., Kim, Y., & Wood, T. K. (2008). Protein translation and cell death: The role of rare tRNAs in biofilm formation and in activating dormant phage killer genes. *PloS One*, *3*(6), e2394.
- Gehring, A. M., Bradley, K. A., & Walsh, C. T. (1997). Enterobactin biosynthesis in escherichia coli: Isochorismate lyase (EntB) is a bifunctional enzyme that is phosphopantetheinylated by EntD and then acylated by EntE using ATP and 2,3-dihydroxybenzoate. *Biochemistry*, *36*(28), 8495-8503.
- Gill, H. S., & Boron, W. F. (2006). Expression and purification of the cytoplasmic N-terminal domain of the Na/HCO₃ cotransporter NBCe1-A: Structural insights from a generalized approach. *Protein Expression and Purification*, *49*(2), 228-234.
- Grass, G. (2006). Iron transport in escherichia coli: All has not been said and done. *Biometals : An International Journal on the Role of Metal Ions in Biology, Biochemistry, and Medicine*, *19*(2), 159-172.
- Hamdan, H., Weller, D. M., & Thomashow, L. S. (1991). Relative importance of fluorescent siderophores and other factors in biological control of gaeumannomyces graminis var. tritici by pseudomonas fluorescens 2-79 and M4-80R. *Applied and Environmental Microbiology*, *57*(11), 3270-3277.
- Hantke, K. (1990). Dihydroxybenzoylserine--a siderophore for E. coli. *FEMS Microbiology Letters*, *55*(1-2), 5-8.
- Hider, R. C., & Kong, X. (2010). Chemistry and biology of siderophores. *Natural Product Reports*, *27*(5), 637-657.

Johnson, W. C. (1999). Analyzing protein circular dichroism spectra for accurate secondary structures. *Proteins*, 35(3), 307-312.

Khalil, S., & Pawelek, P. D. (2009). Ligand-induced conformational rearrangements promote interaction between the escherichia coli enterobactin biosynthetic proteins EntE and EntB. *Journal of Molecular Biology*, 393(3), 658-671.

Kitagawa, M., Ara, T., Arifuzzaman, M., Ioka-Nakamichi, T., Inamoto, E., Toyonaga, H., *et al.* (2005). Complete set of ORF clones of escherichia coli ASKA library (a complete set of E. coli K-12 ORF archive): Unique resources for biological research. *DNA Research : An International Journal for Rapid Publication of Reports on Genes and Genomes*, 12(5), 291-299.

Kurtz, D. M., Holm, R. H., Ruzicka, F. J., Beinert, H., Coles, C. J., & Singer, T. P. (1979). The high potential iron-sulfur cluster of aconitase is a binuclear iron-sulfur cluster. *The Journal of Biological Chemistry*, 254(12), 4967-4969.

Lai, J. R., Fischbach, M. A., Liu, D. R., & Walsh, C. T. (2006). Localized protein interaction surfaces on the EntB carrier protein revealed by combinatorial mutagenesis and selection. *Journal of the American Chemical Society*, 128(34), 11002-11003.

Lai, J. R., Fischbach, M. A., Liu, D. R., & Walsh, C. T. (2006). A protein interaction surface in nonribosomal peptide synthesis mapped by combinatorial mutagenesis and selection. *Proceedings of the National Academy of Sciences of the United States of America*, 103(14), 5314-5319.

Lambalot, R. H., Gehring, A. M., Flugel, R. S., Zuber, P., LaCelle, M., Marahiel, M. A., *et al.* (1996). A new enzyme superfamily - the phosphopantetheinyl transferases. *Chemistry & Biology*, 3(11), 923-936.

Lambalot, R. H., & Walsh, C. T. (1995). Cloning, overproduction, and characterization of the escherichia coli holo-acyl carrier protein synthase. *The Journal of Biological Chemistry*, 270(42), 24658-24661.

Lambalot, R. H., & Walsh, C. T. (1997). Holo-[acyl-carrier-protein] synthase of escherichia coli. *Methods in Enzymology*, 279, 254-262.

Leduc, D., Battesti, A., & Bouveret, E. (2007). The hotdog thioesterase EntH (YbdB) plays a role in vivo in optimal enterobactin biosynthesis by interacting with the ArCP domain of EntB. *Journal of Bacteriology*, 189(19), 7112-7126.

Mercer, A. C., La Clair, J. J., & Burkart, M. D. (2005). Fluorescent multiplex analysis of carrier protein post-translational modification. *ChemBiochem : A European Journal of Chemical Biology*, 6(8), 1335-1337.

- Miethke, M., & Marahiel, M. A. (2007). Siderophore-based iron acquisition and pathogen control. *Microbiology and Molecular Biology Reviews : MMBR*, 71(3), 413-451.
- Morgan, J. W., & Anders, E. (1980). Chemical composition of earth, venus, and mercury. *Proceedings of the National Academy of Sciences of the United States of America*, 77(12), 6973-6977.
- Otaka, E., & Ooi, T. (1989). Examination of protein sequence homologies: V. new perspectives on evolution between bacterial and chloroplast-type ferredoxins inferred from sequence evidence. *Journal of Molecular Evolution*, 29(3), 246-254.
- Quadri, L. E., Weinreb, P. H., Lei, M., Nakano, M. M., Zuber, P., & Walsh, C. T. (1998). Characterization of sfp, a bacillus subtilis phosphopantetheinyl transferase for peptidyl carrier protein domains in peptide synthetases. *Biochemistry*, 37(6), 1585-1595.
- Raymond, K. N., Dertz, E. A., & Kim, S. S. (2003). Enterobactin: An archetype for microbial iron transport. *Proceedings of the National Academy of Sciences of the United States of America*, 100(7), 3584-3588.
- Reuter, K., Mofid, M. R., Marahiel, M. A., & Ficner, R. (1999). Crystal structure of the surfactin synthetase-activating enzyme sfp: A prototype of the 4'-phosphopantetheinyl transferase superfamily. *The EMBO Journal*, 18(23), 6823-6831.
- Roberts, A. A., Copp, J. N., Marahiel, M. A., & Neilan, B. A. (2009). The synechocystis sp. PCC6803 sfp-type phosphopantetheinyl transferase does not possess characteristic broad-range activity. *Chembiochem : A European Journal of Chemical Biology*, 10(11), 1869-1877.
- Salerno, J. C., & Ohnishi, T. (1976). Tetranuclear and binuclear iron-sulfur clusters in succinate dehydrogenase: A method of iron quantitation by formation of paramagnetic complexes. *Biochemical and Biophysical Research Communications*, 73(3), 833-840.
- Sambrook J., Russell D.W. (2001). *Molecular Cloning: A laboratory Manual*. 3rd ed. New York: Cold Springs Harbor Laboratory Press.
- Schuck P. (2000) Size distribution analysis of macromolecules by sedimentation velocity ultracentrifugation and Lamm equation modeling. *Biophysical Journal* (78), 1606-1619.
- Smith, C.A. (1998). Estimation of sedimentation coefficients and frictional ratios of globular proteins. *Biochemical education*, 16(2), 104-106.
- Sridharan, S., Howard, N., Kerbarh, O., Blaszczyk, M., Abell, C., & Blundell, T. L. (2010). Crystal structure of escherichia coli enterobactin-specific isochorismate synthase (EntC) bound to its reaction product isochorismate: Implications for the enzyme mechanism

and differential activity of chorismate-utilizing enzymes. *Journal of Molecular Biology*, 397(1), 290-300.

Staab, J. F., & Earhart, C. F. (1990). EntG activity of escherichia coli enterobactin synthetase. *Journal of Bacteriology*, 172(11), 6403-6410.

Thulasiraman, P., Newton, S. M., Xu, J., Raymond, K. N., Mai, C., Hall, A., *et al.* (1998). Selectivity of ferric enterobactin binding and cooperativity of transport in gram-negative bacteria. *Journal of Bacteriology*, 180(24), 6689-6696.

Vance D.B. (1994). Iron-the environmental impact of a universal element. *National Environmental Journal*, 4(5), 24-25.

Venkatesh, B., Arifuzzaman, M., Mori, H., Suzuki, S., Taguchi, T., & Ohmiya, Y. (2005). Use of GFP tags to monitor localization of different luciferases in E. coli. *Photochemical & Photobiological Sciences : Official Journal of the European Photochemistry Association and the European Society for Photobiology*, 4(9), 740-743.

Vogt, G., Woell, S., & Argos, P. (1997). Protein thermal stability, hydrogen bonds, and ion pairs. *Journal of Molecular Biology*, 269(4), 631-643.

Vogt, G., Woell, S., & Argos, P. (1997). Protein thermal stability, hydrogen bonds, and ion pairs. *Journal of Molecular Biology*, 269(4), 631-643.

Weinglass, A. B., Whitelegge, J. P., Verner, G. E., Faull, K. F., and Kaback, H. R. (2003). Elucidation of substrate binding interactions in a membrane transport protein by mass spectrometry. *EMBO*, (22), 1497-1477.

Wilderman, P. J., Sowa, N. A., FitzGerald, D. J., FitzGerald, P. C., Gottesman, S., Ochsner, U. A., *et al.* (2004). Identification of tandem duplicate regulatory small RNAs in pseudomonas aeruginosa involved in iron homeostasis. *Proceedings of the National Academy of Sciences of the United States of America*, 101(26), 9792-9797.

Xu, Z., Horwich, A. L., & Sigler, P. B. (1997). The crystal structure of the asymmetric GroEL-GroES-(ADP)₇ chaperonin complex. *Nature*, 388(6644), 741-750.

Yin, J., Lin, A. J., Buckett, P. D., Wessling-Resnick, M., Golan, D. E., & Walsh, C. T. (2005). Single-cell FRET imaging of transferrin receptor trafficking dynamics by sfp-catalyzed, site-specific protein labeling. *Chemistry & Biology*, 12(9), 999-1006.

Zeng, X., Xu, F., & Lin, J. (2009). Molecular, antigenic, and functional characteristics of ferric enterobactin receptor CfrA in campylobacter jejuni. *Infection and Immunity*, 77(12), 5437-5448.

Zhou, Z., Lai, J. R., & Walsh, C. T. (2006). Interdomain communication between the thiolation and thioesterase domains of EntF explored by combinatorial mutagenesis and selection. *Chemistry & Biology*, 13(8), 869-879.