

Spring 5-2009

## **Molecular Interactions Involved in the Biogenesis of Bacterial Microcompartments**

Balaraj Balaram Menon  
*University of Southern Mississippi*

Follow this and additional works at: <https://aquila.usm.edu/dissertations>

 Part of the [Chemistry Commons](#)

---

### **Recommended Citation**

Menon, Balaraj Balaram, "Molecular Interactions Involved in the Biogenesis of Bacterial Microcompartments" (2009). *Dissertations*. 1037.  
<https://aquila.usm.edu/dissertations/1037>

This Dissertation is brought to you for free and open access by The Aquila Digital Community. It has been accepted for inclusion in Dissertations by an authorized administrator of The Aquila Digital Community. For more information, please contact [aquilastaff@usm.edu](mailto:aquilastaff@usm.edu).

The University of Southern Mississippi

MOLECULAR INTERACTIONS INVOLVED IN THE BIOGENESIS  
OF BACTERIAL MICROCOMPARTMENTS

by

Balaraj Balaram Menon

Abstract of a Dissertation  
Submitted to the Graduate Studies Office  
of The University of Southern Mississippi  
in Partial Fulfillment of the Requirements  
for the Degree of Doctor of Philosophy

May 2009

COPYRIGHT BY  
BALARAJ BALARAM MENON

2009

The University of Southern Mississippi

MOLECULAR INTERACTIONS INVOLVED IN THE BIOGENESIS  
OF BACTERIAL MICROCOMPARTMENTS

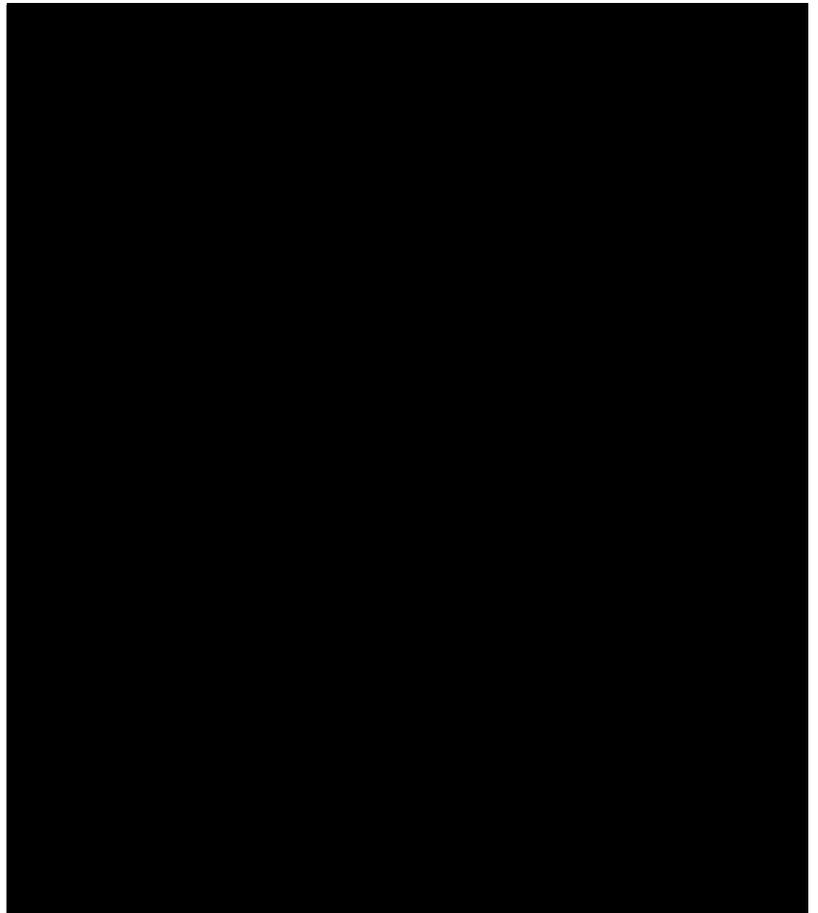
by

Balaraj Balaram Menon

A Dissertation

Submitted to the Graduate Studies Office  
of The University of Southern Mississippi  
in Partial Fulfillment of the Requirements  
for the Degree of Doctor of Philosophy

Approved:



May 2009

## ABSTRACT

# MOLECULAR INTERACTIONS INVOLVED IN THE BIOGENESIS OF BACTERIAL MICROCOMPARTMENTS

by

Balaraj Balaram Menon

May 2009

This study was undertaken with the goal of gaining better insights into the assembly pathway of carboxysomes and related polyhedra. Aside from their similarity in size and shape, all known microcompartments package enzymes that mediate key reactions of metabolic pathways [1]. It remains unclear whether the sequestered enzymes participate in microcompartment assembly and/or contribute to the overall shape of the polyhedra. Genetic studies in *Salmonella enterica* have suggested that the shells of organelles involved in propanediol utilization assemble in the absence of encapsulated enzymes [2]. In this study, *in vivo* yeast two-hybrid screens, involving components of the ethanolamine utilization (Eut) organelles of *S. enterica*, revealed strong interactions between the putative shell proteins – EutN and EutL [3, 4], and the large subunit of the ethanolamine ammonia lyase enzyme. This result suggested that similar interactions between shell components and sequestered enzymes may play an important role in the biogenesis of other microcompartments as well. Since the Eut polyhedra could not be purified and characterized, conclusions drawn from the yeast two-hybrid protein interaction studies remain speculative. A more direct *in vivo* approach for addressing the role of packaged enzymes in

microcompartment assembly was employed by constructing Form IA RubisCO mutants of the model chemoautotrophic bacterium, *Halothiobacillus neapolitanus*, and analyzing their phenotypes. In these mutants, the genes encoding Form IA RubisCO, the enzyme sequestered within carboxysomes, were either deleted or replaced with orthologs from another autotrophic bacterium, *Thiomicrospira crunogena*. Phenotypic characterization studies revealed that a mutant lacking Form IA RubisCO assembled empty carboxysome shells of apparently normal size, shape, and composition. Furthermore, carboxysomes of *H. neapolitanus* readily packaged chimeric and heterologous species of Form IA RubisCO. The large subunit of these foreign RubisCO species was identified as an important determinant of the enzyme's packagability. The impact of these findings on the current understanding of carboxysome architecture and function will be discussed.

*To Papa and Mamma*

## ACKNOWLEDGMENTS

To my parents, who I miss in every walk of life, I dedicate this piece of work. Their unfathomable love, unflinching encouragement, and bountiful blessings resonate throughout all I have done and will ever do.

I am blessed and honored to have Vinitha as my best friend and soul mate. Her confidence in me has been the biggest driving force behind the completion of my doctoral degree. Countless thanks to her for the best graduation gift I could possibly ask for – Akash Balaram Menon, my soon-to-be-born son.

I remain indebted to Drs. Gordon C. Cannon and Sabine Heinhorst, my advisors, for their excellent direction, leadership, patience, and support during my stint with them as a graduate student. Under their tutelage I have learned to do good science and emerge as a better scientist.

I am grateful to my committee members Drs. Robert Bateman Jr., Jeffrey Evans, and Glenmore Shearer for their feedback and advice from time to time.

To current and past members of the Cannon/Heinhorst labs – Martha, Scott, Zhicheng, Fei, Kristina, Jenifer, Nicole, Sarah, Evan, Eric, Steve, Andrea, Daniel, Sara, and Lakeshia, I thank you for your camaraderie during these years. You definitely made life easier at work. Special thanks to Dr. Zhicheng Dou a.k.a. Big Bean for his friendship and scientific inputs and to Martha, the lab mom.

Thanks to Drs. Cheryl A. Kerfeld and Kathleen Scott for scientific discussions and for critically reading our *PLoS ONE* manuscript. Dr. Kathleen Scott provided the *Thiomicrospira crunogena* strain that was used in this research.

Special thanks to Dr. Richard J. Roberts, 1993 Nobel Laureate in Physiology  
or Medicine, for his encouragement and advice during difficult times.

## TABLE OF CONTENTS

ABSTRACT.....	ii
DEDICATION.....	iv
ACKNOWLEDGEMENTS.....	v
LIST OF ILLUSTRATIONS.....	ix
LIST OF TABLES.....	xii
LIST OF ABBREVIATIONS.....	xiii
CHAPTER	
I. INTRODUCTION .....	1
II. LITERATURE REVIEW .....	4
The CO <sub>2</sub> -Concentrating Mechanism	
Carboxysomes – The Prototype Bacterial Microcompartment	
RubisCO and its Multiple Forms	
Protein Interactions and Carboxysome Biogenesis	
Other Polyhedral Microcompartments	
1,2-Propanediol-induced Microcompartments in <i>Salmonella enterica</i>	
Ethanolamine-induced Microcompartments in <i>Salmonella enterica</i>	
Objective of this Study	
III. EXPERIMENTAL PROCEDURES.....	28
Routine Apparatus	
Routine Chemicals	
Media	
Buffers	
Cultures and Growth Conditions	
Methods	
IV. RESULTS .....	62
Protein interactions in the Eut microcompartment of <i>Salmonella enterica</i>	
Evaluation of the role of Form IA RubisCO in carboxysome biogenesis	
Sequestration of foreign Form IA RubisCO within carboxysomes of <i>Halothiobacillus neapolitanus</i>	

V. DISCUSSION .....	135
Interactions between EutB, EutN, and EutL may be central to Eut organelle assembly	
<i>H. neapolitanus</i> assembles empty carboxysomes in the absence of Form IA RubisCO	
The <i>csa</i> -operon encoded copy of Form IA RubisCO is packaged within carboxysomes of <i>Thiomicrospira crunogena</i>	
Carboxysomes of <i>T. crunogena</i> and <i>H. neapolitanus</i> are indistinguishable in morphology but display variations in shell protein composition	
<i>H. neapolitanus</i> carboxysomes accommodate chimeric and heterologous Form IA RubisCO species	
The large subunit, CbbL, determines whether Form IA RubisCO is incorporated into carboxysomes	
VI. CONCLUSIONS AND FUTURE WORK .....	153
APPENDIX.....	157
REFERENCES .....	162

## LIST OF ILLUSTRATIONS

### Figure

1. Carboxysomes – The Prototype Bacterial Microcompartment.....	7
2. Carboxysome shell homologs of the Pdu and Eut microcompartments .....	21
3. Strategy used for constructing the pUC18-Tc-NC-cbbL-kan <sup>r</sup> plasmid .....	47
4. Primary sequence alignment of the CsoS1 and putative Eut shell proteins of <i>H. neapolitanus</i> and <i>S. enterica</i> , respectively.....	64
5. Primary sequence alignment of the CsoS4A and CsoS4B proteins ( <i>H. neapolitanus</i> ) and putative shell protein, EutN ( <i>S. enterica</i> ).....	65
6. Identification of pGADT7- <i>eutB</i> , - <i>C</i> , - <i>K</i> , - <i>L</i> , - <i>M</i> , - <i>N</i> , and - <i>S</i> ‘prey’ constructs via restriction digest analysis .....	68
7. Identification of pGBKT7- <i>eutB</i> , - <i>C</i> , - <i>K</i> , - <i>L</i> , - <i>M</i> , - <i>N</i> , and - <i>S</i> ‘bait’ constructs via restriction digest analysis.....	69
8. Examples of interacting and non-interacting protein pairs.....	70
9. Protein-protein interaction grid of the putative Eut shell proteins and EAL .....	72
10. Purification of recombinant His <sub>6</sub> -tagged Eut - <i>K</i> , - <i>L</i> , - <i>M</i> , - <i>N</i> , and - <i>S</i> proteins.....	75
11. Purification of recombinant His <sub>6</sub> -tagged EutC protein.....	76
12. Western blotting with pre-immune serum and serum containing anti-EutC antibodies .....	77
13. Western blotting with pre-immune serum and serum containing anti-EutC antibodies .....	78
14. EAL activity in extracts of <i>S. enterica</i> cells grown in different carbon sources ...	80
15. Band enriched for EAL separated on a 35-65% w/w sucrose density gradient.....	83
16. SDS-PAGE and western blot analyses of fractions from Eut microcompartment purification steps .....	84
17. EAL activity and densitometry of fractions from 35-65% w/w sucrose gradient .....	85

18. Electron microscopy of the fraction enriched for EAL .....	86
19. Organization of genes within the <i>cso</i> operon of the Form IA RubisCO deletion mutant .....	88
20. Confirmation of <i>H. neapolitanus</i> Form IA RubisCO mutants via PCR amplification of genomic DNA .....	90
21. Growth patterns of wild type and Form IA RubisCO mutants .....	92
22. RubisCO activity in cell extract of the <i>cbbLS::kan<sup>r</sup></i> mutant .....	93
23. Expression of the large subunit (CbbL) in wild type and Form IA RubisCO mutants .....	95
24. Electron micrographs of wild type and mutant carboxysomes .....	98
25. Polypeptide composition of wild type and mutant carboxysomes.....	99
26. Western blot analysis of wild type and mutant carboxysomes probed with anti-CbbL and anti-CsoS1 antibodies .....	100
27. Western blot analysis of wild type and mutant carboxysomes probed with anti-CbbM (Form II RubisCO) antibodies.....	102
28. Flowchart for purification of carboxysomes from <i>T. crunogena</i> .....	104
29. SDS-PAGE analysis of fractions obtained during the purification of <i>T. crunogena</i> carboxysomes.....	105
30. Transmission electron micrographs of negatively stained carboxysomes from <i>H. neapolitanus</i> and <i>T. crunogena</i> .....	106
31. Primary sequence alignment of Form IA RubisCO large subunits.....	109
32. Primary sequence alignment of Form IA RubisCO small subunits.....	110
33. Identification of the <i>cso</i> operon-encoded RubisCO large subunit in carboxysomes of <i>T. crunogena</i> .....	111
34. Identification of the <i>cso</i> operon-encoded RubisCO small subunit in carboxysomes of <i>T. crunogena</i> .....	112
35. Separation of <i>H. neapolitanus</i> and <i>T. crunogena</i> carboxysome proteins via SDS-PAGE .....	116

36. Western blotting analyses of <i>H. neapolitanus</i> and <i>T. crunogena</i> carboxysomes	118
37. Organization of genes within the <i>csa</i> operon of the Form IA RubisCO replacement mutants .....	121
38. RubisCO activity in cell extract of the <i>cbbLS::Tc NC cbbLS</i> mutant.....	124
39. RubisCO activity in cell extract of the <i>cbbL::Tc NC cbbL</i> mutant.....	125
40. Identification of <i>T. crunogena</i> carboxysomal large subunit (CbbL) in carboxysomes of the <i>cbbLS::Tc C cbbLS</i> mutant via mass spectrometry .....	129
41. Identification of <i>T. crunogena</i> carboxysomal small subunit (CbbS) in carboxysomes of the <i>cbbLS::Tc C cbbLS</i> mutant via mass spectrometry .....	130
42. Identification of <i>T. crunogena</i> noncarboxysomal small subunit (CbbS) in carboxysomes of the <i>cbbS::Tc NC cbbS</i> mutant via mass spectrometry .....	131

## LIST OF TABLES

### Table

1. Stoichiometric distribution of carboxysomal components in *H. neapolitanus* and *T. crunogena* .....117
2. CO<sub>2</sub> fixation activities of carboxysomes and free RubisCO released from carboxysomes.....134

## LIST OF ABBREVIATIONS

$\beta$ -ME	2-mercaptoethanol
BCA	Bicinchoninic acid
BCIP	5-bromo-4-chloro-3-indolyl phosphate
Bicine	N,N-Bis(2-hydroxyethyl)glycine
BMC	Bacterial microcompartment
bp	Base pair
CA	Carbonic anhydrase
CCM	Carbon dioxide concentrating mechanism
CBB	Calvin-Benson-Bassham
C <sub>i</sub>	Inorganic carbon
dNTP	deoxynucleotide triphosphate
DTT	Dithiothreitol
EA	ethanolamine
Eut	Ethanolamine utilization
<i>hcr</i>	High CO <sub>2</sub> -requiring
HEPES	4-(2-hydroxyethyl)-1-piperazine ethane sulfonic acid
HRP	Horseradish peroxidase
IgG	Immunoglobulin G
IPTG	Isopropyl- $\beta$ -D-thiogalactoside
MALDI-ToF	Matrix Assisted Laser Desorption /Ionization – Time of Flight
NTA	Nitrilotriacetic acid
OD	Optical density
PAGE	Polyacrylamide gel electrophoresis
PBS	Phosphate-buffered saline
PCR	Polymerase chain reaction
PD	Propanediol
Pdu	Propanediol utilization
PG	2-phosphoglycolate
PGA	3-phosphoglycerate
PMSF	Phenylmethylsulfonyl fluoride
PTSF	p-toluenesulfonyl fluoride
RubisCO	Ribulose 1,5-bisphosphate carboxylase/oxygenase
RuBP	Ribulose 1,5-bisphosphate
SDS	Sodium dodecyl sulfate
SPA	Sequential peptide affinity
TEM	Transmission electron microscopy

## CHAPTER I

### INTRODUCTION

Until recently, it was thought that organized subcellular interiors existed only inside eukaryotic cells. However, emerging research suggests that this is not the case. Recent studies have identified foci and hubs within prokaryotic cells that mediate important cellular and metabolic processes [5-8]. One well studied example of organization within prokaryotic cells is the CO<sub>2</sub>-concentrating mechanism (CCM) found in cyanobacteria and many chemoautotrophic bacteria. The CCM in these microbes begins with the cytosolic import of inorganic carbon (C<sub>i</sub>) by transporters, followed by the fixation of C<sub>i</sub> into organic metabolites by the enzyme ribulose 1,5-bisphosphate carboxylase/oxygenase (RubisCO, EC 4.1.1.39). The latter process of C<sub>i</sub> fixation occurs within a highly structured proteinaceous organelle called the carboxysome [9].

The chemoautotrophic bacterium, *H. neapolitanus*, has served as the model organism for studying the architectural and functional aspects of carboxysomes. The molecular arrangement of RubisCO, carbonic anhydrase, and shell components within the carboxysome ensures efficient CO<sub>2</sub> fixation. The working model for carboxysome function suggests that cytosolic bicarbonate is first converted to CO<sub>2</sub> at the shell interface by the shell-associated carbonic anhydrase. This CO<sub>2</sub> is then converted by the Form IA RubisCO enzyme sequestered within the carboxysome to two molecules of 3-phosphoglycerate (3PGA), which is used by the cell to build organic biomass. Recent research in *H. neapolitanus* suggests that the shell limits the

bidirectional diffusion of CO<sub>2</sub> in and out of the carboxysome [10]. Another likely function of the shell might be the preclusion of molecular O<sub>2</sub> from entering the carboxysome and competing with CO<sub>2</sub> for the active site of Form IA RubisCO.

Crystallographic studies of the shell components CsoS1A and CsoS4A from *H. neapolitanus* have provided vital clues about the external architecture of carboxysomes [11-13]. In spite of such significant advancements, fundamental issues concerning the biogenesis of the microcompartment and the involvement of Form IA RubisCO in the assembly pathway remain unresolved. Previous *in vitro* protein interaction studies have suggested that Form IA RubisCO may be an obligatory component of carboxysome assembly [14]. At the same time, electron microscopy studies of a *H. neapolitanus* mutant lacking the large subunit of Form IA RubisCO revealed the existence of polyhedral structures that were empty and smaller in size compared to wild type carboxysomes [15]. However, this study did not ascertain if the empty polyhedra were indeed carboxysome shells [15].

Certain chemoautotrophic bacteria such as *Hydrogenovibrio marinus*, *Acidithiobacillus ferrooxidans*, and the recently sequenced deep sea vent  $\gamma$ -proteobacterium *Thiomicrospira crunogena* possess duplicate copies of Form IA RubisCO-encoding genes [16-18]. In *Hy. marinus* and *T. crunogena*, one copy resides within the *cso* operon while the other is located elsewhere in the genome. Although the large subunits of both copies are nearly 80% identical in their amino acid sequence [17], which copy is incorporated within carboxysomes of these bacteria is not known.

Intriguingly, carboxysome-like microcompartments, also referred to as metabolosomes, have been identified in chemoheterotrophic bacteria such as *Salmonella enterica*, *Escherichia coli*, and *Lactobacillus reuteri* during growth on substrates like 1,2-propanediol (1,2-PD) and ethanolamine (EA) [2, 19-21]. The presumed structural components of such microcompartments share a high degree of homology with the shell proteins of carboxysomes. These polyhedra, like carboxysomes, tend to sequester specific enzymes within their interior that catalyze key reactions of metabolic pathways. Microcompartments involved in the catabolism of 1,2-PD and EA sequester the enzymes PD dehydratase and EA ammonia lyase, respectively. While genetic studies in *S. enterica* have suggested that shells of the PD utilization (Pdu) organelles assemble in the absence of the PD dehydratase enzyme, such studies with regard to the Eut organelles have remained inconclusive. Thus, in the case of all known bacterial microcompartments, it remains unclear whether the sequestered enzymes contribute to their biogenesis and overall polyhedral shape. This study was thus designed with the goal of assessing whether encapsulated enzymes participate in the assembly pathway and/or contribute to the overall organization of their respective microcompartments.

## CHAPTER II

### LITERATURE REVIEW

#### The CO<sub>2</sub>-Concentrating Mechanism

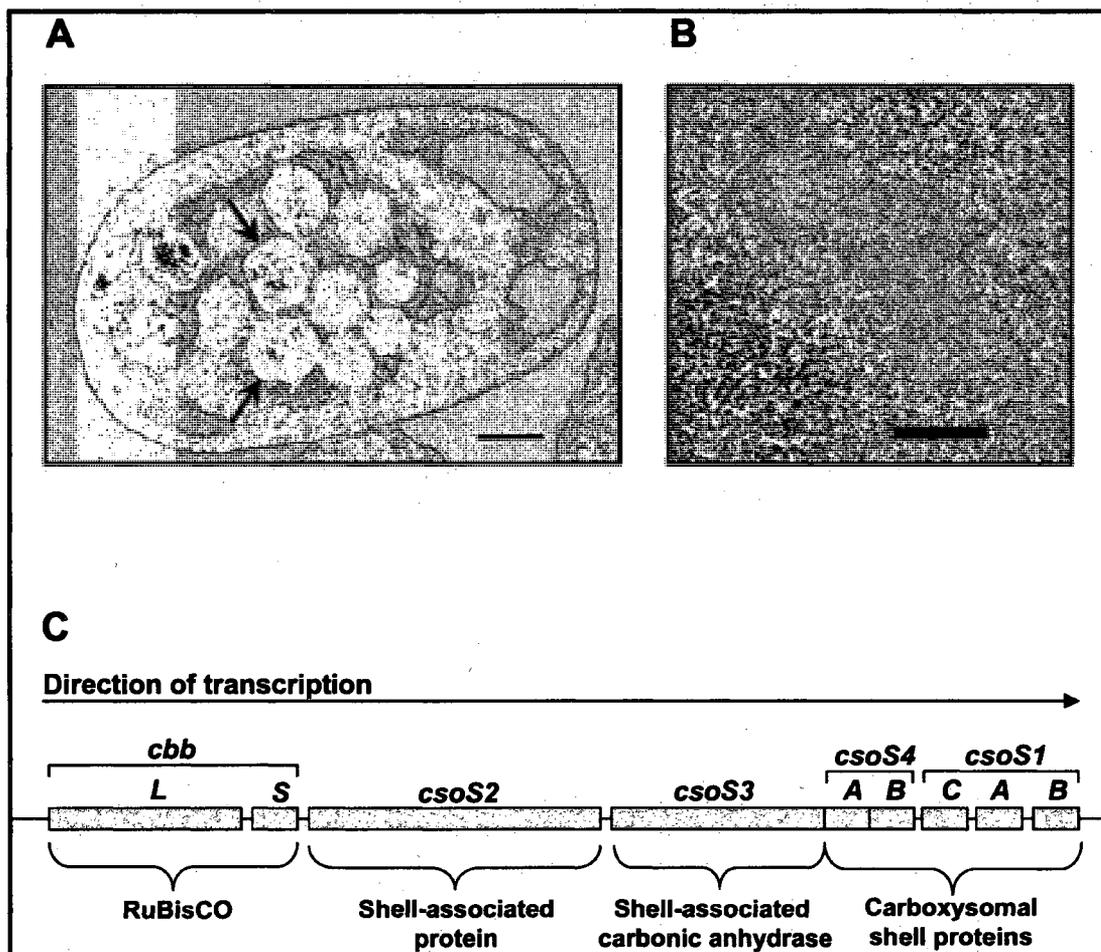
Modern biology has revolutionized the long-held notion that complex intracellular domains exist only within eukaryotic cells. High resolution microscopic studies of prokaryotic cells suggest that proteins and small molecules are not randomly distributed within the cytoplasm [22]. In fact, recent reports have revealed that many proteins involved in processes like cell division, DNA replication and chromosome segregation, transcription and translation, and metabolic pathways are localized at specific foci within the bacterial cell [5-8, 23]. Logistically, having such an organized subcellular architecture provides cells with advantages such as energy conservation and optimization of various processes. One well studied example of intracellular organization is the CO<sub>2</sub>-concentrating mechanism (CCM) observed in many chemo- and photoautotrophic bacteria [9, 24]. The CCM begins with the cytoplasmic import of inorganic carbon (C<sub>i</sub>), mostly bicarbonate, via energy dependent transmembrane transport. This step is followed by utilization of the intracellular C<sub>i</sub> pool thereby generating elevated concentrations of CO<sub>2</sub> within a proteinaceous microcompartment called the carboxysome that is filled with the CO<sub>2</sub>-fixing enzyme ribulose 1,5-bisphosphate carboxylase/oxygenase (RubisCO, EC 4.1.1.39) [9, 25].

Insights into transport-mediated uptake of C<sub>i</sub> and CCMs have mainly arisen from research on cyanobacterial species [24, 25] and data addressing C<sub>i</sub> uptake

mechanisms in chemoautotrophic bacteria is limited. In cyanobacteria,  $C_i$  uptake is mediated by families of  $HCO_3^-$  transporters or dehydrogenase complexes with an increased affinity for  $CO_2$  [26-28]. A recent study of the CCM in the deep sea vent chemoautotrophic proteobacterium, *Thiomicrospira crunogena*, suggested that  $C_i$  uptake is mediated by a membrane potential or an energy-dependent bicarbonate transport mechanism. No apparent homologs of cyanobacterial  $HCO_3^-$  transporters or  $CO_2$  uptake systems were identified in this bacterium. Intriguingly, a gene encoding a SulP-type anion transporter, which was previously found to be associated with  $HCO_3^-$  transporting activity in marine cyanobacteria, was identified in the genome of *T. crunogena* [17, 27, 29]. Irrespective of whether  $HCO_3^-$  or  $CO_2$  is brought into the cell,  $HCO_3^-$  remains the main  $C_i$  species that accumulates in the cytoplasm and reaches concentrations that are 100 to 1000 fold higher than extracellular levels [24, 29]. The absence of a cytoplasmic carbonic anhydrase (CA) ensures that intracellular  $HCO_3^-$  is maintained at higher levels than  $CO_2$ . Expression of human CA in the cytoplasm of the  $\beta$ -cyanobacterium *Synechococcus* PCC 7942 was reported to result in equilibration of  $HCO_3^-$  and  $CO_2$  and subsequent leakage of the latter from the cells. Furthermore, cells expressing the human CA were found to exhibit a high  $CO_2$ -requiring (*hcr*) phenotype [25, 30, 31].

The high concentration of  $HCO_3^-$  that is generated within the cytoplasm needs to be converted to  $CO_2$ , which is the substrate for RubisCO. This conversion process is part of terminal step of the CCM [32]. A highly structured polyhedral microcompartment, called the carboxysome (**Figure 1**), which packages most, if not all, of the cell's RubisCO, also contains a shell-associated CA that aids in the

conversion of  $\text{HCO}_3^-$  to  $\text{CO}_2$  at the microcompartment's shell interface. The local concentration of  $\text{CO}_2$  inside the carboxysome is thereby increased. This mechanism ensures efficient  $\text{CO}_2$  fixation by the encapsulated RubisCO enzyme [24, 33, 34]. In most cyanobacteria, the CA that is thought to mediate the conversion of  $\text{HCO}_3^-$  to  $\text{CO}_2$  within the carboxysome is encoded by the *ccaA/icfA* gene, while in chemoautotrophic bacteria it is encoded by the *csO* operon-associated *csoS3* gene (Figure 1) [35, 36].



**Figure 1. Carboxysomes – the prototype bacterial microcompartment**

[A] Thin section of a *H. neapolitanus* cell harboring multiple carboxysomes indicated by arrows and [B] negatively stained carboxysomes purified from *H. neapolitanus*. Bars represent a scale of 100 nm. [C] The *H. neapolitanus* *cso* operon harbors the genes for RubisCO (*cbbL*, *cbbS*) and the carboxysomal shell and shell-associated proteins (*csoS2*, *csoS3*, *csoS4A*, *csoS4B*, *csoS1C*, *csoS1A*, *csoS1B*).

## Carboxysomes – The Prototype Bacterial Microcompartment

Polyhedral inclusion bodies in bacteria were discovered nearly fifty years ago in the cyanobacterium *Phormidium uncinatum* [1]. Subsequent studies identified such polyhedra in other cyanobacteria and chemoautotrophic bacteria [37, 38]. The inclusions were first purified and biochemically characterized from the chemoautotrophic bacterium *Halothiobacillus neapolitanus*. It was noted that they consist of a shell that is filled with numerous molecules of the CO<sub>2</sub>-fixing enzyme ribulose 1,5-bisphosphate carboxylase/oxxygenase (RubisCO) [39]. This finding led to the coining of the term ‘carboxysomes’ to describe the polyhedra [39]. Visualization of purified carboxysomes by electron microscopy revealed that they are hexagonal in shape with an average diameter of 100 nm [40]. Some preparations of carboxysomes were found to contain the characteristic ‘donut’ shaped RubisCO molecules arranged in paracrystalline arrays. Holthuijzen *et al.* [41] observed the arrangement of RubisCO molecules in a layer adjoining the inside of the shell and proposed that the interior of the carboxysome might be hollow. However, the separation of purified carboxysome preparations by sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) revealed that RubisCO accounts for nearly 60% of carboxysomal protein [42]. Furthermore, size-based calculations had predicted that the carboxysome could accommodate up to 300 molecules of RubisCO [38]. Recent high resolution cryo-electron tomography studies have revealed that RubisCO molecules occupy almost the entire internal volume of carboxysomes [43, 44]. These observations indicate that the interior of carboxysomes are not hollow as proposed by Holthuijzen *et al.* [41]

Carboxysomes are classified as  $\alpha$ -type or  $\beta$ -type based on their protein composition and the arrangement of their genes [45]. The components of the  $\alpha$ -type carboxysomes are encoded by a cluster of genes that is referred to as the carboxysome (*cso*) operon (**Figure 1**). The  $\alpha$ -type carboxysomes are typically found in chemoautotrophic bacteria and certain  $\alpha$ -cyanobacteria such as *Prochlorococcus* and certain marine *Synechococcus* species. On the other hand, genes encoding the putative components of the  $\beta$ -type carboxysomes are organized in multiple clusters. The  $\beta$ -type carboxysomes are found in  $\beta$ -cyanobacteria such as *Synechococcus elongatus* PCC 7942 and *Synechocystis* sp. PCC 6803. Since  $\beta$ -type microcompartments from cyanobacteria have not yet been purified to homogeneity, their exact composition is not known. Hence, studies pertaining to carboxysome structure and function have mainly focused on the  $\alpha$ -type microcompartments of *H. neapolitanus* [9].

It is now known that the  $\alpha$ -carboxysomes of *H. neapolitanus* are composed of ten polypeptides encoded by the nine genes of the *cso* operon (**Figure 1**). The major structural components of the carboxysome shell are the CsoS1C, CsoS1A, and CsoS1B proteins. A *H. neapolitanus* mutant lacking the *csoS1A* gene was found to contain a reduced number of carboxysomes and exhibit the *hcr* phenotype [46]. The CsoS1 proteins, which together account for nearly 15% of carboxysomal protein, are nearly identical to one another and are believed to be products of gene duplication events [1]. In addition, these proteins share significant sequence homology with the CcmK and CcmO proteins that were previously shown via genetic studies to be important for  $\beta$ -carboxysome formation [47]. The CsoS1 and CcmK proteins have subsequently

been classified into the bacterial microcompartment (BMC) family of proteins (Pfam number: PF00936, <http://pfam.sanger.ac.uk/family/bmc>). The crystal structures of CsoS1A from *H. neapolitanus* and CcmK1, CcmK2, and CcmK4 from *Synechocystis* sp. PCC 6803 were recently solved [11, 48, 49]. Interestingly, all of these shell proteins were shown to form hexamers that are comprised of six identical subunits packed cyclically around a central axis of symmetry [1]. This finding led to the speculation that the carboxysome shell is mainly assembled from hexameric building blocks. The observation that individual hexamers of CsoS1A, CcmK1, and CcmK2 associate tightly in crystals to form two dimensional layers led to the hypothesis that such layers form the flat surfaces of the carboxysome shell [11, 48]. However, it remains unknown if such layers are composed of identical or mixed hexamers *in vivo*. Intriguingly, the hexameric structures of BMC proteins contain a central pore ranging from 4 to 7 Å in diameter. These pores bear a positive electrostatic potential, possibly due to lysine and arginine residues, that may serve as a conduit for the transport of substrates and products across the carboxysome shell as proposed by Yeates *et al.* [1].

The minor components of the *H. neapolitanus* carboxysome shell are encoded by the *csoS4A* and *csoS4B* genes, previously known as *orfA* and *orfB*, respectively. The products of these genes are expressed in low copies, which would explain why they escaped detection in purified carboxysome preparations previously [50, 51]. The CsoS4A and CsoS4B proteins are homologous to the CcmL protein of  $\beta$ -cyanobacteria. Intriguingly, a *Synechococcus* sp. PCC 7942 mutant lacking the *ccmL* gene was found to display an *hcr* phenotype and harbor elongated carboxysomes [30].

The effect of deleting the CsoS4 genes on the phenotype of *H. neapolitanus* is

currently being investigated. Recent structural data has shown that CcmL and CsoS4A proteins form pentamers composed of five identical subunits [13]. Such pentamers are thought to occupy the vertices of the icosahedral carboxysome shell and provide the curvature needed to close the flat hexagonal surfaces into a three-dimensional microcompartment [13].

The *cso* operon of *H. neapolitanus* harbors two additional shell genes, *csoS2* and *csoS3*. The *csoS2* gene encodes two polypeptides, CsoS2B and CsoS2A of 130 kDa and 85 kDa, respectively, that together represent 12% of carboxysomal protein [7]. That the CsoS2B and CsoS2A polypeptides are shell-associated was shown by immunogold electron microscopy [52]. Considering that the N-termini of the CsoS2B and CsoS2A polypeptides are identical, the observed difference in their molecular mass has been attributed to differential levels of posttranslational glycosylation [52]. A recent study showed that CsoS2A is a C-terminally truncated form of the CsoS2B polypeptide [53]. However, the exact site of cleavage within the CsoS2 protein could not be identified via mass spectrometric analysis. Due to its high isoelectric point ( $pI > 9$ ), it has been proposed that the positively charged CsoS2 protein may also play a role in the diffusion of negatively charged substrates such as bicarbonate, 3-phosphoglycerate, and ribulose 1,5-bisphosphate across the carboxysome shell [53]. The *csoS3* gene was recently found to encode a carbonic anhydrase (CsoS3 or CsoSCA). Because of lack of sequence homology with known carbonic anhydrases, it was originally assigned to a new  $\epsilon$  class of CAs [36]. However, crystallographic studies revealed that CsoSCA forms a subclass of the  $\beta$ -carbonic anhydrases [54]. It is estimated that the functional dimeric form of CsoSCA is present in nearly forty

copies per carboxysome [10]. CsoSCA is tightly associated with the carboxysome shell and is released only when the shell is disassembled.

The majority of the carboxysomal protein is composed of RubisCO [42]. The *cbbL* and *cbbS* genes of the *cso* operon encode the large and small subunit of the RubisCO holoenzyme, respectively [55]. In  $\beta$ -cyanobacteria, the *rbcL* and *rbcS* genes encode the equivalents of the *cbbL* and *cbbS* genes [56]. Eight large ( $L_8$ ) and small ( $S_8$ ) subunits assemble to form a hexadecameric RubisCO holoenzyme.

Despite such advancements in the understanding of carboxysome structure and function, some fundamental questions remain unanswered. The molecular events that trigger the assembly of carboxysomes and the network of protein interactions that lays the foundation for the assembly process remain unknown. This study is aimed at addressing whether interactions between RubisCO and the shell proteins determine the assembly and/or architecture of the carboxysome.

### **RubisCO and its Multiple Forms**

To date, five different metabolic routes have been identified through which autotrophs assimilate  $\text{CO}_2$  into biomass. These routes are 1) the Calvin cycle, also known as the Calvin-Benson-Bassham (CBB) reductive pentose phosphate pathway, 2) the reductive citric acid cycle, 3) the reductive acetyl-CoA pathway, 4) the 3-hydroxypropionate/malyl CoA cycle, and 5) the recently discovered 3-hydroxypropionate/4-hydroxybutyrate cycle prevalent in the archaea [57]. Of these pathways, cyanobacteria and aerobic proteobacteria employ the CBB cycle to fix  $\text{CO}_2$  [57, 58]. The key enzyme of this cycle is RubisCO, which is nature's most abundant

enzyme. RubisCO catalyzes the addition of  $\text{CO}_2$  onto the enediol form of ribulose 1,5-bisphosphate (RuBP), as part of a carboxylase reaction, to form a highly unstable six-carbon phosphorylated intermediate known as 3-keto-2-carboxyarabinitol 1,5-bisphosphate, which is converted instantaneously to two molecules of 3-phosphoglycerate (3-PGA). Oxygen competes with  $\text{CO}_2$  and can combine with the enediol form of RuBP, as part of an oxygenase reaction, to yield one molecule each of 3-PGA and 2-phosphoglycolate (2-PG) [59]. The carboxylation reaction is beneficial as it releases 3-PGA for build-up of sugars like glucose. The oxygenation reaction, on the other hand, is an energy wasteful process since 2-PG has to be excreted as otherwise toxic glycolate waste or converted to 3-PGA for re-entry into the CBB cycle.. Plants convert the 2-PG into RuBP via photorespiration, which consumes ATP [59]. The inability of RubisCO to discriminate between its substrates  $\text{CO}_2$  and  $\text{O}_2$  along with its low  $K_m$  and  $V_{\max}$  values makes it a catalytically imperfect enzyme.

Four different forms of RubisCO, Forms I, II, III, and IV, have been identified in nature [60]. These classifications have been made based on differences in amino acid sequence of the enzyme's catalytic subunit and in holoenzyme structure. Form I RubisCO, which is the most abundant form, is widespread in eukaryotes and in autotrophic bacteria. The structural unit is comprised of four dimers of eight large and eight small subunits in an  $(L_2)_4(S_2)_4$  conformation [59]. Form II RubisCO shares only 25-30% identity in primary structure with the large subunit of the Form I enzyme. Structurally, Form II RubisCO resembles the basic  $L_2$  dimer or multimers thereof of the Form I enzyme; however both forms have distinct catalytic properties.

From a historical perspective, Form II RubisCO found in the photosynthetic anaerobic bacterium *Rhodospirillum rubrum* was the first RubisCO structure to be solved [60]. In addition, it has served as a model for elucidating several structural and functional characteristics of RubisCO in general. This form is found in eukaryotes, some autotrophic bacteria and some archaea. Data mining of several recently sequenced genomes of diverse prokaryotes led to the identification of a third form of RubisCO, Form III, in methanogens and thermophilic archaea. For the first time, RubisCO-encoding genes were discovered in bacteria that inhabit O<sub>2</sub>-free environments. Although this form has been shown to possess *bona fide* RubisCO activity, its function in these organisms seems unclear, especially since CO<sub>2</sub> is fixed via a modified acetyl CoA pathway rather than the CBB cycle in these organisms. In addition, Form III RubisCO has been implicated in AMP metabolism [61]. The most recently discovered form of RubisCO, Form IV, is encoded by genes that are distinct from those encoding the other forms. Since members of this protein family lack RubisCO activity, they have been referred to as RubisCO-like proteins (RLPs). Their inability to catalyze CO<sub>2</sub> fixation has been attributed to substitutions of crucial active site residues. Nevertheless, the crystal structure of an RLP from the green sulfur bacterium *Chlorobium tepidum* revealed that it was a dimer closely resembling the L<sub>2</sub> structures of Forms II and III [62]. This form has been found in eukaryotes, auto- and heterotrophic bacteria, and in archaea and has been implicated in sulfur metabolism [63].

Based on primary sequence diversification and occurrence, Form I RubisCO has been divided into four different clades. Forms IA, and IB form the 'green' group

and are prevalent in proteobacteria, cyanobacteria, green algae, and higher plants; Forms IC and ID, which represent the 'red' group, occur in proteobacteria and non-green algae [56]. Form IA has undergone a further sub-classification into Form IAc (carboxysomal) and Form IAq (noncarboxysomal) based on gene organization and differences in small subunit sequences. Similarly, Form IB contains the Form IBc (carboxysomal) that is populated by  $\beta$ -carboxysomal cyanobacterial enzymes. It is interesting to note that certain proteobacteria possess genes encoding up to three forms of RubisCO belonging to Form IAc and/or IAq, IC, and II, while cyanobacteria contain genes encoding either Form IAc or IB/IBc. The absence of Form IBc in proteobacteria may suggest that this form evolved more recently from Form IA [56]. In the proteobacteria *Thiomicrospira crunogena*, *Hydrogenovibrio marinus*, and *Acidithiobacillus ferrooxidans* the organization of the Forms IAc and IAq RubisCO-encoding genes is similar. The Form IAc gene set is embedded within an  $\alpha$ -type *cso* operon, while the Form IAq gene set is located upstream of a *cbbQ/cbbO* gene cluster. The CbbQ/CbbO proteins are thought to encode proteins involved in RubisCO assembly. The Form II RubisCO-encoding gene *cbbM* of *A. ferrooxidans* and *Hy. marinus* is located upstream of another *cbbQ/cbbO* gene cluster, which is in turn followed by a  $\beta$ -class carbonic anhydrase-encoding gene. The *T. crunogena* Form IAq and Form II RubisCO-encoding genes are juxtaposed and divergently transcribed [17].

To date, no study has provided direct evidence of whether carboxysomes of *T. crunogena*, *Hy. marinus*, and *A. ferrooxidans* accommodate Form IAc and/or IAq RubisCO. A study in *Hy. marinus* had suggested a strong correlation between the

expression of Form IAc RubisCO and the presence of carboxysomes [16]. Expression profiles of the three forms indicated that only Form II RubisCO is expressed at high CO<sub>2</sub> concentrations (15%). At CO<sub>2</sub> levels of 2%, Forms IAq and II RubisCO are expressed, while at 0.15% CO<sub>2</sub> concentrations, Form IAc is the predominant enzyme. All three forms are expressed at similar levels at ambient CO<sub>2</sub>. However, one drawback of this study is that the formation of carboxysomes was not evaluated at ambient CO<sub>2</sub> levels, the condition under which all forms of RubisCO are expressed and one would expect to observe the largest number of microcompartments [64, 65]. Thus, it remains inconclusive if carboxysomes of *Hy. marinus* package Forms IAc or IAq RubisCO.

Primary sequence alignment of Form IA RubisCO large subunits from autotrophic bacteria revealed that they are nearly 80% identical to one another. However, their small subunits display considerable divergence. Badger and Bek [56] made an important observation with respect to the N-terminus of some RubisCO small subunit sequences. The N-terminus of the Form IAq (noncarboxysomal) RubisCO small subunits harbors a six amino acid insertion that is missing in the Form IAc (carboxysomal) CbbS species. Interestingly, this motif is also absent in the cyanobacterial Form IBc RubisCO. Structure-based predictions suggested that the six amino acid N-terminal insertion may be located at the surface of the RubisCO holoenzyme. Based on these results, the inserted amino acids are thought to possibly interfere with protein interactions that may be crucial to the assembly of carboxysomes. This assumption is reasonable considering that carboxysomes have not yet been observed in the chemoautotrophic bacterium *Thiobacillus denitrificans*,

which harbors only genes for Form IAq (noncarboxysomal) and Form II RubisCO. Nevertheless, it remains to be determined experimentally if the insertion motif found in the Form IAq RubisCO small subunit species influences sequestration of the holoenzyme into carboxysomes and/or interferes with microcompartment assembly.

### **Protein Interactions and Carboxysome Biogenesis**

Earlier electron microscopy studies had reported that RubisCO molecules within carboxysomes are packed tightly in paracrystalline arrays [40, 41]. These studies suggested possible interactions between RubisCO and the carboxysomal shell components. Holthuijzen *et al.* [66] reported that the small subunit of RubisCO associates tightly with the carboxysome shell of *H. neapolitanus* after mechanical disruption of intact carboxysomes. A yeast two-hybrid screen of *H. neapolitanus* carboxysomal proteins indicated strong interactions between the RubisCO small subunit and the CsoS1B and CsoS4A shell proteins [14]. Furthermore, disruption of carboxysomes by freeze/thaw treatment [36, 67] revealed a significant portion of RubisCO to co-purify with the shell-enriched fraction, likewise suggesting a strong interaction between RubisCO and the shell proteins [14]. On the other hand, recent high resolution cryo-electron tomography studies of purified carboxysomes from the  $\alpha$ -cyanobacterium *Synechococcus* sp. strain WH8102 indicated that the RubisCO molecules that were arranged in three to four concentric layers within the microcompartment did not appear to make any contact with the shell component [43].

To address the assembly pathway of carboxysomes, two models have been proposed. Price and Badger [32] made observations with regards to the different

stages of carboxysome maturation in thin sections of the filamentous cyanobacterium *Anabaena variabilis* M3 and expressed the view that a carboxysome shell is assembled first and that RubisCO and carbonic anhydrase are inserted through the existing shell. An alternate cascade of events for assembly was suggested by Orús *et al.* [68]. Since, even the earliest carboxysome precursor stages visible in thin sections of the  $\beta$ -cyanobacterium *Synechococcus* PCC 7942 contain regular arrays of RubisCO, these researchers proposed that the carboxysome shell assembles around a preassembled RubisCO core [68]. *In vitro* reconstitution studies involving recombinant shell proteins and purified endogenous RubisCO enzyme indicated that the latter was needed for carboxysome-like polyhedra to assemble. These results likewise suggested a pathway for carboxysome assembly that is dependent on RubisCO [14]. The most direct way to test if RubisCO is involved in the biogenesis of carboxysomes is to generate a mutant in which the RubisCO large and small subunit-encoding genes are deleted and evaluate the mutant's ability to form carboxysomes. Baker *et al.* [15] constructed a *H. neapolitanus* mutant in which the RubisCO-large subunit gene, *cbbL* gene of the *csa* operon, was replaced with a kanamycin cassette. The resulting *cbbL::Km* mutant displayed an *hcr* phenotype and thin sections of this mutant revealed empty polyhedral structures that were smaller in size than wild type carboxysome shells. This study did not characterize the polyhedra observed in the mutant. Thus, it remains inconclusive if the inclusions observed were indeed carboxysome shells.

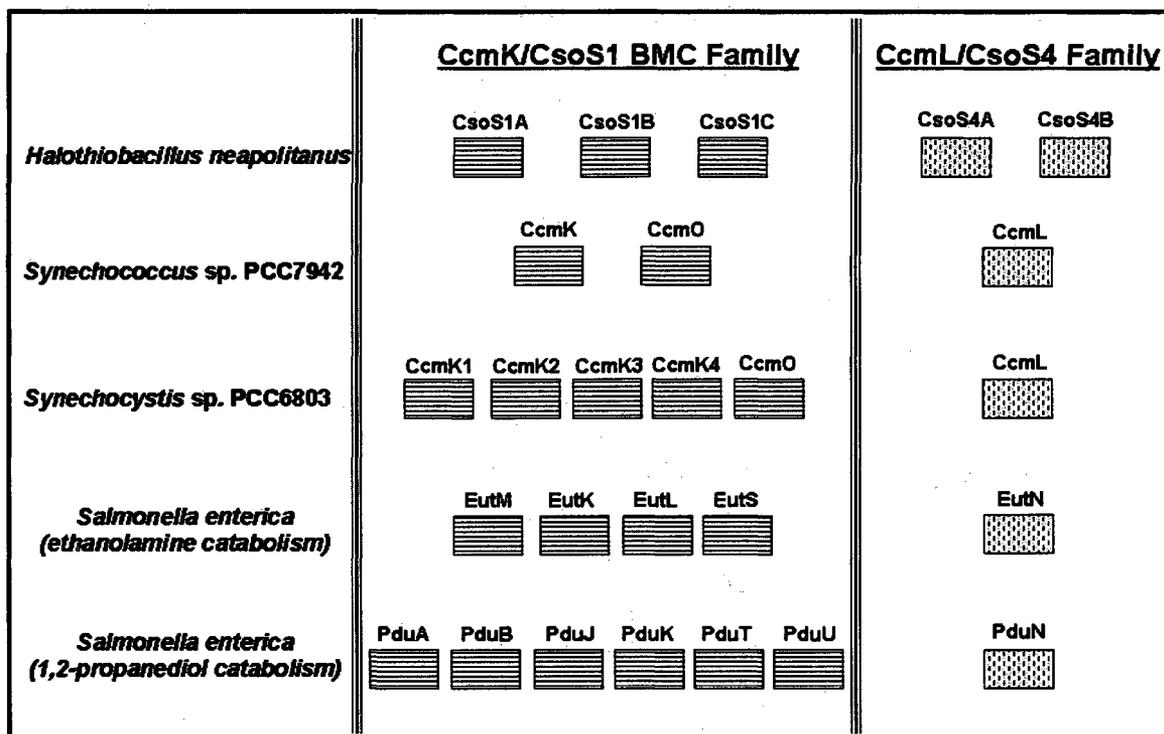
Recent protein interaction studies identified subcomplexes that may be involved in the assembly of the carboxysome. Affinity chromatography-based

experiments involving presumed *Synechococcus* sp. PCC 7942  $\beta$ -carboxysomal proteins identified two major complexes, one between the carbonic anhydrase CcaA, the putative shell protein CcmM and RubisCO and another between a truncated form of CcmM and RubisCO [69]. It was thought that such subcomplexes may be involved in forming a scaffold for other shell proteins to assemble. Similar results implicating CcmM as an important component of carboxysome biogenesis in the  $\beta$ -cyanobacterium *Synechocystis* sp. PCC 6803 were reported by Cot *et al.* [70]. Interestingly, the CcmM protein harbors three to four small subunit repeats at its C-terminus and a domain commonly found in carbonic anhydrases and acetyltransferases at its N-terminus. A *Synechococcus* sp. PCC 7942 mutant lacking CcmM was shown to lack carboxysomes and display an *hcr* phenotype [71]. However, since  $\beta$ -carboxysomes have not been purified to homogeneity, their composition and the stoichiometry of individual components remain unclear. Thus, observations from *in vitro* protein-protein interactions remain purely speculative and are of limited value in understanding the biogenesis of carboxysomes.

### **Other Polyhedral Microcompartments**

The next best studied polyhedral microcompartments are ones involved in the 1,2-propanediol (1,2-PD) and ethanolamine (EA) catabolism pathways in *Salmonella enterica* [72]. In addition to possessing a similar size and shape as carboxysomes, their putative shell proteins share a high degree of homology with carboxysomal shell subunits (**Figure 2**) [3]. However, the physiological functions of these microcompartments are different from that of carboxysomes, since *S. enterica* cells

lack the apparatus to assimilate inorganic carbon. One model suggested by Havemann *et al.* [19] predicts that the 1,2-PD utilization (Pdu) and EA utilization (Eut) organelles aid in minimizing the intracellular levels of toxic propionaldehyde and acetaldehyde, which are products released from 1,2-PD and EA catabolism, respectively. A newer model proposed by Penrod and Roth [73] suggests that the Pdu and Eut microcompartments serve to conserve the volatile aldehydes by providing a low pH environment. Clearly, the physiological function of microcompartments in chemoheterotrophs remains debatable.



**Figure 2. Carboxysome shell homologs of the Pdu and Eut microcompartments**

Putative shell proteins of the Pdu and Eut organelles of *Salmonella enterica* bear resemblance to the carboxysomal shell components at the primary sequence level. Based on homology patterns, the Pdu and Eut shell proteins can be classified into the CcmK/CsoS1 BMC protein family or the CcmL/CsoS4 family of proteins. Members of the CcmK1/CsoS1 and CcmL/CsoS4 families of proteins typically form hexamers and pentamers, respectively, [1].

### **1,2-Propanediol-induced Microcompartments in *Salmonella enterica***

Studies indicating the high degree of homology between certain genes of the 1,2-PD utilization pathway in *S. enterica* and carboxysomal shell components have been previously reported [74]. In addition, the ability of *S. enterica* to synthesize polyhedral microcompartments in a coenzyme B<sub>12</sub>-dependent manner during aerobic growth on 1,2-PD as a sole source of carbon and energy was recently discovered by Bobik *et al.*[2]. Intriguingly, the Pdu organelles do not assemble when *S. enterica* cells are grown in routine media. It is now known that the assortment of proteins mediating the breakdown of 1,2-PD is encoded by a cluster of 23 genes, collectively referred to as the propanediol utilization, or *pdu*, operon [2]. These genes, with minor sequence variations, have also been identified in other enteric species of *E. coli* and other enteric species of *Klebsiella*, *Shigella*, and *Yersinia* [4, 74].

The first step in the Pdu pathway is the conversion of 1,2-PD to propionaldehyde, which is catalyzed by the enzyme diol dehydratase. This enzyme is encoded by the *pduCDE* genes and requires the cofactor adenosylcobalamin (AdoCbl) for catalysis. AdoCbl can be obtained either via *de novo* synthesis under strict anaerobic conditions, or by import of exogenous source in the form of cyanocobalamin. The propionaldehyde generated is then either reduced to propanol or oxidized to propionic acid via a series of reactions catalyzed by an array of enzymes, including alcohol dehydrogenase, CoA-dependent propionaldehyde dehydrogenase, phosphotransacetylase, and propionate kinase [19, 75]. Reduction to propanol serves in the regeneration of NAD from NADH, while oxidation to propionic acid provides

ATP and an intermediate carbon source in the form of propionyl CoA. Propionyl CoA can be utilized in the generation of pyruvate via the methyl citrate pathway [19].

Recent genetic and biochemical evidence has suggested that the enzymatic breakdown of 1,2-PD occurs in a polyhedral microcompartment, which may serve to control the toxic propionaldehyde liberated in the first step of the 1,2-PD utilization pathway [19, 76]. It was observed that *S. enterica* mutants lacking the shell-encoding genes failed to grow in the presence of 1,2-PD. Further, this growth defect was found to be 1,2-PD concentration-dependent [76]. Since of all the intermediates formed in the 1,2-PD utilization pathway propionaldehyde is known to be cytotoxic, it was proposed that the primary function of the polyhedral organelle is to control and reduce toxic aldehyde formation possibly by moderating intra-organellar AdoCbl levels. The alleviation of toxicity may be achieved by the ability of the organelle to selectively block the entry of excess AdoCbl [19]. Alternatively, it has also been hypothesized that the Pdu organelles ensure that the excess aldehyde within the organelles inhibits the diol dehydratase enzyme activity via a feedback mechanism, thereby preventing more aldehyde formation [75]

One of the most striking features of the Pdu organelles is their close resemblance, in size and shape, to carboxysomes. However, in comparison to carboxysomes, the Pdu organelles exhibit a certain degree of geometrical irregularity. The presumed shell proteins of the Pdu organelles, which are encoded by the *pduA*, *pduJ*, *pduK*, *pduT*, *pduB*, and *pduU* genes, bear varying degrees of homology to the carboxysomal CsoS1 and CcmK BMC family of proteins and the CsoS4 and CcmL protein family [72, 74]. Immunoelectron microscopy studies suggested that the diol

dehydratase enzyme is encapsulated within the Pdu organelles [2, 19]. Although the *pdu* locus is comprised of 23 genes, only 14 genes have been found to encode proteins that are associated with the organelle. These include the shell components PduA, PduB, PduB', PduJ, PduK, PduN, PduT, and PduU. PduB' is thought to be a truncated form of PduB that lacks the N-terminal 37 residues [75]), diol dehydratase (Pdu-CDE), diol dehydratase reactivating factor (PduG, PduH), propionaldehyde dehydrogenase (PduP), and cobalamin adenosyltransferase (PduO). The PduN protein, which is related to the CsoS4A and CsoS4B proteins of *H. neapolitanus*, was not found to be a component of the Pdu organelles [75]. Since the CsoS4A and CsoS4B proteins are predicted to be present in such low copy numbers, it remains possible that the PduN protein, too, is expressed at low levels but is evading detection by conventional SDS-PAGE analysis. Intriguingly, Bobik *et al.* [2, 77] reported that deletion of the *pduCDE* genes, which encode the encapsulated diol dehydratase enzyme, does not interfere with the assembly of the Pdu shell proteins. However, this conclusion was made by observing thin sections of *pduCDE* mutant cells. The composition and stoichiometry of the Pdu shells in this mutant await biochemical characterization and comparison with those of wild type organelles.

### **Ethanolamine-induced Microcompartments in *Salmonella enterica***

Analogous to the 1,2-PD utilization pathway, *S. enterica* is capable of utilizing ethanolamine (EA) as a sole source of carbon and energy in an AdoCbl-dependent manner [78]. The assortment of proteins needed for EA breakdown is encoded by a cluster of 17 genes collectively referred to as the ethanolamine utilization (*eut*) operon. Several genetic studies have indicated the possible

involvement of a polyhedral microcompartment, also referred to as a metabosome, in the Eut pathway [3, 4, 21].

The Eut pathway begins with the conversion of ethanolamine to acetaldehyde and free ammonia, a reaction that is catalyzed by the enzyme ethanolamine ammonia lyase (EAL). The functional form of this enzyme, encoded by the *eutBC* genes, requires a reactivation factor encoded by the *eutA* gene and adenosylcobalamin as a cofactor [79, 80]. The acetaldehyde generated in the initial step is later converted to ethanol and acetate, which is subsequently activated to acetyl CoA by the alcohol and aldehyde dehydrogenase enzymes. The ammonia generated in the first step is utilized as a source of nitrogen and the acetyl CoA generated in the subsequent steps is fed into the TCA cycle [81, 82].

Similar to the 1,2-PD pathway, the breakdown of EA is thought to be accomplished within a polyhedral microcompartment [21]. Several findings support this notion. The 1,2-PD and EA utilization pathways have common reactions that produce similar intermediates and end products in the form of aldehydes, alcohols and acids. Furthermore, the *eut* operon is comprised of certain genes, namely, *eutK*, *eutL*, *eutM*, *eutN*, and *eutS*, that encode carboxysomal and Pdu shell homologs. Finally, deletion of the certain shell protein-encoding *eut* genes resulted in the inability of *S. enterica* cells to grow in the presence of EA [21]. It is plausible that the Eut polyhedral microcompartment, like the Pdu organelle, may aid in controlling acetaldehyde production, thereby minimizing its toxic effects. However, a recent report suggested that the microcompartment, instead of posing a barrier to the entry of AdoCbl, may serve to concentrate low levels of ethanolamine catabolic enzymes like

the acetaldehyde dehydrogenase and alcohol dehydrogenase, which keep the acetaldehyde generated during EA catabolism from accumulating by rapidly consuming it [21]. It was also reported that mutants lacking one or more of the putative shell or EAL genes assembled a variant of the Eut organelle that was capable of supporting growth in medium containing EA. Thin sections of these mutants exhibited the presence of intracellular electron dense regions, the nature of which is unclear. These observations suggested that, unlike what is observed for the Pdu organelles, the assembly of the Eut microcompartments may rely on the presence of the EAL enzyme that is thought to be encapsulated. However, since the Eut microcompartments have not yet been purified and characterized, the composition and stoichiometric distribution of the individual Eut polypeptides remain unknown. Thus, results from genetic studies like those mentioned above cannot be corroborated with direct biochemical approaches at this point.

A recent study in *Vibrio furnissii* M1 identified a BMC gene cluster encoding proteins that are similar to the shell components of the Pdu and Eut organelles [83]. It was hypothesized that this bacterium synthesizes a microcompartment for the metabolism of pyruvate to ethanol [83]. In addition, recent comparative genomics and bioinformatics analyses have identified close to 200 bacterial species and strains that harbor BMC protein-encoding gene clusters [1]. The widespread distribution of these genes suggests that many bacteria have the potential to form microcompartments as part of a general metabolic strategy.

### Objective of this study

Similar to eukaryotic assemblies like the nuclear pore complex, the assembly of polyhedral bacterial microcompartments is likely to be directed by specific protein interactions. Although recent structural studies have aided in building a rudimentary atomic level model of the carboxysome, the prototype bacterial microcompartment, several issues concerning the biogenesis of carboxysomes and related polyhedra remain unsolved. *In vitro* studies of  $\alpha$ - and  $\beta$ -carboxysomes have provided clues about the network of protein interactions that may be required for the assembly process, but *in vivo* studies validating such interactions are lacking. This study was undertaken to examine the role of encapsulated enzyme(s) in microcompartment biogenesis. For this purpose, yeast two hybrid screens of the putative shell proteins and the encapsulated ethanolamine ammonia lyase enzyme of the ethanolamine utilization organelles of *Salmonella enterica* were evaluated. Further, the ability of *Halothiobacillus neapolitanus* to assemble carboxysomes *in vivo* in the absence of a Form IA RubisCO-encoding gene set was tested. Finally, the possibility of *H. neapolitanus* carboxysomes to package foreign RubisCO species was investigated.

## CHAPTER III

### EXPERIMENTAL PROCEDURES

#### Routine Apparatus

Liquid cultures of *H. neapolitanus* and *T. crunogena* were grown in an Infors-HT Multitron shaker or maintained in an Infors-HT labfors 2 L fermentor, while solid cultures were grown on plates incubated in an Equatherm 1572 or a Fisher Scientific Isotemp Plus D650 incubator. Cultures of *E. coli* and *Saccharomyces cerevisiae* were grown in a Fisher Scientific Isotemp or an Equatherm 1572 incubator, a New Brunswick TC-7 culture rotator, or a New Brunswick Series 25 incubator shaker. Long term storage of cultures, clones, and DNA was accomplished using a Thermo Scientific Revco Elite series -80°C freezer. Sterilization of media and solutions were done in a Steris Amsco Lab 250 or a 3031-S autoclave. Protocols necessitating sterile techniques were carried out in a SterilGuard Class II Type A/B3 hood by Baker Company Inc. Cell breaking was done using a Branson model 450 sonifier. Short term storage of cell extracts at 4°C was accomplished using a Sears or Whirlpool refrigerator. Optical density measurements were made in a Beckman Coulter DU800 UV-vis spectrophotometer. Large scale high speed centrifugations were carried out in Beckman Avanti J-30I and Avanti J-26XP centrifuges using the JA-14, JA-20, JA-25.5, JSP-24.38, and F14BCI-6x250y series rotors, while microcentrifugation was done in an Eppendorf 5417C and 5417R microcentrifuge. Ultracentrifugation was performed in a Beckman L7-65 ultracentrifuge as well as a Beckman Optima Max ultracentrifuge. Polymerase chain reactions for DNA amplification were run on

BioRad's MyCycler thermal cycler. DNA restriction digests (37°C) and ligation reactions (16°C) were incubated in Fisher Versa-Bath model 137 and 133 waterbaths respectively. DNA gel electrophoresis was done using electrophoresis cells by Continental Lab Protocols model 75.710 and GibcoBRL Horizon 58 Model 200. DNA concentration measurements were done using a Nanodrop ND-1000 spectrophotometer. Protein gel electrophoresis was performed using a BioRad Criterion or Mini-Protean III cell. Western blotting was done using a BioRad Mini Trans-Blot or Criterion electrophoretic transfer cell. Visualization and image capturing of DNA and protein gels and immunoblots were done using BioRad's VersaDoc Imaging System. For electron microscopy, Zeiss EM 109 and JEOL JEM-100 electron microscopes were used. Scanning of negative images was done using an Epson Perfection V700 Flatbed Photo scanner. Mass spectrometry of trypsin-digested proteins was done using a Bruker Microflex-LT MALDI-ToF instrument (Bruker Daltonics Inc.). Microliter volumes were measured and pipetted using Rainin Pipetman P-1000, P-200, and P20 pipettes. For pipetting larger volumes, a Drummond pipet-aid was used. pH measurements were made using an Orion 8102BNUWP Ross Ultra Combination pH probe connected to an Orion 720A+ pH meter by Thermo Electron. Deionized water was obtained from a Barnstead NANOpure Diamond water system. Mass measurements were done using a Mettler Toledo AB54-S or PV303-S balance.

Software programs used were DU800 v 2.0 Build 83 by Beckman Coulter (DNA concentration measurements), Quantity One v 4.6.3 by BioRad (visualization and densitometry of DNA and protein bands), pDRAW32 v 1.1.100 by Acaclone software (generation of DNA vector maps for cloning purposes), ClustalX v 1.83

(DNA and protein sequence alignments, FlexControl v 2.4 and FlexAnalysis v 2.4 (mass spec data acquisition and analyses), Adobe Photoshop v 7.0.1 (image processing)).

### **Routine Chemicals**

Routine chemicals were purchased from VWR, Sigma and Fisher Scientific. B-PER II and GelCode Blue stain, One-Step NBT-BCIP reagent, BCA protein assay reagents were purchased from Thermo Scientific, Criterion PreCast 4–20% polyacrylamide gradient Tris-HCl protein gels and nitrocellulose blotting paper, were from BioRad, oligonucleotides were from IDTDNA, restriction enzymes, Deep Vent DNA polymerase, DNA ligase were from New England Biolabs, deoxynucleotide triphosphate (dNTPs) mixture was from Takara Bio Inc., plasmid isolation and gene cleaning kits were from Qiagen and Zymo Research, chicken anti-RbcL antibodies (product No. AS01 017) were from Agrisera, rabbit anti-CsoS1 polyclonal antibodies were generated at Cocalico Biologicals, goat anti-chicken and anti-rabbit HRP-tagged antibodies were from Santa Cruz Biotechnology, formvar coated copper grids, ammonium molybdate stain were from Electron Microscopy Services, peptide calibration standards for mass spectrometry were from Bruker Daltonics, Matchmaker Gal4 Two Hybrid System 3 for yeast two-hybrid screens was from Clontech,

## Media

### ***H. neapolitanus* liquid medium**

4 g/L  $K_2HPO_4$   
 4 g/L  $KH_2PO_4$   
 10 g/L  $Na_2S_2O_4$   
 0.4 g/L  $NH_4Cl$   
 0.4 g/L  $MgSO_4$  (anhydrous)  
 10 ml Trace element solution  
 0.1% v/v of 1% w/v phenol red solution prepared in 50% ethanol  
 Adjust pH to 6.8 and autoclave for 20 min at 121°C/15 psi

### ***H. neapolitanus* solid medium**

0.8 g/L  $K_2HPO_4$   
 0.2 g/L  $KH_2PO_4$   
 10 g/L  $Na_2S_2O_4$   
 1 g/L  $NH_4Cl$   
 0.01 g/L  $CaCl_2$   
 0.24 g/L  $MgSO_4$  (anhydrous)  
 1% v/v trace elements solution  
 15 g/L agar  
 0.1% v/v of 1% w/v phenol red solution prepared in 50% ethanol  
 Adjust pH to 6.8 and autoclave for 20 min at 121°C/15 psi

### **Trace elements for *H. neapolitanus* medium (100X)**

50 g/L EDTA  
 5.44 g/L  $CaCl_2$   
 1.61 g/L  $CoCl_2$   
 1.57 g/L  $CuSO_4 \cdot 5H_2O$   
 4.99 g/L  $FeSO_4 \cdot 7H_2O$   
 5.06 g/L  $MnCl_2 \cdot 4H_2O$   
 1.10 g/L  $(NH_4)_6Mo_7O_{24} \cdot 4H_2O$   
 2.20 g/L  $ZnSO_4 \cdot 7H_2O$

EDTA was dissolved in water first and pH adjusted to 6.0 with 20% w/v KOH. Next, chemicals were added sequentially maintaining pH at 6.0. The final pH was adjusted to 6.8 and the solution stored at 4°C

### ***T. crunogena* liquid medium**

Solution A (700 ml)  
 20.46 g NaCl  
 0.866 g  $NH_4SO_4$   
 0.732 g  $MgSO_4$   
 0.1% v/v of 1% w/v phenol red solution prepared in 50% ethanol  
 Adjust pH to 7.6

Solution B (230 ml)

0.294 g CaCl<sub>2</sub>

26.01 g HEPES

Adjust pH to 7.6

Solution C (20 ml)

0.539 g K<sub>2</sub>HPO<sub>4</sub>

Autoclave solutions A, B, and C separately for 20 min at 121°C/15 psi and when cool, mix with 30 ml of 1M Na<sub>2</sub>S<sub>2</sub>O<sub>3</sub> and 0.1% v/v trace elements

### ***T. crunogena* solid medium**

Solution A (700 ml)

25.16 g NaCl

0.866 g NH<sub>4</sub>SO<sub>4</sub>

0.732 g MgSO<sub>4</sub>

0.1% v/v of 1% w/v phenol red solution prepared in 50% ethanol

15 g agar

Solution B (230 ml)

0.294 g CaCl<sub>2</sub>

3.152 g Tris

Adjust pH to 7.6

Solution C (20 ml)

0.539 g K<sub>2</sub>HPO<sub>4</sub>

Na<sub>2</sub>S<sub>2</sub>O<sub>3</sub> (10 ml, filter sterilized)

0.1% v/v Trace elements solution

Autoclave solutions A, B, and C separately for 20 min at 121°C/15 psi and when cool, mix with Na<sub>2</sub>S<sub>2</sub>O<sub>3</sub> and trace elements

### **Luria Bertani (LB) broth**

10 g/L Bacto tryptone

5 g/L Bacto yeast extract

10 g/L NaCl

Autoclave for 20 min at 121°C/15 psi

For solid LB medium, add 1.5% w/v agar before autoclaving

### **Super optimal with catabolite repression (SOC) medium**

20 g/L Bacto tryptone

5 g/L Bacto yeast extract

0.58 g/L NaCl

0.186 g/L KCl

0.952 g/L MgCl<sub>2</sub>

3.6 g/L glucose

Autoclave above components for 20 min at 121°C/15 psi

**Yeast extract peptone dextrose (YEPD) medium**

20 g/L Difco peptone

10 g/L Yeast extract

0.1 g/L adenine hemisulfate

Autoclave above components for 20 min at 121°C/15 psi and once cool and add 2% v/v sterile dextrose. For solid medium, add 1.5% w/v agar before autoclaving

**Synthetic complete dropout (SCD) medium**

6.7 g/L Yeast nitrogen base without amino acids

0.6 g/L of Complete synthetic mixture (CSM) – X (where X represents the amino acid(s) to be left out for selection purposes)

10 mM 3-amino-1,2,4-triazol (3-AT)

2% v/v sterile glucose

Filter sterilize the above components. For solid medium, autoclave 1.5% w/v agar separately and once cool, add above components. For protein-protein interaction screening, add 5-bromo-4-chloro-3-indolyl- beta-D-galactopyranoside (X-gal) to a final concentration of 50 µg/ml before pouring plates

**No carbon E (NCE) medium**

0.68 g/L  $\text{KH}_2\text{PO}_4$

0.12 g/L  $\text{MgSO}_4$

1.04 g/L  $\text{NaNH}_4\text{HPO}_4$

Adjust pH to 7.0 with MOPS buffer and autoclave above components for 20 min at 121°C/15 psi. Once medium cools add 150 nM cyanocobalamin and 30 mM filter-sterilized ethanolamine

**Trace elements for NCE medium**

0.3 µM  $\text{CaCl}_2$

0.1 µM  $\text{ZnSO}_4$

0.045 µM  $\text{FeSO}_4$

0.2 µM  $\text{Na}_2\text{Se}_2\text{O}_3$

0.2 µM  $\text{Na}_2\text{MoO}_4$

0.2 µM  $\text{MnSO}_4$

0.1 µM  $\text{CuSO}_4$

3 µM  $\text{CoCl}_2$

0.1 µM  $\text{NiSO}_4$

**Antibiotic solutions**

Stock: 100 mg/ml ampicillin prepared in deionized water and filter sterilized  
Final concentration in medium: 100 µg/ml

Stock: 50 mg/ml kanamycin prepared in deionized water and filter sterilized  
Final concentration in medium: 50 µg/ml

## Buffers

### **TBE buffer pH 8.0**

89 mM Tris base  
89 mM Boric acid  
2 mM EDTA

### **TAE buffer pH 8.0**

40 mM Tris-HCl (pH 7.8)  
20 mM Na-acetate  
2 mM EDTA

### **TE buffer pH 8.0**

10 mM Tris-HCl (pH 8.0)  
1 mM EDTA  
Autoclave above components

### **Alkaline lysis solution I (for plasmid isolation)**

50 mM glucose  
25 mM Tris-HCl (pH 8.0)  
10 mM EDTA (pH 8.0)

### **Alkaline lysis solution II (for plasmid isolation)**

0.2 N NaOH  
1 % w/v Sodium dodecyl sulfate  
Always prepare fresh

### **Alkaline lysis solution III (for plasmid isolation)**

5 M potassium acetate  
Add glacial acetic acid to a pH of 4.8

### **TEL buffer for yeast transformation**

10 mM Tris-HCl (pH 8.0)  
1 mM EDTA  
100 mM lithium acetate  
Autoclave for 20 min at 121°C/15 psi

### **PEG solution for yeast transformation**

50% w/v PEG 4000 prepared in deionized water,  
Filter sterilize

**Phosphate buffered saline buffer (PBS), pH 7.4**

137 mM NaCl

2.7 mM KCl

4.3 mM Na<sub>2</sub>HPO<sub>4</sub>1.4 mM KH<sub>2</sub>PO<sub>4</sub>

Add 0.1% v/v Triton X-100 to prepare PBS-Triton buffer

**Tris-EDTA-magnesium-bicarbonate (TEMB) buffer, pH 8.0**

10 mM Tris

1 mM EDTA

10 mM MgCl<sub>2</sub>20 mM NaHCO<sub>3</sub>

Adjust pH to 8.0 with HCl and filter sterilize

**Tris-EDTA-magnesium-ethanolamine (TEME) buffer, pH 8.0**

10 mM Tris

1 mM EDTA

10 mM MgCl<sub>2</sub>

0.2% v/v ethanolamine

Adjust pH to 8.0 with HCl and filter sterilize

**Bicine-EDTA-magnesium-bicarbonate (BEMB) buffer, pH 8.0**

10 mM Bicine

1 mM EDTA

10 mM MgCl<sub>2</sub>20 mM NaHCO<sub>3</sub>

Adjust pH to 8.0 with HCl and filter sterilize

**PMSF/PTSF stock solution**

100 mM phenylmethylsulfonylfluoride

100 mM p-toluenesulfonylfluoride

Dissolve in 95% ethanol

**Breaking buffer (pPROEX-based expression and purification), pH 8.0**

50 mM Tris HCl (pH 8.0)

5 mM 2-mercaptoethanol

1 mM PMSF/PTSF added immediately before use

**Column equilibration buffer (pPROEX-based expression and purification), pH 8.0**

20 mM Tris HCl (pH 8.0)

100 mM KCl

5 mM 2-mercaptoethanol

10% v/v glycerol

20 mM imidazole

**High salt buffer (pPROEX-based expression and purification), pH 8.0**

20 mM Tris HCl (pH 8.0)  
1 M KCl  
5 mM 2-mercaptoethanol  
10% v/v glycerol

**Elution buffer (pPROEX-based expression and purification), pH 8.0**

20 mM Tris HCl (pH 8.0)  
100 mM KCl  
5 mM 2-mercaptoethanol  
10% v/v glycerol  
150 mM imidazole

**Protein dialysis buffer for recombinant proteins, pH 8.0**

10 mM Tris HCl (pH 8.0)  
400  $\mu$ M PMSF/PTSF

**4X SDS-PAGE loading buffer**

200 mM Tris HCl (pH 6.8)  
40% v/v glycerol  
8% w/v SDS  
10% v/v 2-mercaptoethanol  
Sterile deionized water to adjust final volume

**SDS-PAGE running buffer (Laemmli buffer)**

25 mM Tris  
192 mM glycine  
1% w/v SDS

**Western blot transfer buffer**

25 mM Tris  
192 mM glycine  
20% v/v methanol  
Store at 4°C

**Immunoblot blocking buffer**

5% non-fat dry milk dissolved in PBS buffer containing 0.1% v/v Triton X-100  
Buffer always prepared fresh

**Trypsin stock solution**

Dissolve trypsin at 1  $\mu$ g/ $\mu$ l in 50 mM acetic acid

**Trypsin digestion buffer**40 mM NH<sub>4</sub>HCO<sub>3</sub>

10% acetonitrile

Dissolve stock trypsin in this buffer to a final concentration of 20µg/ml

**Cultures and Growth Conditions*****Escherichia coli***

For routine cloning and expression purposes, Top10 cells (genotype: *F*- *mcrA*  $\Delta$ (*mrr-hsdRMS-mcrBC*)  $\phi$ 80*lacZ* $\Delta$ *M15*  $\Delta$ *lacX74* *nupG* *recA1* *araD139*  $\Delta$ (*ara-leu*))7697 *galE15* *galK16* *rpsL*(Str<sup>R</sup>) *endA1*  $\lambda$ ) from Invitrogen were grown either in liquid LB medium in a culture shaker (250 rpm, 37°C) or on solid LB medium plates incubated at 37°C. During chemical transformations, Top10 cells were heat shocked at 42°C and recovered in SOC liquid medium. For cloning of plasmids via homologous recombination, DY330 cells (genotype: *F*  $\lambda$  *rph-1* *INV(rrnD, rrnE)*  $\Delta$ *lacU169* *gal490* *lcI857*  $\Delta$ (*cro-bioA*)) were grown in liquid and solid LB medium incubated at 30°C (recombination repressive condition) or 42°C (recombination permissive condition) [84].

***Salmonella enterica* serovar Typhimurium LT2**

For routine purposes, wild type *S. enterica* serovar Typhimurium LT2 (ATCC 19585) culture was maintained in liquid LB medium in a culture shaker (250 rpm, 37°C) or on solid LB medium plates incubated at 37°C. For purification of the ethanolamine utilization (Eut) microcompartments and ethanolamine ammonia lyase assays, LT2 cells were grown to saturation in 750 ml of NCE medium supplemented

with 150 nM cyanocobalamin and 1% v/v succinic acid in a 2 L baffled flask incubated at 37°C.

### ***Halothiobacillus neapolitanus***

Wild type *H. neapolitanus* (ATCC 23641) was grown as a chemostat culture using the medium described by Vishniac and Santer [85] in a 2 L bioreactor with a dilution rate of 0.08 h<sup>-1</sup> and maintained at 30°C. *H. neapolitanus* Form IA RubisCO mutants were grown as 750 ml cultures (2 L baffled flasks) in air supplemented with or without 5% CO<sub>2</sub> in shaker incubators (200 rpm, 30°C).

### ***Thiomicrospira crunogena***

For routine maintenance, wild type *Thiomicrospira crunogena* strain XCL-2 [17] was grown in 50 ml of medium in a 250 ml baffled flask in a shaker incubator (200 rpm, 30°C). For carboxysome isolation, *T. crunogena* was maintained as a chemostat culture grown at 30°C in a 2 L bioreactor with a medium dilution rate of 0.12 h<sup>-1</sup>.

### ***Saccharomyces cerevisiae***

Wild type *S. cerevisiae* AH109 (genotype: *Mat a, trp1-901, leu2-3, 112, ura3-52, his3-200, gal4Δ, gal80Δ, LYS2::GAL1<sub>UAS</sub>-GAL1<sub>TATA</sub>-HIS3, MEL1 GAL2<sub>UAS</sub>-GAL2<sub>TATA</sub>-ADE2, URA3::MEL1<sub>UAS</sub>-MEL1<sub>TATA</sub>-lacZ*) was grown on solid and liquid YEPD medium supplemented with 0.1 g/L adenine at 30°C. Cells co-transformed with the pGADT7 and pGBKT7 constructs were selected on SCD-leu-trp medium supplemented with 0.1 g/L adenine at 30°C. To screen for protein-protein interactions, co-transformants were patched onto SCD-leu-trp-his-ade plates containing 50 μg/ml X-gal and incubated at 30°C.

## Methods

### Isolation of genomic DNA from *T. crunogena*, *H. neapolitanus*, and *S. enterica*

50 ml cultures of *T. crunogena*, *H. neapolitanus*, and *S. enterica*, grown to saturation, were harvested by centrifuging at 12,000 x g for 10 min at 4°C. The resulting cell pellet was resuspended in 5 ml TE buffer containing 0.5% SDS, 2 mg/ml Proteinase K and incubated at 37°C for an hour. To this slurry, 5 ml of TE-saturated phenol was added and the mixture was shaken vigorously and centrifuged in a 30 ml Corex glass tube at 12,000 x g for 15 min at 4°C. The aqueous phase was carefully transferred to a separate 30 ml Corex tube. 500 µl of 3 M sodium acetate pH 5.2 and 10 ml of ice cold isopropanol were added and the resulting DNA precipitate was pelleted by centrifuging at 12,000 x g for 15 min at 4°C. The DNA pellet was resuspended in 400 µl TE buffer and transferred to a 1.5 ml microfuge tube. To this suspension, 40 µl of 3 M sodium acetate pH 5.2 and 800 µl of 100% ethanol were added. The tube was then incubated on ice for 5 min and centrifuged at 21,000 x g for 10 min. The resulting DNA pellet was again resuspended in 400 µl TE buffer and the sodium acetate-ethanol precipitation step was repeated. The final DNA pellet obtained was air dried and resuspended in 100 µl TE buffer or sterile deionized water. For routine polymerase chain reactions, DNA purified by this method was used as template. For sequencing of specific genomic loci, DNA was purified and isolated using the Zymo Fungal/Bacterial DNA kit, following the manufacturer's protocol.

### Polymerase chain reactions

Reactions, which were set up in a total volume of 50 µl in thin-walled 0.2 ml tubes, contained ThermoPol buffer (20 mM Tris HCl, 10 mM (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub>, 10 mM

KCl, 2 mM MgSO<sub>4</sub>, 0.1% Triton X-100, pH 8.8), 0.2 μM of each dNTP (dATP, dGTP, dCTP, and dTTP), 2 U of Deep Vent Polymerase (New England Biolabs), 0.5 μM of forward and reverse primers, and 10 ng of plasmid DNA or 100 ng of genomic DNA. Typically, all samples were heated at 95°C for 5 min to ensure complete denaturation of template DNA prior to subjecting them to 30 cycles of the following three steps: 1) denaturation at 95°C for 1 min, 2) annealing at  $T_m - 5^\circ\text{C}$  (5°C below the melting temperature ( $T_m$ ) of the primers being used) for 45 s, and 3) extension at 72°C for anywhere from 30 s to 4 min depending on the length of DNA being amplified, following the general rule of thumb that the enzyme polymerizes 1 kb DNA per minute. At the end of 30 cycles, an additional extension step at 72°C for 10 min was included.

### **Restriction digest of DNA**

Plasmid DNA was cut using restriction endonucleases (New England Biolabs) in a total reaction volume of 100 μl. Typically, the reaction consisted of 10 uL of the corresponding 10X buffer, 10 μl of 10X bovine serum albumin if necessary, 1-5 μg of plasmid DNA, 5-20 U of the restriction endonuclease, and sterile deionized water to adjust the final volume. Control reactions were included wherever necessary. The reactions were incubated at 37°C anywhere from 1 to 4 h. The resulting DNA fragments were separated by DNA gel electrophoresis for further analyses.

### **DNA Ligation**

DNA ligation reactions were set up in a total volume of 10 μl and consisted of 50-100 ng of vector DNA with a 5-fold molar excess of insert, T4 DNA ligase buffer (50 mM Tris HCl pH 7.5, 10 mM MgCl<sub>2</sub>, 1 mM ATP, 10 mM dithiothreitol), 400 U

of T4 DNA ligase, and sterile deionized water to adjust the final volume. Reactions were incubated at 16°C overnight. Aliquots (2 µl) from these reactions were used to transform chemically competent *E. coli* cells.

To ligate PCR products into the pCR-Blunt II TOPO vector (Invitrogen), the Zero Blunt TOPO PCR cloning kit was used. Typically, the reaction contained 1 µl (10 ng) of the TOPO vector and anywhere from 1 to 4 µl of the PCR product, maintaining a 1:3 molar ratio of TOPO vector to linear PCR product, 1 µl of salt solution (1.2 M NaCl, 0.06 M MgCl<sub>2</sub>), and sterile deionized water to adjust the final volume. After incubation for 30 min at room temperature, the reaction was used to transform chemically competent Top10 *E. coli* cells.

#### **DNA gel electrophoresis, visualization, and fragment purification**

DNA concentrations were estimated by measuring the absorbance of 1 µl samples at 260 and 280 nm in a Nanodrop ND-1000 spectrophotometer. The instrument quantifies the concentration of DNA in a sample based on the relation that one A<sub>260</sub> unit corresponds to 50 µg/ml of dsDNA or 40 µg/ml ssDNA. Only samples whose A<sub>260/280</sub> ratios fell between 1.8 and 2.0 were considered for further experiments.

DNA samples were prepared in 6X gel loading dye (NEB Inc.) and separated by electrophoresis on a 0.8% w/v agarose gel prepared in TAE or TBE buffer. Electrophoresis was allowed to proceed at 5 V/cm until the bromophenol blue had migrated <sup>3</sup>/<sub>4</sub> the length of the gel, followed by staining of the gel in deionized water containing 0.5 µg/ml ethidium bromide for 15 to 20 min. After de-staining for 30 min in deionized water, gels were placed on a UV transilluminator and visualized. Gel

images were captured using BioRad's VersaDoc imaging system and Quantity One v 4.6.3 software.

To purify DNA fragments from agarose gels, DNA was separated on 0.8% w/v agarose gels prepared in TAE or TBE buffer as described above. Bands of interest were visualized under UV illumination, excised using clean razor blades, and further processed using the Zymoclean gel DNA recovery kit (Zymo Research).

### **DNA sequence analysis**

Plasmid and genomic DNA preparations were purified using the Qiagen Plasmid Purification kit and Zymo Fungal/Bacterial DNA kit, respectively, and shipped to the University of Maine DNA Sequencing Facility, Orono, ME, at a concentration of 0.1 µg/ µl. For the generation of DNA restriction maps, the pDRAW32 v 1.1.100 software (Acaclone) was used. Sequence alignments were done using the Local Alignment Tool software (<http://www.expasy.ch/tools/sim-nucl.html>).

### **Chemical transformation and electroporation of *E. coli* cells**

Plasmids or ligation reactions were mixed with 50 µl of chemically competent *E. coli* cells and incubated on ice for 30 min. Cells were heat-shocked at 42°C for exactly 45 s and immediately transferred to ice. 300 µl of room temperature SOC liquid medium was added to the cells and the contents were transferred to a 2057 polystyrene tube (Falcon), which was incubated at 37°C in a culture shaker for 1 h. Cells were then plated on LB solid medium containing antibiotic corresponding to the resistance marker borne on the plasmid and incubated at 37°C for overnight.

For preparation of electrocompetent DY330 cells, 50 ml LB liquid medium (250 ml baffled flask) was inoculated with an overnight culture of DY330 cells at a dilution of 1:50 and incubated at 30°C in a shaker incubator. Upon reaching an OD<sub>600</sub> of 0.5, 25 ml culture was transferred to a new 250 ml baffled flask, which was incubated with constant shaking for 15 min at 42°C (induced condition, wherein recombination functions are transiently expressed), while the remaining 25 ml culture was allowed to continue growing at 30°C for 15 min (uninduced condition). Both flasks were then transferred to an ice slurry and allowed to cool for 15 min. The cultures were then centrifuged at 5,500 x g for 10 min at 4°C and washed by resuspending the cells in 1 ml ice-cold sterile deionized water. The cell suspension was transferred to a 1.5 ml microfuge tube and centrifuged at 21,000 x g for 5 min at 4°C. The resulting cell pellet was washed three times, as described, and finally resuspended in 100 µl of sterile ice-cold deionized water containing 10% v/v sterile glycerol. For electroporation, 50 µl of induced and uninduced (control) electrocompetent cells were mixed with equimolar ratios of linear donor DNA and plasmid acceptor DNA in a pre-cooled electroporation cuvette (0.1 cm). Following electroporation, which was performed using a BioRad Gene Pulser set at 1.8 kV, 25 µF, with pulse controller at 200 Ohms, cells were immediately recovered in 1 ml LB liquid medium and incubated at 30°C for 1.5 h. Recombinants were selected by screening for *amp<sup>r</sup>* and *kan<sup>r</sup>* colonies.

#### **Generation of clones for yeast two-hybrid screens**

The *eutK*, *-L*, *-M*, *-N*, and *-S*, genes encoding putative shell proteins and the *eutB* and *-C* genes encoding the heavy and light chain subunits of the ethanolamine

ammonia lyase enzyme of the ethanolamine-induced microcompartments [6] were amplified from *Salmonella enterica* LT2 genomic DNA and cloned into the BamHI and NdeI restriction sites of the bait (pGBKT7) and prey (pGADT7) vectors of the Matchmaker GAL4 Two Hybrid System III (Clontech). Constructs were verified using restriction digests and sequencing of plasmid DNA as described in the results. All constructs were propagated and maintained in Top10 *E. coli* cells (Invitrogen) prior to transforming *Saccharomyces cerevisiae* AH109 cells.

### **Yeast transformation**

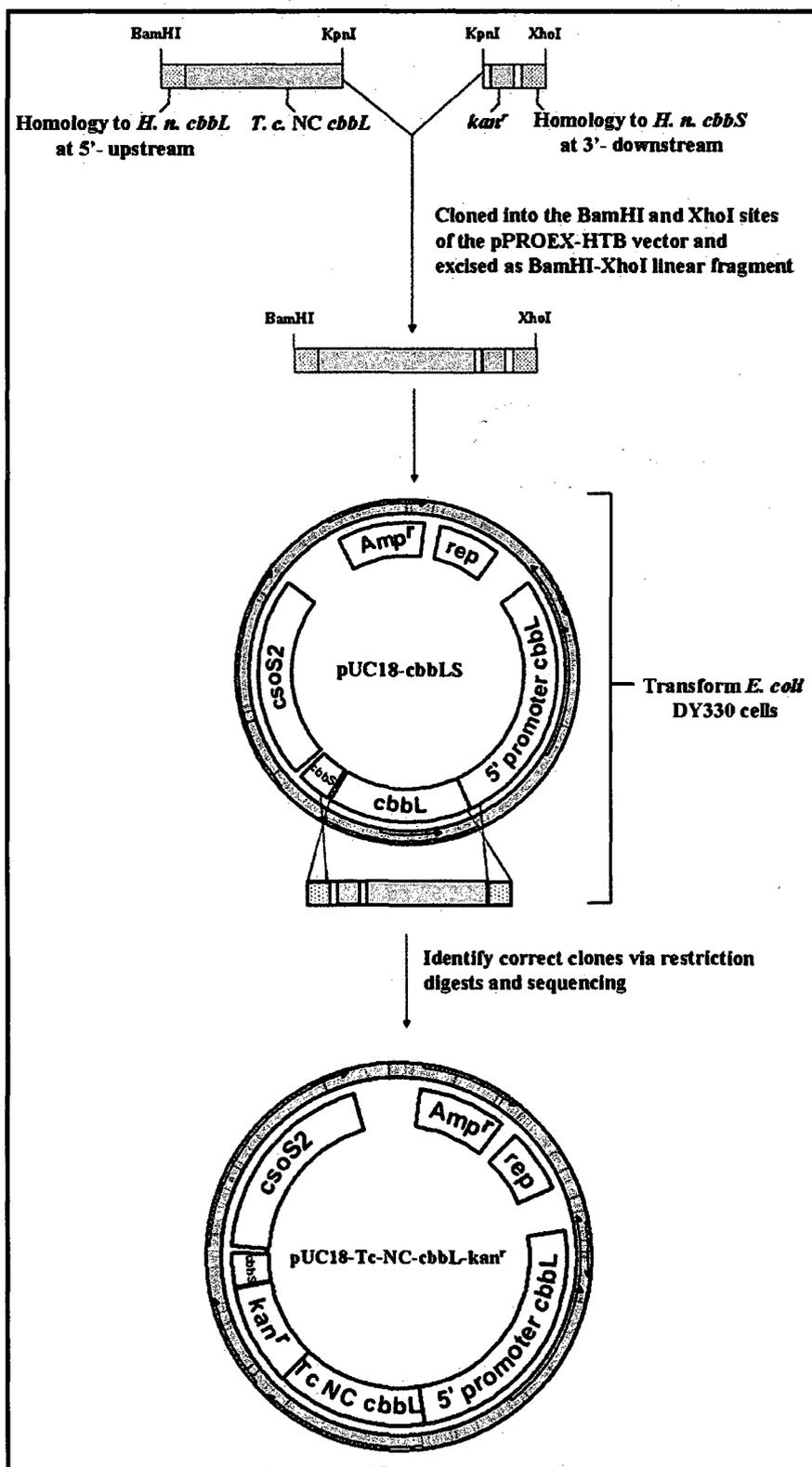
Transformation of wild type *Saccharomyces cerevisiae* AH109 was done using the method of Geitz *et al* [86]. A glycerol stock of *S. cerevisiae* AH109 was streaked onto solid YEPD medium and incubated at 30°C until colonies had reached 2-3 mm in size. A single colony was then picked and transferred to 5 ml liquid YEPD medium and incubated at 30°C for 2 days on a wheel rotating at full speed. Once growth of this culture was saturated, it was used to inoculate 50 ml liquid YEPD medium in a 250 ml Erlenmeyer flask at a 1:5000 dilution. The 50 ml YEPD culture was incubated at 30°C overnight on a shaker. When the culture reached an OD<sub>600</sub> of approximately 0.5, cells were harvested by centrifuging at 1000 x g for 10 min at room temperature. Pelleted cells were resuspended in 25 ml sterile deionized water and centrifuged again at 1000 x g for 10 min at room temperature. The supernatant was discarded and the cell pellet was resuspended in 500 µl of TEL buffer (TE buffer, pH 8.0 containing 0.1 M lithium acetate). The approximately 600 µl cell suspension was divided into 100 µl aliquots in sterile 1.5 ml microfuge tubes. To each 100 µl of cell suspension was added 10 µl of denatured herring sperm carrier DNA supplied at

a concentration of 1 mg/ml (Clontech), 1  $\mu$ g of the bait and/or prey constructs, and 600  $\mu$ l of PEG:TEL buffer (TEL buffer mixed with 50% w/v PEG 4000 solution in a 1:5 ratio). The contents were mixed by pipetting and the microfuge tube was incubated at 30°C for 30 min on a rotating wheel. Next, 100  $\mu$ l DMSO was added to each tube and the contents were mixed by inversion. The cells were heat shocked for 15 min by placing the tubes in a water bath maintained at 42°C and subsequently pelleted by centrifuging the tubes at 1000 x g for 1 min at room temperature. The cell pellet was resuspended in 100  $\mu$ l of sterile deionized water and plated on solid SCD medium lacking leucine (SCD-leu) for pGADT7 based constructs or solid SCD medium lacking tryptophan (SCD-trp) for pGBKT7 based constructs. Yeast cells co-transformed with both pGADT7 and pGBKT7 constructs were plated on solid SCD-leu-trp medium. Plates were incubated at 30°C until transformant colonies had reached a diameter of 2 mm. Ten colonies were selected and patch-plated onto solid SCD medium lacking the specified amino acid(s). Plates were stored at 4°C to be used as master plates for protein-protein interaction screening. For long term storage, 25% v/v glycerol stocks of the transformants were made and stored at -80°C.

#### **Generation of *H. neapolitanus* Form IA RubisCO mutants**

Briefly, genes encoding the carboxysomal RubisCO (large subunit GenBank CP000109, GeneID: 3760532; small subunit GenBank CP000109, GeneID: 3760533) and noncarboxysomal RubisCO (large subunit GenBank CP000109, GeneID: 3761246; small subunit GenBank CP000109, GeneID: 3761247) were amplified from *Thiomicrospira crunogena* XCL-2 genomic DNA and cloned into the pCR-BluntII-TOPO vector (Invitrogen). The resulting insert was excised by digestion with BamHI

and KpnI, and ligated along with a *kan<sup>r</sup>* cassette containing KpnI–XhoI ends into the BamHI–XhoI sites of the pPROEX-HTb vector (Invitrogen). In the case of the RubisCO deletion mutant (*cbbLS::kan<sup>r</sup>*), a *kan<sup>r</sup>* cassette was used for ligation into the BamHI–XhoI pPROEX-HTb vector. The resulting construct was digested with BamHI and XhoI to release the insert. *Escherichia coli* DY330 cells were co-transformed with this fragment and with pUC18 containing the *cbbL-cbbS* region of the *H. neapolitanus* *csO* operon to replace the wild type *cbbL-cbbS* region on the plasmid with that of the insert by homologous recombination (**Figure 3**). The resulting plasmid containing the mutated *cbbL-cbbS* region was electroporated into exponentially growing *H. neapolitanus* cells using the method of English *et al.*[46]. The presence of the desired changes in the *csO* operon of the *H. neapolitanus* mutants was confirmed by genomic sequencing (University of Maine DNA Sequencing Facility) and by PCR amplification.



**Figure 3. Strategy used for constructing the pUC18-Tc-NC-cbbL-kan<sup>r</sup> plasmid**  
 Plasmids used in the generation of the other Form IA RubisCO mutants were constructed using an identical strategy.

**Electroporation of *H. neapolitanus* cells [46]**

A 40 ml aliquot of exponentially growing *H. neapolitanus* cells was collected from the bioreactor and pelleted by centrifugation at 12,000 x g for 10 minutes at 4°C. The resulting cell pellet was resuspended in 40 ml of ice-cold sterile water and centrifuged at 12,000 x g for 10 minutes at 4°C. The wash step was repeated at least three times and the final cell pellet was resuspended in 200 µl of chilled sterile water. Typically, 1 µg of purified plasmid DNA was added to 50µl of freshly prepared electrocompetent *H. neapolitanus* cells. The mixture was transferred to a pre-cooled 0.1 cm gap electroporation cuvette (BioRad), which was pulsed at 19 kV/cm, 200 Ω, and 25 µF. The typical pulse time was around 4.5 ms. The transformation mixture was immediately transferred to 5 ml of chilled liquid *H. neapolitanus* medium and placed on ice for 5 minutes before being incubated in a shaker at 30°C for 24 hours in air supplemented with 5% CO<sub>2</sub>. Electroporated cells were allowed to recover for 48 h and 500 µl from this culture was inoculated into 5 ml of liquid *H. neapolitanus* medium containing 50 µg/ml kanamycin. Cultures were incubated for 4 to 5 days in air supplemented with 5% CO<sub>2</sub>. 100 µl from this culture was spread onto a plate containing solid *H. neapolitanus* medium until single colonies could be seen. Individual colonies were selected and analyzed by PCR amplification and genomic DNA sequencing.

**Protein quantification and separation**

Protein concentrations were determined using the bicinchoninic acid assay [87]. Reactions were read using a DU 800 UV/vis spectrophotometer (Beckman Coulter) at a wavelength of 562 nm.

Separation of polypeptides was accomplished by sodium-dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) in a Mini Protean III cell or Criterion cell (BioRad) using 12% polyacrylamide gels or precast 4-20% gradient Tris HCl gels, respectively. Electrophoresis conditions employed were 100 V for 2 h for the 12% gels and 150 V for 1.6 h for the precast gels. Post electrophoresis, gels were rinsed in deionized water for 30 min, followed by staining with GelCode Blue stain (Thermo Scientific) for 1 h. Gels were destained in deionized water overnight before visualizing and image-capturing of the protein bands.

### **Mass spectrometry**

Proteins to be analyzed by matrix-assisted laser desorption/ionization time of flight mass spectrometry (MALDI-ToF MS) were first separated by electrophoresis on a denaturing 4-20% SDS-polyacrylamide Tris HCl gradient gel. After electrophoresis, the gel was rinsed with deionized water three times for fifteen min each, followed by staining with GelCode blue for 1 h. Following overnight destaining of the gel in deionized water, the protein bands of interest were excised using a clean razor blade and transferred to a 1.5 ml microfuge tube that was pre-washed twice with 1 ml of 50% v/v acetonitrile (ACN)/0.1% v/v trifluoroacetic acid (TFA). Gel slices were destained twice with 0.2 ml of gel destaining buffer (100 mM  $\text{NH}_4\text{HCO}_3$ /50% v/v ACN) for 45 min each at 37°C. The destained gel slices were dehydrated in 0.1 ml of 100% ACN for 5 min at room temperature and dried in a Savant Speed-Vac concentrator (vacuum set to read between 150-200 mtorr) connected to a refrigerated vapor trap. The dried gel slices were allowed to rehydrate in 20  $\mu\text{l}$  of digestion buffer containing 20  $\mu\text{g/ml}$  trypsin for 1 h at room temperature. Following rehydration, 200-

300  $\mu$ l of digestion buffer was added and the reaction was incubated overnight at 37°C. The next day, 150  $\mu$ l of ultra-pure HPLC-grade water was added to the digested mix and incubated at room temperature for 10 min with frequent vortexing. The gel slices were extracted with 50  $\mu$ l 50% v/v ACN/5% v/v TFA twice for 60 min each at room temperature. Extracts were pooled and dried under vacuum as before. Dried extracts were resuspended in 10  $\mu$ l of 0.1% TFA and cleaned and concentrated using a C<sub>18</sub> Reversed-Phase Ziptip (Millipore). Prior to use, Ziptips were wetted in 10  $\mu$ l of 100% ACN and equilibrated three times in 10  $\mu$ l of 0.1% v/v TFA equilibration/wash buffer. Extracts were bound to the C<sub>18</sub> matrix by pipetting and aspirating at least ten times. The matrix was washed three times using 10  $\mu$ l of 0.1% TFA and purified bound proteins were eluted using 10  $\mu$ l 50% v/v ACN/0.1% v/v TFA.

Eluted proteins (0.5  $\mu$ l) were mixed with  $\alpha$ -cyano-4-hydroxycinnamic acid matrix (0.5  $\mu$ l) prepared in 50% ACN/0.05% TFA, and spotted onto a MSP Anchorchip 600/96 (Bruker Daltonics) sample well and allowed to dry. The evaporated mixture was inserted into the source of a benchtop Microflex MALDO-ToF MS instrument onto which the laser was fired at 30% intensity. Data acquisition and analyses of the resulting ionization spectra were performed using the FlexControl v 2.4 and FlexAnalysis v 2.4 software programs, respectively.

### **Electroblotting**

For immunoblotting purposes, proteins that were separated by SDS-PAGE were electro-transferred to 0.45  $\mu$ m pore size nitrocellulose membranes (Millipore) at 200 mA for 1.5 h at 4°C in the BioRad Mini Trans Blot cell or Criterion transfer cell

containing pre-chilled western blot transfer buffer. All blots were dried overnight at room temperature prior to performing immunoblotting.

### **Development of immunoblots**

Nitrocellulose membranes containing electro-transferred proteins were incubated in immunoblotting blocking buffer containing 5% w/v non-fat dry milk in PBS-Triton X-100 for 30 min at room temperature with slow rocking. The blots were then incubated for 1 h at room temperature in blocking buffer containing pre-determined dilutions of primary (1°) polyclonal antibodies raised in rabbit (IgG) or chicken (IgY) as indicated. Carboxysomal proteins were probed with 1° antibody dilutions ranging from 1:2500 – 1:20,000. The light chain of ethanolamine ammonia lyase, EutC, was probed with a 1:2500 dilution of 1° antibody. Blots were washed three times, 15 min each, with PBS-Triton X-100 buffer, followed by incubation with blocking buffer containing a 1:10,000 dilution of goat anti-rabbit or anti-chicken horseradish peroxidase (HRP)-tagged secondary (2°) antibody for 1 h at room temperature. After a series of washes as described above, membranes were treated with 5 ml of the NBT/BCIP reagent (Thermo Scientific) for 3 min and sandwiched between sheets of plastic films (Saran Wrap) prior to visualization of immunoreactive bands using the VersaDoc imaging system.

### **Purification of carboxysomes from *H. neapolitanus***

Wild type *H. neapolitanus* cells were maintained as a chemostat culture in the mid-exponential growth phase in a 2 L bioreactor with a medium dilution rate of 0.08 h<sup>-1</sup>, as described previously [36]. Stationary phase cells that were collected in an aerated 8 L carboy connected to the bioreactor were concentrated using a 0.2 μ pore

Millipore Pellicon filtration unit driven by a peristaltic pump. Concentrated cells were pelleted by centrifugation at 12,000 x g for 10 min. The cell pellet, which typically weighed approximately 6 g, was resuspended in 15 ml of TEMB buffer containing 2 mg/ml lysozyme, 10 mM MgSO<sub>4</sub>, and 0.2 μM PMSF/PTSF. The cell suspension was transferred to a sterile 50 ml conical tube and 20 ml B-PER II was added. The tube was inverted several times to ensure mixing of the detergent with the cell suspension. At this point, the consistency of the cell suspension became mucoid, which is an indication of cell lysis. The cell slurry was then sonicated using a disruption horn (9.5 mm diameter) connected to a Branson model 450 sonifier set at a constant duty cycle and a power output control of 7 for 10 s. This step ensured disruption of genomic DNA, cell membranes and release of intracellular contents. In addition, 150 μl of a 1 mg/ml bovine pancreatic DNase I solution in TEMB buffer was added to get rid of any remaining genomic DNA. The sonicated cell lysate was then agitated for 30 min at room temperature on a Reliable Scientific D55 shaker set at approximately 60 cycles per min. Contents were transferred to a 40 ml polypropylene tube and centrifuged at 10,000 x g for 10 min. The resulting cell pellet was discarded and the supernatant was carefully transferred to a fresh polypropylene tube and centrifuged at 48,000 x g for 30 min. The pellet obtained from this high speed spin, also referred to as P<sub>20K</sub> pellet, was resuspended in 3 ml TEMB buffer and centrifuged at 1000 x g for 3 min. This pre-clearing step yielded a supernatant fraction highly enriched in carboxysomes, which was then loaded onto a 36 ml 10-60% w/v continuous sucrose gradient in TEMB buffer prepared in a 40 ml 25 X 89 mm clear centrifuge tube using a gradient former (Bethesda Research Laboratories). The tube was then centrifuged in

a JS 28.38 swinging bucket rotor (Beckman) at 104,000 x g for 30 min at 4°C. Typically, carboxysomes migrated as a single milky white band towards the middle of the gradient. This band was harvested and transferred to a 25 x 89 mm polycarbonate centrifuge bottle. After adjusting the volume to nearly 36 ml with TEMB buffer, the tube was then centrifuged at 150,000 x g for 2 h at 4°C in a Type 70Ti rotor (Beckman). The resulting pellet was resuspended in 1 ml TEMB buffer and stored at 4°C until further use.

The protocol used to purify carboxysomes from *H. neapolitanus* Form IA RubisCO mutants was essentially a scaled-down version of that used for wild type cells. Cells were grown as 2 x 750 ml cultures in 2 L baffled flasks at 30°C in air supplemented with 5% CO<sub>2</sub>. Cells were resuspended in 5 ml of TEMB buffer containing lysozyme and MgSO<sub>4</sub>. 7.5 ml B-PER II reagent and 25 µl of DNase I solution were added and sonicated as described above, but using a 3 mm tapered microtip disruptor horn connected to a Branson 450 model sonifier. After the series of differential centrifugations described above, the P<sub>20K</sub> pellet was resuspended in 500 µl of TEMB buffer, clarified by a low speed centrifugation, and the resulting supernatant was loaded onto a 4.25 ml 10-60% w/v continuous sucrose gradient in TEMB buffer and centrifuged at 104,000 x g in ultra-clear 13 x 51 mm centrifuge tubes for 30 min at 4°C in an MLS 50 swinging bucket rotor (Beckman). Carboxysomes were collected, transferred to a 11 x 34 mm PC centrifuge tube, diluted in TEMB buffer, and centrifuged at 150,000 x g in an MLA 130 fixed angle rotor (Beckman) for 2 h at 4°C. The final carboxysome pellet was resuspended in 250 µl TEMB buffer and stored at 4°C until further use.

### **Isolation of carboxysomes from *T. crunogena***

Carboxysomes from wild type *T. crunogena* XCL-2 were isolated from a mid-exponential growth phase chemostat culture maintained in a 2 L bioreactor, as described for wild type *H. neapolitanus* cells, but with two major modifications. The pellet fraction resulting from the 10,000 x g spin was used as the starting material for carboxysome isolation rather than the supernatant fraction. This pellet fraction was sonicated for four 30 s bursts with 1 min cooling intervals between each burst. In addition, 1% v/v Nonidet P-40 was added to the sonicated lysate, in place of B-PER II reagent, for solubilization of membranes. Contents were transferred to a clean 100 ml glass beaker and stirred on a magnetic stir plate for 1 h at room temperature. The remaining steps were as described in the protocol for *H. neapolitanus* carboxysome purification.

### **Purification of ethanolamine utilization (Eut) organelles from *S. enterica***

The method used for purifying the Eut organelles was adapted from Havemann and Bobik [19] with minor modifications. *S. enterica* LT2 cells were grown to saturation ( $OD_{600}$  of approximately 1.2) in NCE medium supplemented with 150 nM cyanocobalamin and 1% v/v succinic acid in 2 L baffled flasks incubated in a shaker (250 rpm) at 37°C. Cells from 2 L of medium were harvested by centrifugation at 4,000 x g for 10 min at room temperature. Pelleted cells were washed in 300 ml of TEME buffer containing 2 mg/ml lysozyme, 10 mM  $MgSO_4$ , and 0.2  $\mu$ M PMSF/PTSF and incubated in 30 ml of the same buffer for 2 h at 37°C in a shaker incubator (100 rpm). The treated cells were pelleted by centrifugation at 7,740 x g for 10 min at 4°C and resuspended in 30 ml of TEME buffer. Cells were

then lysed by sonication for four 30 s bursts with 1 min cooling intervals as described for large scale purification of *H. neapolitanus* carboxysomes. After sonication, 30 ml of Bacterial Protein Extraction Reagent (B-PER II) solution containing 400 mM NaCl and 20 mM MgCl<sub>2</sub> were added to the cell extract and incubated in a culture shaker (50 rpm) for 30 min at 4°C. The detergent-treated cell extract was then centrifuged at 12,000 x g for 10 min at 4°C to remove the cell debris. The resulting supernatant was subjected to centrifugation at 48,000 x g for 90 min at 4°C. The pellet obtained from the high speed spin was resuspended in 2 ml of TEME buffer, loaded onto a 36 ml 35-65% w/w continuous sucrose gradient in TEME buffer prepared in a 40 ml 25 X 89 mm clear centrifuge tube. The gradient was subjected to centrifugation at 104,000 x g for 12 h at 4°C in a SW28 rotor. A translucent band that migrated 2/3<sup>rd</sup> into the sucrose gradient was harvested, transferred to a 25 x 89 mm polycarbonate centrifuge bottle, diluted in TEME buffer, and centrifuged at 150,000 x g for 2 h at 4°C in a Type 70Ti rotor. The resulting pellet was resuspended in 1 ml of TEME buffer and stored at 4°C until further use.

### **Transmission electron microscopy**

For transmission electron microscopy, 4 µl of diluted protein sample was placed on formvar coated copper grids (300 mesh, Electron Microscopy Services) and allowed to sit for 4 min. Excess sample was wicked off using clean tissue paper (Kimwipes). After air-drying the grids for about 2 to 5 min, samples were negatively stained for 40 s using 4 µl of 1% w/v ammonium molybdate stain prepared in TEMB buffer pH 8.0. Grids were again air-dried for nearly 2-5 min before making observations under the Zeiss EM 109 or JEOL JEM-100 electron microscope. Images

obtained with the Zeiss EM 109 electron microscope were captured by exposing Kodak EM 4489 films (Electron Microscopy Services). Exposed films were developed in D-19 developer solution (Kodak) for 4 min, rinsed in deionized water for 2 min, and fixed in fixer solution (Kodak) for 10 min. Fixed films were rinsed in deionized water for 5 min and allowed to air dry before being scanned on an Epson Perfection V700 photo scanner and processed using Adobe Photoshop v 7.0.1. Images of samples observed under the JEOL JEM-100 electron microscope were digitally captured by the built-in Gatan multiscan CCD camera. Raw images were converted to JPEG and TIFF file formats using the Gatan software and processed using Adobe Photoshop v 7.0.1.

#### **Ethanolamine ammonia lyase (EAL) assays**

EAL assays were performed as described by Toraya *et al.* [88] in a total volume of 1 ml. This assay uses 3-methyl-2-benzothiazolinone hydrazone (MBTH) to detect the amount of acetaldehyde formed from the breakdown of ethanolamine by EAL. The assay buffer contained 30 mM potassium phosphate pH 8.0, 50 mM KCl, 10 mM ethanolamine, 1.5  $\mu$ M adenosylcobalamin (AdoCbl), 3 mM ATP, and 3 mM  $MgCl_2$ . The reaction was initiated by the addition of 5 to 25  $\mu$ l of cell extract followed by incubation at 37°C for 30 min in the dark to avoid photolysis of cyanocobalamin. Subsequently, 0.1 ml of 1 M potassium citrate buffer (pH 3.6) was added to stop the reaction, followed by the addition of 0.5 ml of 0.1% w/v MBTH. After an additional 15 min at 37°C, the amount of aldehyde formed by reaction with MBTH was determined spectrophotometrically at 305 nm.

#### **RubisCO assays**

Carboxysomes and RubisCO samples were desalted on a Micro Bio-Spin 6 column (BioRad) and stored in 50 mM Bicine-NaOH, pH 8.0, prior to use. Radioactive  $\text{NaH}^{14}\text{CO}_3$  solution was prepared by adding 10  $\mu\text{l}$  of 2  $\mu\text{Ci}/\mu\text{l}$   $\text{NaH}^{14}\text{CO}_3$  stock solution (GE Healthcare) to 1 ml of freshly prepared 200 mM  $\text{NaHCO}_3$  solution to a final specific radioactivity of 0.1  $\mu\text{Ci}/\mu\text{mol}$ . The desalted RubisCO samples were activated in activation buffer (50 mM Bicine-NaOH, pH 8.0, 10 mM  $\text{MgCl}_2$  and 10 mM of 0.1  $\mu\text{Ci}/\mu\text{mol}$   $\text{NaH}^{14}\text{CO}_3$  for 10 minutes prior to the start of the assay. Each assay vial contained 0.5 ml of 50 mM Bicine-NaOH, pH 8.0, 20 mM  $\text{MgCl}_2$ , 0.5 mM ribulose 1,5-bisphosphate (RuBP, from Fluka) and 20 mM of  $\text{NaHCO}_3$ . A 10  $\mu\text{l}$  aliquot of activated RubisCO was added to initiate the reaction. The reaction was stopped by pipetting 100  $\mu\text{l}$  aliquots of the assay mixture into 300  $\mu\text{l}$  of glacial acetic acid every minute for a total of 3 minutes. The vials, which contained the acidified samples were heated for 4 minutes to evaporate excess  $^{14}\text{CO}_2$  and allowed to cool to room temperature before adding 4 ml of ScintiVerse II cocktail solution. Contents of the vials were mixed thoroughly by inversion. The radioactivity in acid-stable reaction products was quantified by scintillation counting in an LS 6000SC Beckman scintillation counter. A blank sample containing the same assay mixture, but no enzyme, was included in the assay. Quenching and counting efficiency were determined with the help of a standard curve that was constructed from various amounts of  $^{14}\text{C}$  labeled n-hexadecane added to 0.3 ml glacial acetic acid followed by addition of 4 ml of ScintiVerse II cocktail. All assays were performed in triplicates and repeated at least three times with independent carboxysome preparations.

### Recombinant protein purification

All recombinant proteins purified in this study were based on the pPROEX HTb prokaryotic expression system. Genes cloned within the multiple cloning site of the pPROEX HTb vector would result in an N-terminal polyhistidine fusion (His<sub>6</sub> tag for affinity purification) with the corresponding proteins. In this study, the *eutK*, *-L*, *-M*, *-N*, and *-S*, genes encoding putative shell proteins [21] of the ethanolamine-induced microcompartments as well as the *eutC* gene encoding the light chain of the ethanolamine ammonia lyase enzyme [21], from *S. enterica*, were cloned into the BamHI and XhoI restriction sites of the pPROEX HTb vector. Correct constructs were identified by restriction digest analyses and sequencing (University of Maine DNA Sequencing Facility). Clones of the *eutK*, *-L*, *-M*, *-N*, *-S*, and *-C* genes were used to inoculate 3 ml of LB + ampicillin liquid medium, which were grown overnight in a shaker incubator at 37°C. A 1:100 dilution of the overnight cultures served as inoculum for the 500 ml of LB + ampicillin liquid cultures, which were grown at 37°C in a shaker incubator until the OD<sub>600</sub> was 0.5. At that point, 0.6 mM isopropyl-β-D-galactoside (IPTG) was added to the cultures, which were allowed to grow at 37°C for an additional 3 h in a shaker incubator to induce expression of recombinant proteins. Optimal expression conditions were pre-determined in pilot scale studies and were found to be 37°C for 3 h post induction with 0.6 mM IPTG for all proteins. Expression in the pPROEX HTb based constructs is driven by the *trc* (*trp lac*) promoter and *lacI*<sup>q</sup> repressor gene. After induction, cells were harvested by centrifugation at 10,000 x g for 10 min at 4°C. Cell pellets were resuspended in 30 ml of breaking buffer (50 mM Tris HCl, pH 8.0, 5 mM 2-mercapoethanol, 1 mM

PMSF/PTSF) and sonicated on ice using a 9.5 mm disruption horn connected to a Branson model 450 sonifier set at a constant duty cycle and a power output control of 7 for 30 s six times with 1 min cooling intervals in between. The cell lysate was then centrifuged at 10,000 x g for 10 min at 4°C and the resulting supernatant fractions were incubated with 5 ml of Ni-NTA agarose resin pre-equilibrated with column equilibration buffer (20 mM Tris HCl, pH 8.0, 100 mM KCl, 5 mM 2-mercaptoethanol, 10% v/v glycerol, and 20 mM imidazole), and incubated overnight at 4°C on a rotating wheel (speed set at 15 rpm). The slurry was loaded, 9 ml at a time, into a 0.8 x 4 cm Poly-prep chromatography column at 4°C. The bottom of the column was left open for the flow-through fraction to elute. The column was then washed with 25 ml of equilibration buffer, followed by 25 ml of high salt buffer (20 mM Tris HCl, pH 8.0, 1 M KCl, 5 mM 2-mercaptoethanol, and 10% v/v glycerol), and lastly with 5 ml of equilibration buffer. The resin was then incubated with 5 ml of elution buffer (20 mM Tris HCl, pH 8.0, 100 mM KCl, 5 mM 2-mercaptoethanol, 10% v/v glycerol, and 150 mM imidazole) for 30 min at 4°C. During this incubation period, the bottom of the column was sealed and the resin was mixed with the elution buffer every 5 min. The eluate was collected and another round of elution was repeated with 5 ml of elution buffer. The two eluate fractions were pooled, transferred to a 1.5 cm x 1 cm.ml<sup>-1</sup> dialysis membrane with a 6000 to 8000 MW cut off, and dialyzed twice against 4 L recombinant protein dialysis buffer (10 mM Tris HCl, pH 8.0, 400 µM PMSF/PTSF) at 4°C to remove the imidazole and other salts. The dialyzed samples were stored at -20°C until further use.

### **Yeast two-hybrid reporter assay**

To screen for protein-protein interactions, single patches of co-transformants from master plates were re-patched onto SCD-leu-trp plates containing 50 µg/ml 5-bromo-4-chloro-3-indolyl- beta-D-galactopyranoside (X-gal). These plates were incubated at 30°C for 5 to 7 days and visually checked for blue color formation. Expression of the *MEL1* gene, which occurs only if the two proteins being tested interact, encodes an  $\alpha$ -galactosidase that hydrolyzes the X-gal, resulting in blue color formation. As controls, patches of co-transformants expressing empty vectors and constructs provided by the manufacturer (Clontech) were tested for blue color formation. As an additional line of confirming interaction partners, a nutritional marker based screen was also performed. Patches from master plates were re-patched on SCD-leu-trp-his plates containing 10 mM 3-amino-1,2,4-triazole (3-AT) and incubated at 30°C for 5 to 7 days. The ability of patches to grow in the absence of histidine (his) would indicate expression of the *HIS3* gene, which is possible only if two proteins being tested for interaction truly interact in this system.

### **Generation of polyclonal antibodies against EutS and EutC**

Anti-EutS and -EutC polyclonal antibodies in rabbit were raised at Cocalico Biologicals, Inc., PA. About 2 mg of purified recombinant EutS and EutC proteins were separated on 12% polyacrylamide gels under denaturing conditions. After electrophoresis, gels were washed with deionized water three times, 10 min each, and stained with GelCode Blue stain (Thermo Scientific). The stained gels were washed overnight with deionized water and the protein bands excised using clean razor

blades. The excised bands were transferred to sterile 1.5 ml microfuge tubes and shipped on ice to Cocalico Biologicals.

## CHAPTER IV

### RESULTS

Experiments designed in this study were targeted to advance the current understanding of the participation of encapsulated enzymes in microcompartment biogenesis. For this purpose, protein interactions between the ethanolamine ammonia lyase enzyme and the putative shell components of the ethanolamine utilization organelles of *Salmonella enterica* were evaluated using a yeast two-hybrid screen. Further, the involvement of Form IA RubisCO in the assembly of carboxysomes was assessed by characterizing a *Halothiobacillus neapolitanus* mutant, wherein genes encoding the large and small subunits of the enzyme were deleted. Finally, the potential of *H. neapolitanus* carboxysomes to incorporate foreign RubisCO species was explored.

#### **I. Protein interactions in the Eut microcompartment of *Salmonella enterica***

##### **Five ethanolamine utilization (*eut*) genes encode carboxysome shell homologs**

Of the seventeen genes that comprise the *eut* operon, five (*eutK*, *-L*, *-M*, *-N*, and *-S*) were previously shown to encode homologs of carboxysome shell proteins [3, 4]. Based on sequence homology, the EutK, *-L*, *-M*, and *-S* proteins have been classified as belonging to bacterial microcompartment (BMC) proteins of the CcmK or CsoS1-type, while the EutN protein has been grouped into the CcmL or CsoS4A and CsoS4B (previously known as OrfA and OrfB)-type shell proteins [3]. A closer

look at their amino acid sequences reveals that the EutK and EutM proteins exhibit greater than 45% primary structure identity with the CsoS1 proteins (**Figure 4**). Although the amino acid sequence of the EutL protein is longer than that of the other putative Eut BMC proteins, it, too, harbors stretches of residues that share close to 35% identity with the CsoS1 proteins (**Figure 4**). The N- and C- termini of the EutL primary sequence have extensions that span beyond the BMC homology domains. A similar extension is also observed at the C-terminus of the EutK protein. In addition, non-BMC domain related interspersed amino acid stretches are found within the EutL and EutK sequences. The structural and functional importance of these extensions or stretches remains unknown. Compared to EutK and EutM, the EutS protein is less identical to the CsoS1 proteins (**Figure 4**) and its C-terminal portion along with that of EutL is found to be similar to PduA, a shell protein of the propanediol utilization (Pdu) organelles [3]. EutN possesses nearly 40% identity with the CsoS4A, and CsoS4B proteins (**Figure 5**) and has a four and seven amino acid extension at its N- and C-terminus, respectively. The extent of homology shared with carboxysome components suggests that the EutK, -L, -M, -N, and -S proteins have the potential to form shells of microcompartments.

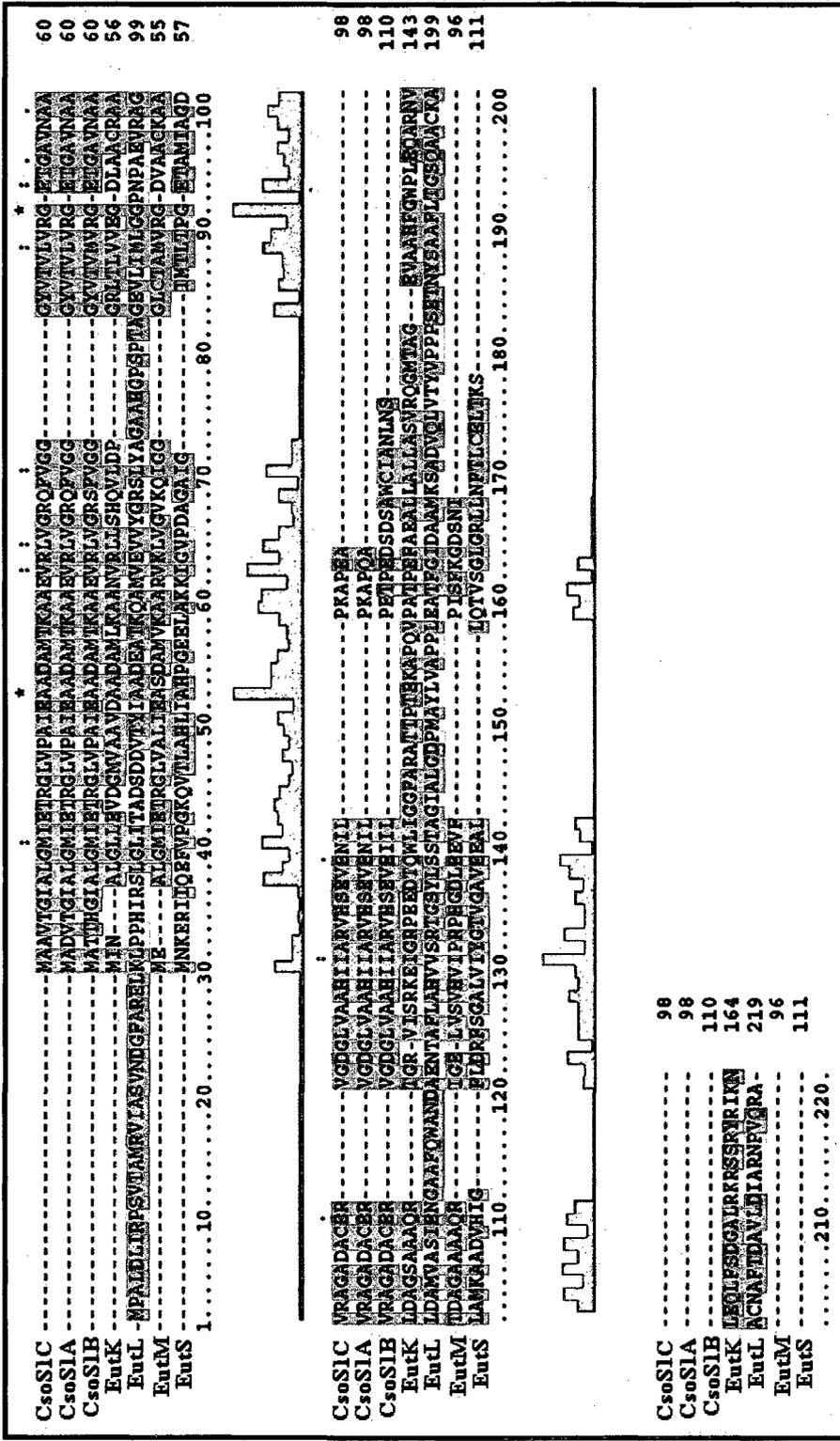


Figure 4. Primary sequence alignment of the CsoS1 and putative Eut shell proteins of *H. neapolitanus* and *S. enterica*, respectively

Amino acid sequences of the CsoS1C, -1A, and -1B proteins and the putative shell-forming EutK, -L, -M, and -S proteins of *S. enterica* were aligned using ClustalX 1.83. The *eutK*, -L, -M, and -S genes are believed to encode homologs of the CsoS1-type shell proteins.

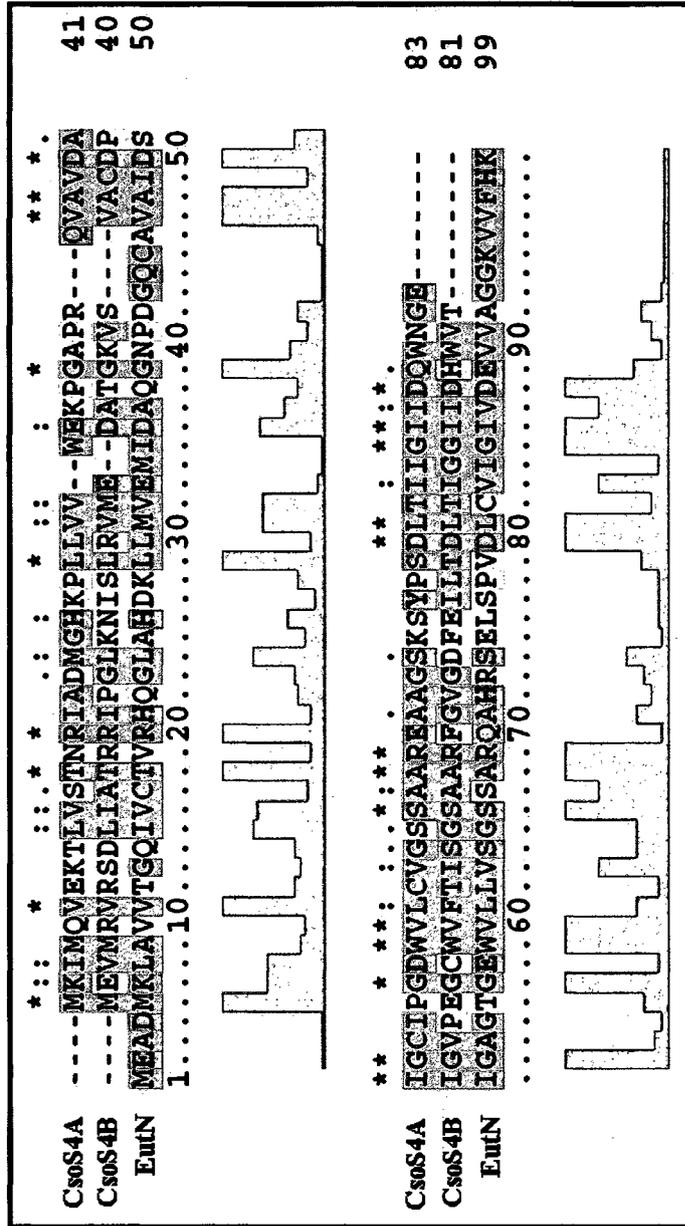


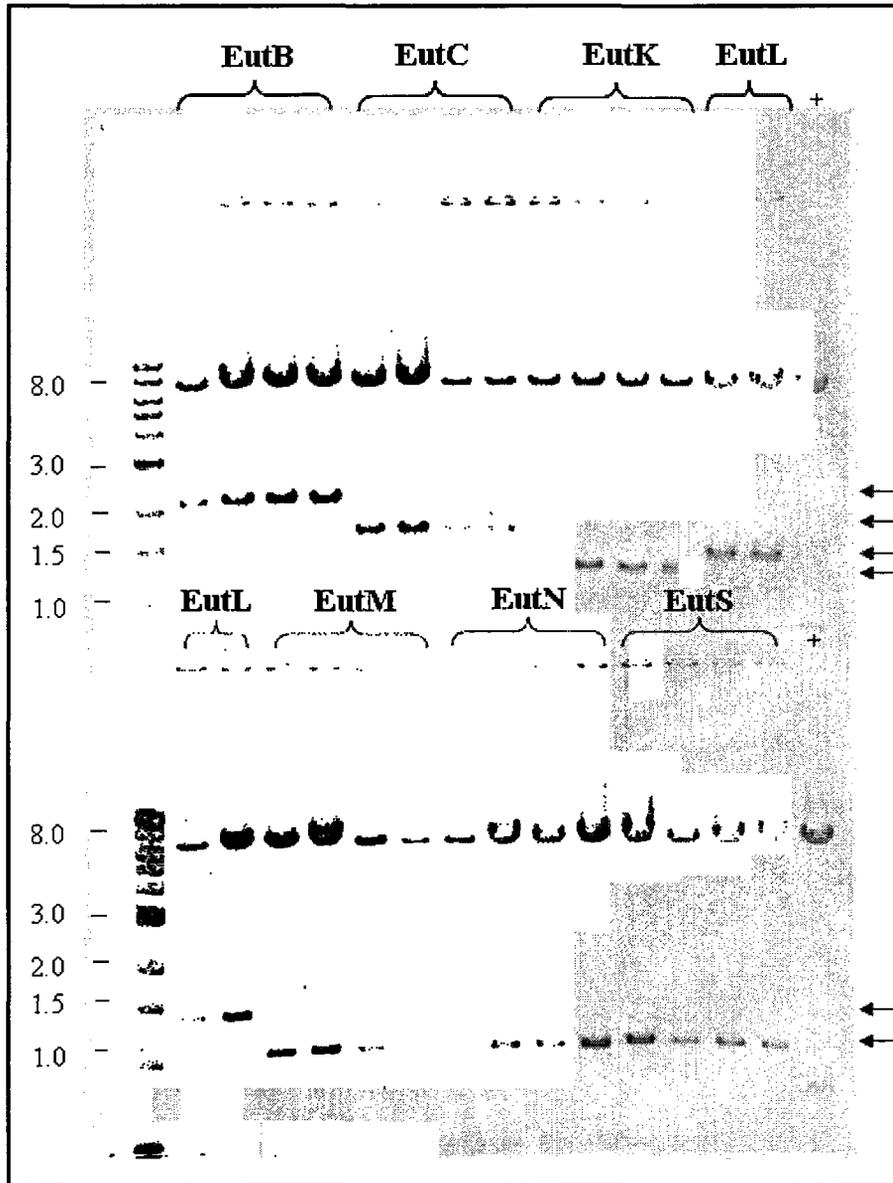
Figure 5. Primary sequence alignment of the CsoS4A and CsoS4B proteins (*H. neapolitanus*) and putative shell protein, EutN (*S. enterica*)  
 Amino acid sequences of the CsoS4A, CsoS4B, and the putative shell-forming EutN protein of *S. enterica* were aligned using ClustalX 1.83. The EutN protein is moderately homologous to the CsoS4A and CsoS4B-type shell proteins.

### **Interactions between the putative shell proteins EutN and EutL with ethanolamine ammonia lyase**

To identify interactions between the putative shell proteins (EutK, -L, -M, -N, and -S) and the heavy and light chains of the ethanolamine ammonia lyase enzyme (EutB, -C), a yeast two-hybrid screen [89] was performed. The pGBKT7 (bait) and pGADT7 (prey) vector-based expression constructs were purified from Top10 *E. coli* cells and verified by DNA sequencing and by analyses of restriction digests generated by the Hind III endonuclease. Upon digestion and separation by DNA gel electrophoresis, correct pGBKT7- and pGADT7-based constructs were identified by the presence of following insert DNA fragments: 2000 bp (*eutB*), 1650 bp (*eutC*), 1350 bp (*eutK*), 1500 bp (*eutL*), 1100 bp (*eutM*, -N), 1150 bp (*eutS*) (**Figure 6**) and 2200 bp (*eutB*), 1700 bp (*eutC*), 1250 bp (*eutK*), 1400 bp (*eutL*), 1100 bp (*eutM*, -N), 1150 bp (*eutS*) (**Figure 7**).

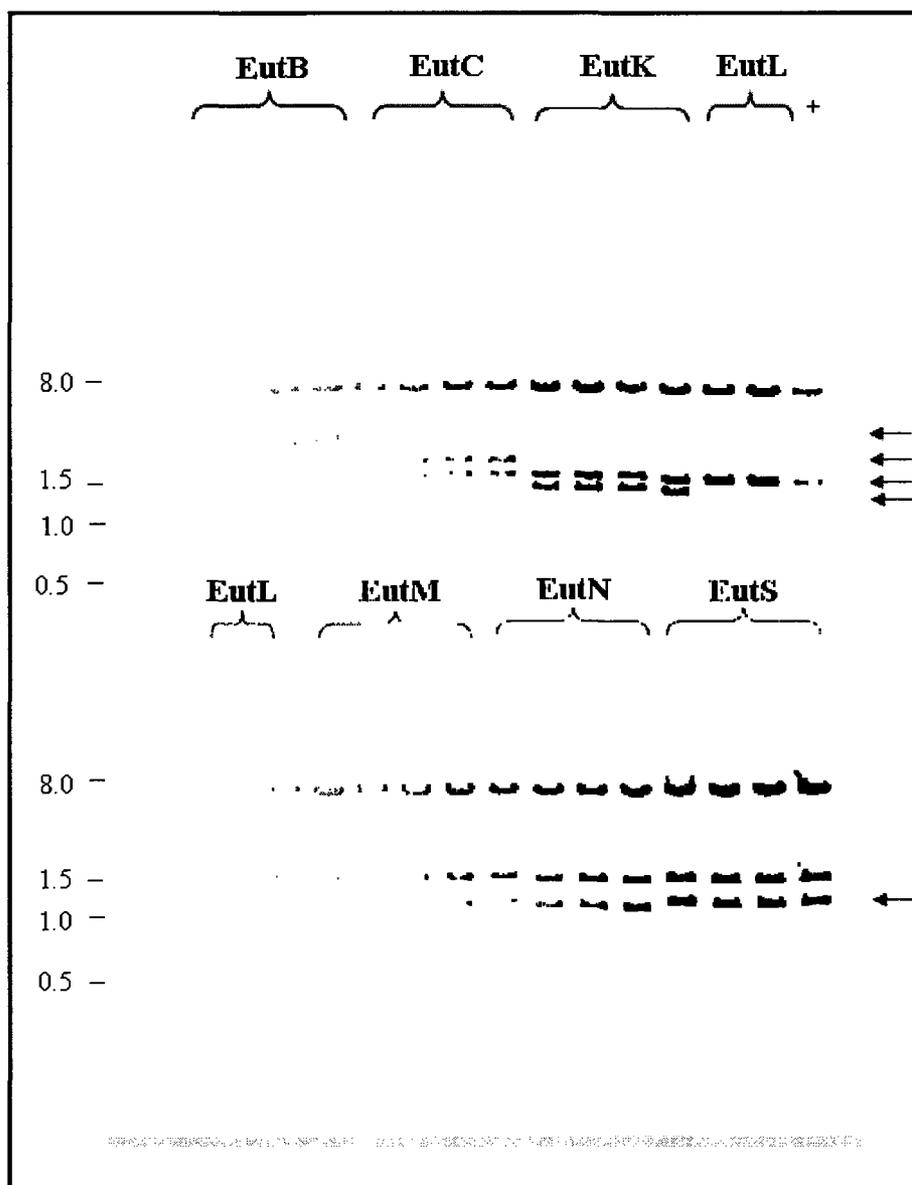
Protein interactions were identified in two ways. First, transformants harboring either bait or prey constructs were tested for their ability to auto-activate the *MEL1* gene, the expression of which results in the formation of blue colonies (**Figure 8**). In this study, no such auto-activating transformants were encountered. Second, vector swapping was done such that each Eut protein could be tested independently as bait and prey. As an additional verification measure, co-transformants harboring potential interacting protein pairs were tested for their ability to grow on triple selective medium (SCD-leu-trp-his) plates containing 10 mM 3-amino-1,2,4-triazole (3-AT). 3-AT is a competitive inhibitor of the *HIS3* gene product, imidazoleglycerol-phosphate dehydratase [90], and its presence in the

medium eliminates background colony formation by non-interacting protein pairs. On the other hand, protein pairs that interact cause high levels of His3 protein expression, which overcomes the inhibitory effects of 3-AT and induces colony formation on triple selective medium containing the inhibitor. In this study, all blue colony-forming co-transformants were able to grow on SCD-leu-trp-his plates and 3-AT-containing solid medium, indicating that they harbored legitimate interacting protein partners.



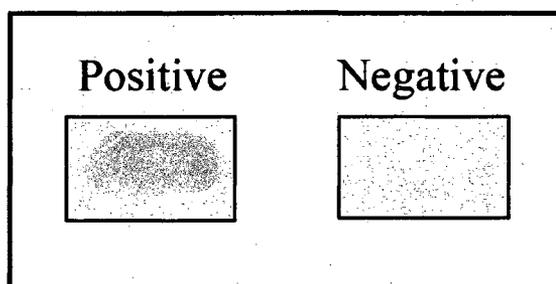
**Figure 6. Identification of pGADT7-*eutB*, -*C*, -*K*, -*L*, -*M*, -*N*, and -*S* ‘prey’ constructs via restriction digest analysis**

Plasmid DNA from four different clones for each *eut* gene construct was digested with the Hind III restriction endonuclease. The lane labeled ‘+’ represents the control pGADT7 vector digested with the same enzyme. Arrows indicate the expected size of the insert released after enzyme digestion. Digests were separated on a 0.8% agarose/TBE gel, stained with ethidium bromide and visualized on a UV transilluminator. DNA molecular weight standards in kbp are indicated on the left of the gel picture.



**Figure 7. Identification of pGBKT7-*eutB*, -*C*, -*K*, -*L*, -*M*, -*N*, and -*S* 'bait' constructs via restriction digest analysis**

Plasmid DNA from four different clones for each *eut* gene construct was digested with the Hind III restriction endonuclease. The lane labeled '+' represents the control pGBKT7 vector digested with the same enzyme. Arrows indicate the expected size of the insert released after enzyme digestion. Digests were separated on a 0.8% agarose/TBE gel, stained with ethidium bromide and visualized on a UV transilluminator. DNA molecular weight standards in kbp are indicated on the left of the gel picture.



**Figure 8. Examples of interacting and non-interacting protein pairs**

Wild type yeast cells were co-transformed with bait and prey plasmids that encoded interacting protein pairs [**Positive**] and non-interacting proteins [**Negative**]. Interacting protein pairs activated expression of the *MEL1* gene and secretion of  $\alpha$ -galactosidase, which hydrolyzed the 5-bromo-4-chloro-3-indolyl- $\alpha$ -D-galactopyranoside (*X- $\alpha$ -Gal*) substrate contained in the medium, resulting in the formation of blue colonies. Non-interacting proteins that failed to activate the *MEL1* gene resulted in the formation of white colonies. Strong interactions, like that shown in the left panel were scored as ‘+++.’ No interaction was scored as ‘-.’

Strong (+++) homo-protein interactions were observed between the EutL, EutN, and EutS proteins when they were used as bait and prey (**Figure 9**). A similar interaction grid of the *H. neapolitanus* carboxysome proteins showed a comparable homo-protein interaction pattern between the CsoS1A and CsoS1B and the CsoS4A and CsoS4B proteins [14], which are homologous to the EutL/EutS and EutN proteins, respectively. The other CsoS1 homologs, EutK and EutM, exhibited moderate (++) homo-protein interactions. The ability of these putative shell proteins to self-interact suggests that their individual protomers might have a tendency to oligomerize, which is thought to be a requirement for microcompartment self-assembly [12]. Moderate homo- and hetero-protein interactions were also observed between the individual subunits of the glutamine ammonia lyase (EAL) enzyme (**Figure 9**). The most pronounced hetero-protein interactions identified were those between the shell protein EutN and EA lyase large subunit, EutB. This pattern, too, was identical to the CsoS4B and CbbL interactions observed in similar screens

involving carboxysome proteins of *H. neapolitanus* [14]. The EutL protein was found to have strongest interactions with the EutB, -C, -K, and -N proteins when used as bait, while the EutN protein was found to have strongest interactions with EutB, -K, -L, and -M proteins when used as prey. The EutK protein showed moderate hetero-protein interactions with other shell proteins and the EAL large subunit, EutB. Minimal hetero-protein interactions were observed for the EutM and EutS proteins (Figure 9).

Results from the yeast two-hybrid screen suggested that the EA lyase large subunit, EutB, makes strong contacts with the EutN and EutL shell components. This is in contrast to the interactions observed in yeast two-hybrid screens of *H. neapolitanus* carboxysomal proteins, wherein the small subunit of RubisCO, CbbS, was found to associate extensively with the CsoS1 and CsoS4 shell components [14]. Based on the results from this study, EutL and EutN may represent essential shell proteins involved in microcompartment assembly.

Although the yeast two-hybrid screens are widely used as a tool for identifying *in vivo* protein-protein interactions, false positive interactions are potential outcomes. In some cases, high expression of bait and prey fusion proteins may drive non-physiological interactions, while in some others, the fusion protein might be unstable or improperly folded, thereby masking essential interaction sites. To substantiate results from the yeast two-hybrid screens, other *in vitro* methods for testing protein-protein interactions such as affinity-based pull-down assays or co-immunoprecipitations will have to be performed, which would require large amounts of recombinant Eut proteins and antibodies raised against them.

B A I T	P R E Y									
	EutB	EutC	EutK	EutL	EutM	EutN	EutS	pGADT7-T (+) AD		
EutB	++	++	++	++	+	+++	++	ND		
EutC	++	++	+	+	-	++	-	ND		
EutK	++	+	++	++	+	+++	+	ND		
EutL	+++	+++	+++	+++	++	+++	++	ND		
EutM	-	-	+	-	++	+++	+	ND		
EutN	+++	++	++	++	+	+++	-	ND		
EutS	++	+	-	-	+	+	+++	ND		
pGBKT7-p53 (+)BD	ND	ND	ND	ND	ND	ND	ND	ND	++	++
pGBKT7-lam (-) BD	ND	ND	ND	ND	ND	ND	ND	ND	-	-

**Figure 9. Protein-protein interaction grid of the putative Eut shell proteins and EAL**

Interactions between protein pairs were scored as strong (+++), moderate (++), weak (+) or none (-) based on the criteria described in Figure 8. EutK, -L, -M, -N, and -S comprise putative shell proteins while EutB, and -C form the large and small subunits of the EAL enzyme, respectively. The pGBKT7-53 (+)BD, pGBKT7-lam (-) BD, and pGADT7-T (+) AD vectors, provided by the manufacturer, encode proteins that serve as positive and negative controls for the yeast two-hybrid screen. ND = Proteins interactions that were not determined.

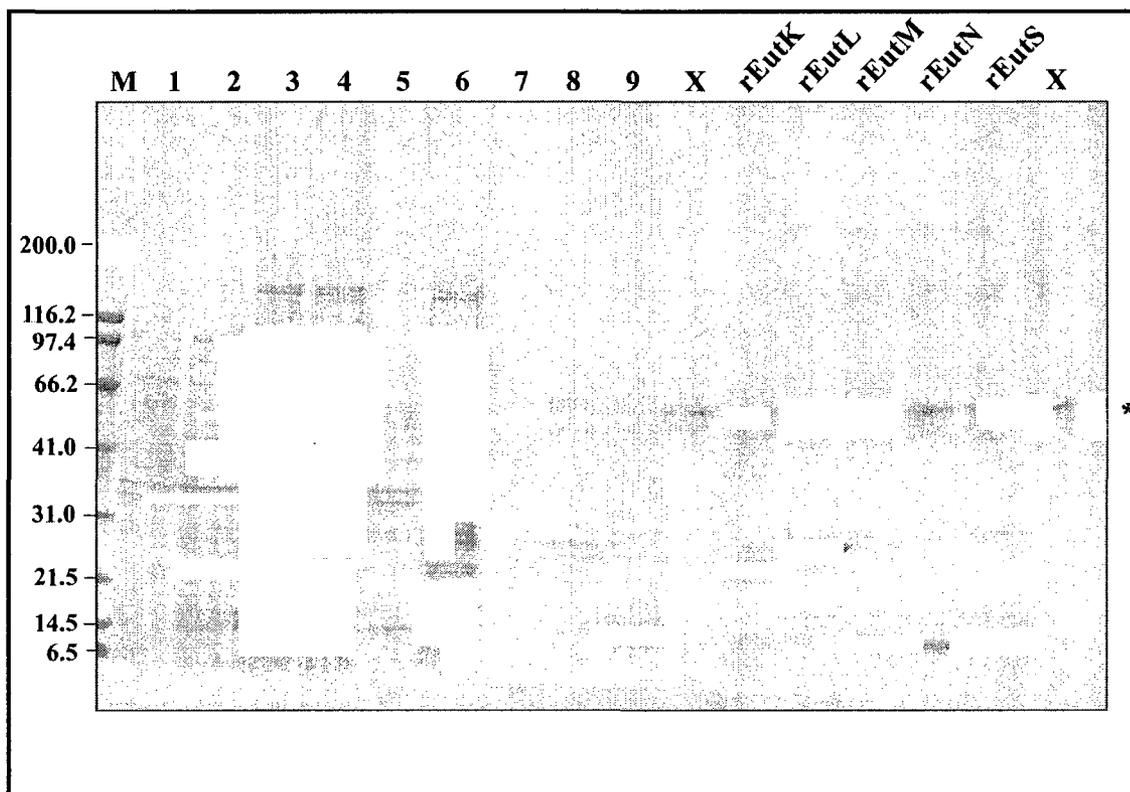
### **Purification of recombinant EutK, -L, -M, -N, -S, and -C proteins**

The recombinant EutK, -L, -M, -N, -S, and -C proteins were purified using the pPROEX-HTb prokaryotic expression system, which exploits a His<sub>6</sub> tag-based affinity purification protocol. This system allows for a one-step purification of recombinant protein on a Ni-NTA resin-packed column. Optimal expression of the EutK, -L, -M, -N, -S, and -C proteins was tested by performing a small scale induction experiment at various temperatures (16 °C, 30°C and 37°C). In all cases, maximal protein expression was found to occur at 37°C after induction with IPTG for 3 h. The recombinant proteins appeared to be approximately equally distributed between the supernatant and pellet fractions obtained after centrifuging the sonicated cell lysate (**Figures 10 and 11**). The supernatant was used as starting material for the Ni-NTA affinity-based protein purification, which typically yielded 6 to 8 mg of recombinant protein from approximately 4 g of wet cell pellet.

### **Generation of anti-EutS and anti-EutC recombinant antibodies**

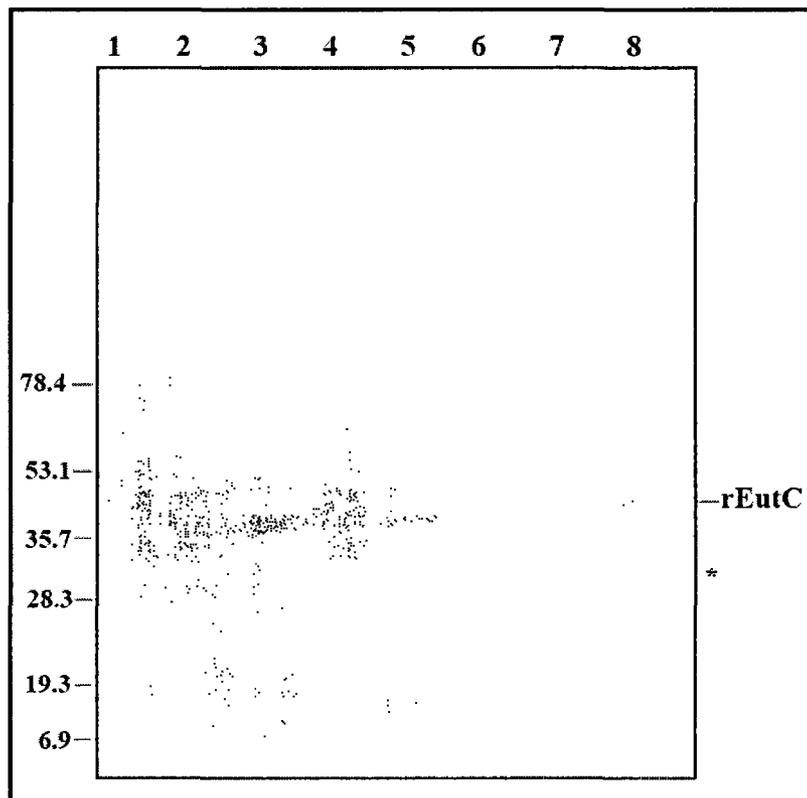
To generate antibodies against the EutS and EutC proteins, pre-immune sera from three different rabbits were initially tested for cross-reactivity with proteins from cell extracts of ethanolamine-grown *S. enterica* or purified recombinant EutS or EutC proteins. The rabbits whose pre-immune serum showed least cross-reactivity were chosen for injection with the recombinant proteins. Representative blots probed with pre-immune serum and serum obtained eight weeks after injecting the host rabbit are shown in **Figures 12 and 13**. The anti-EutS antibodies failed to recognize

endogenous EutS protein in extracts of ethanolamine-grown *S. enterica* cells, while the anti-EutC antibodies recognized recombinant as well as endogenous EutC protein.



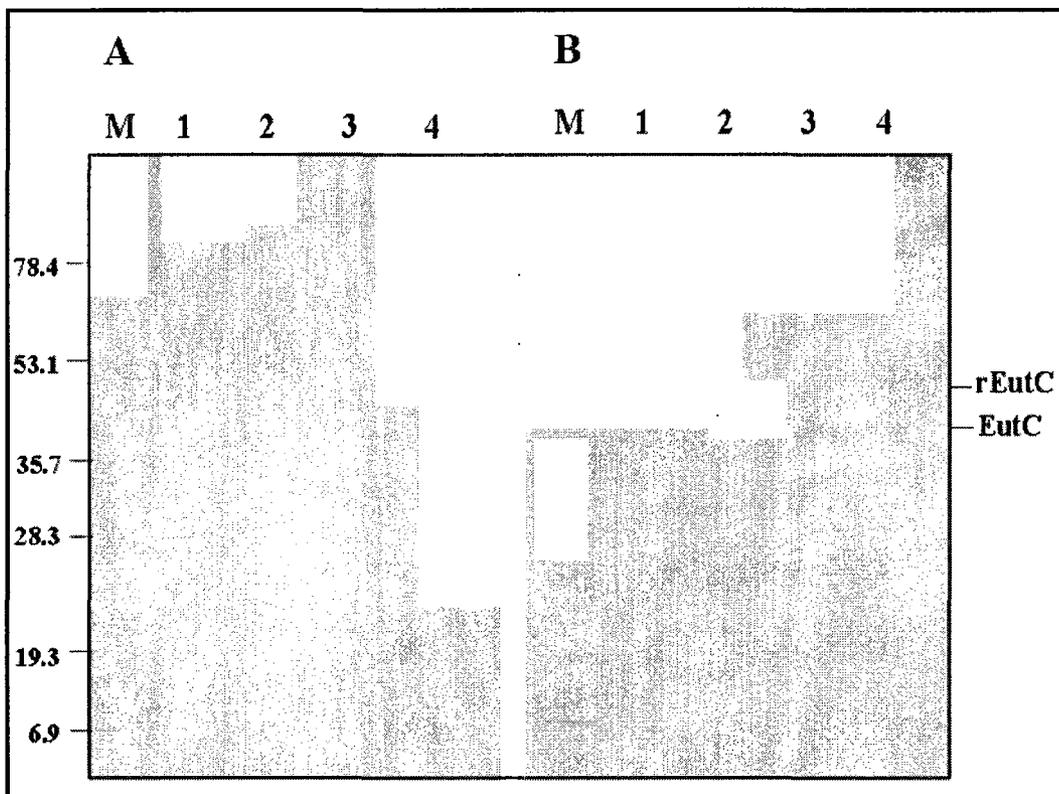
**Figure 10. Purification of recombinant His<sub>6</sub>-tagged Eut -K, -L, -M, -N, and -S proteins**

Samples from the different fractions obtained during the purification of recombinant EutM protein, along with purified EutK, -L, -M, -N, and -S proteins were separated on a denaturing 4-20% SDS-polyacrylamide gradient gel and stained with GelCode Blue. The band labeled with an asterisk appears to be a gel artifact, since it is observed in lanes that do not contain any protein. Protein molecular weight standards in kDa are indicated on the left of the gel. (1 = uninduced, 2 = induced, 3 = sonicated cell lysate, 4 and 5 = supernatant and pellet fractions resulting from centrifugation of the sonicated cell lysate at 10,000 x g for 10 min, 6 = flow-through after overnight binding of supernatant to Ni-NTA beads, 7 = final wash, 8 = eluate 1, 9 = eluate 2, X = blank lanes)



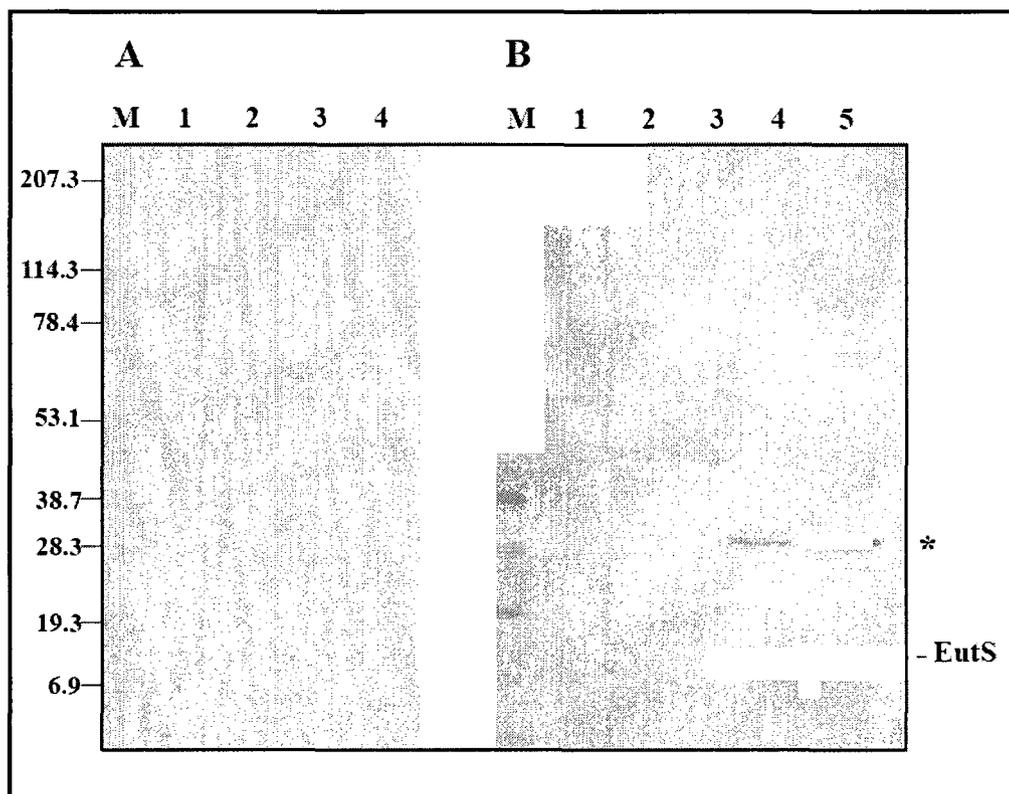
**Figure 11. Purification of recombinant His<sub>6</sub>-tagged EutC protein**

Samples from the different fractions obtained during the purification of recombinant EutC protein were separated on a denaturing 4-20% SDS-polyacrylamide gradient gel and stained with GelCode Blue. The band labeled with an asterisk may be a degradation product of EutC or an *E. coli* protein contaminant. Protein molecular weight standards in kDa are indicated on the left of the gel. (1 = uninduced sample, 2 = induced sample, 3 and 4 = supernatant and pellet fractions resulting from centrifugation of the sonicated cell lysate at 10,000 x g for 10 min, 5 = flow-through after overnight binding of supernatant to Ni-NTA beads, 6 = final wash, 7 = eluate 1, 8 = eluate 2)



**Figure 12. Western blotting with pre-immune serum and serum containing anti-EutC antibodies**

Purified recombinant EutC (lanes 1 and 2 containing 100 ng and 200 ng protein, respectively) and extracts of *S. enterica* cells grown in EA-containing medium (lanes 3 and 4 containing 50 µg and 25 µg protein, respectively) were probed with [A] pre-immune serum and [B] serum containing anti-EutC antibodies. Protein molecular weight standards (lane M) in kDa, are shown on the left of each blot.

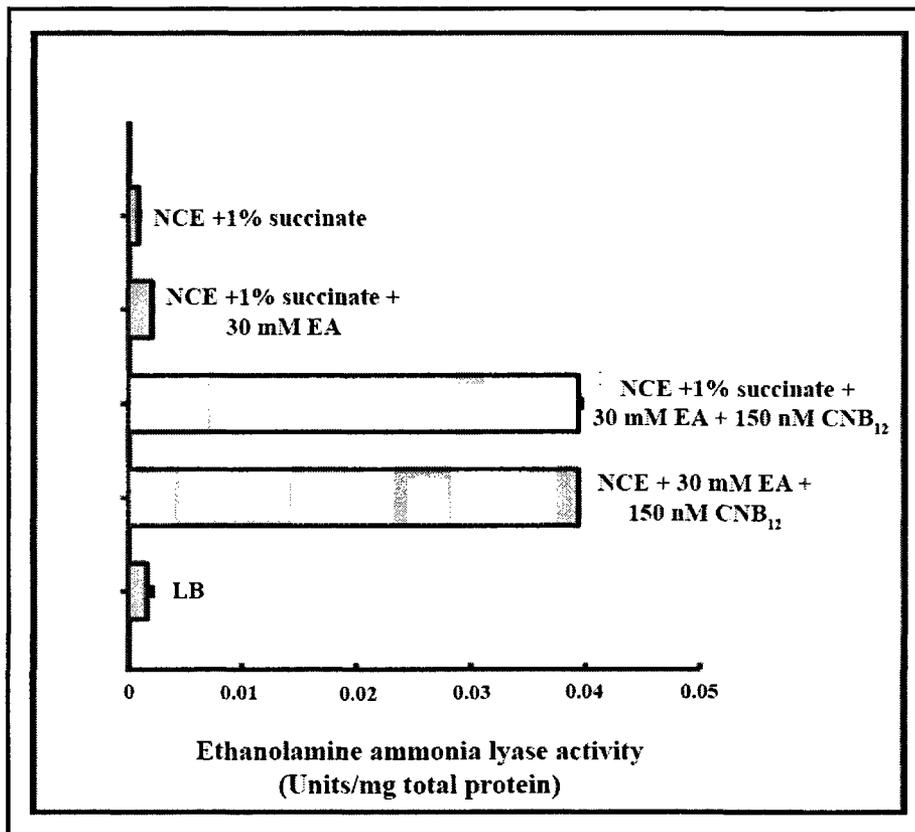


**Figure 13. Western blotting with pre-immune serum and serum containing anti-EutC antibodies**

Extracts of *S. enterica* cells grown in ethanolamine-containing medium (lanes 1 and 2 containing 25  $\mu$ g and 50  $\mu$ g protein, respectively) and purified recombinant EutS (lanes 3 and 4 containing 100 ng and 200 ng protein, respectively) were probed with pre-immune serum [A] and extracts of *S. enterica* cells grown in LB medium (lane 1, 25  $\mu$ g protein), ethanolamine-containing medium (lanes 2 and 3 containing 25  $\mu$ g and 50  $\mu$ g protein, respectively) and purified recombinant EutS (lanes 4 and 5 containing 100 ng and 200 ng protein, respectively) were probed with serum containing anti-EutS antibodies [B]. The band labeled with an asterisk may represent an oligomeric form of the EutS protein. Protein molecular weight standards (lane M) in kDa, are shown on the left of each blot.

### **Determination of ethanolamine ammonia lyase (EAL) activity in *S. enterica* extracts**

Since the anti-EutS antibodies generated in this study failed to recognize the endogenous EutS shell protein of the Eut microcompartments, the activity of the EAL enzyme in *S. enterica* cells grown in different carbon sources was determined to evaluate expression of the *eut* operon. The EAL enzyme converts ethanolamine to acetaldehyde and ammonia in a cobalamin-dependent manner [3, 79]. The amount of acetaldehyde formed from ethanolamine is measured spectrophotometrically at 305 nm [88]. In this study, EAL activity was observed only when cells were grown in cyanocobalamin-containing NCE medium (**Figure 14**). This result was not surprising since cyanocobalamin is known to be an important cofactor of *eut* operon expression under aerobic growth conditions [78]. Hence, NCE medium containing 30 mM ethanolamine, 1% v/v succinate, and 150 nM cyanocobalamin was used to grow *S. enterica* cells and induce Eut microcompartment formation.



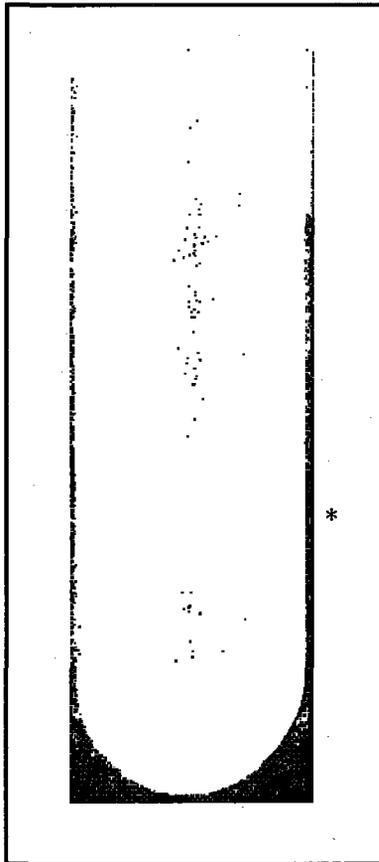
**Figure 14. EAL activity in extracts of *S. enterica* cells grown in different carbon sources**

Extracts from *S. enterica* cells grown in different carbon sources, as indicated, were assayed for EAL activity. The assay was performed in triplicates and the error bars represent the standard error of the mean. Maximum activity was associated with extracts obtained from cells grown in NCE medium containing ethanolamine (EA) and/or succinate and 150 nM cyanocobalamin (CNB<sub>12</sub>).

### **Purification of the Eut microcompartments of *S. enterica***

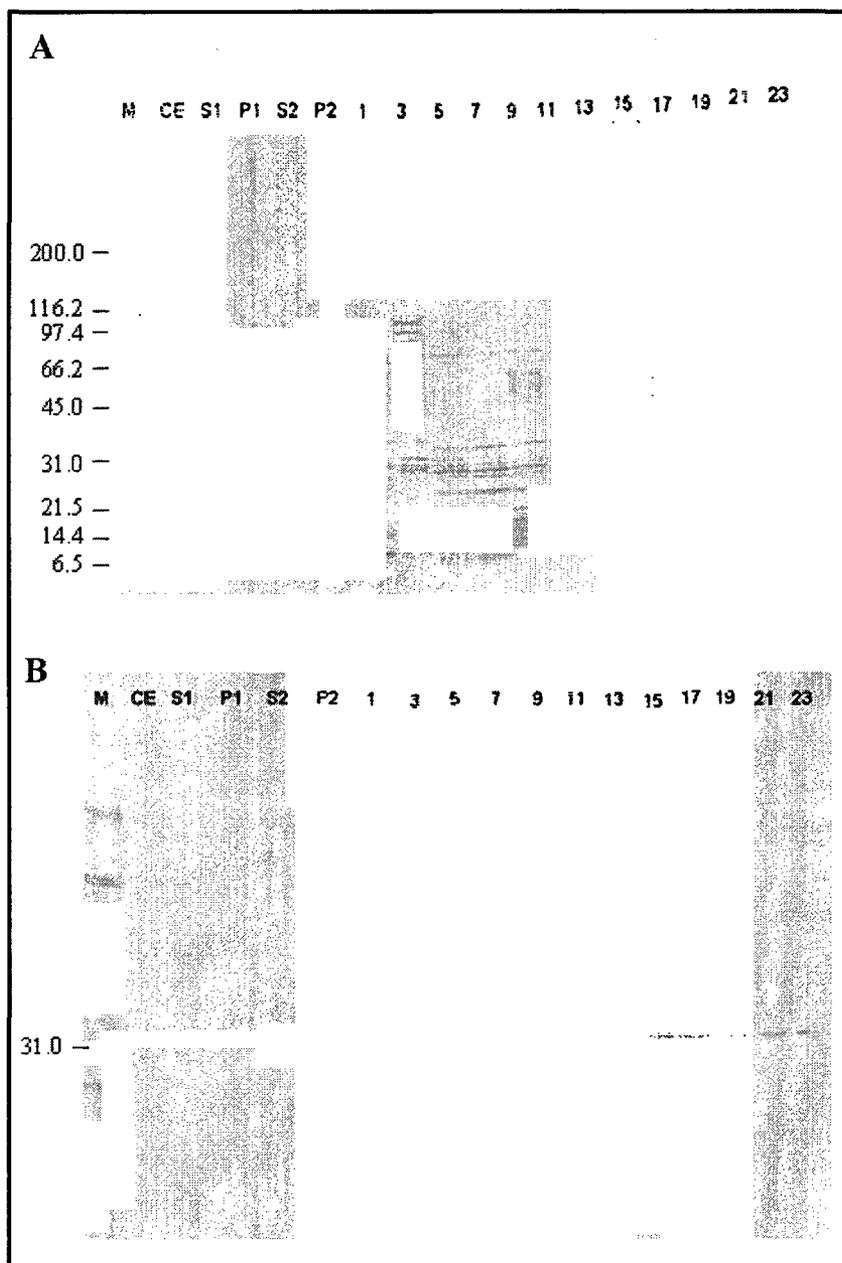
To purify the Eut microcompartments of *S. enterica*, the method of Havemann and Bobik [19] was used, which involves sonication and detergent treatment of cells combined with differential and density gradient centrifugation. To assess the progress of the purification procedure, the various fractions were subjected to EAL assays, SDS-PAGE, anti-EutC western blotting, and electron microscopy analyses. The sucrose density gradient centrifugation step of the purification protocol yielded a white translucent band two thirds of the way down the tube (**Figure 15**), which was thought to be enriched for the Eut microcompartments because a band that migrated to a similar position on a gradient was observed by Havemann and Bobik [19] during the purification of the propanediol utilization microcompartments. The gradient was divided into 25 x 1.5 ml fractions, top to bottom. Each fraction was assayed for EAL activity and probed for the EutC protein by western blotting. Fraction 15 that corresponded to the band on the sucrose gradient exhibited maximum EAL activity with a matching immunoblot signal of the EutC protein (**Figures 16B and 17**). However, SDS-PAGE analysis revealed no significant difference between the protein compositions of crude cell extract and fraction 15, suggesting that the latter may not be enriched for the Eut microcompartments (**Figure 16A**). Although electron microscopy revealed few polyhedra, which seemed to be 50 to 100 nm in diameter, the majority of fraction 15 was found to contain membrane fragments and flagella-like structures (**Figure 18**). Since antibodies that would recognize endogenous Eut BMC proteins could not be generated, the presence of microcompartments could not be assessed by probing for Eut shell components.

Since the Eut organelles have not yet been purified to date, the polypeptide composition of these polyhedra remains unknown. Based on results from the yeast two hybrid screens, one can only speculate about protein interactions that may be involved in the assembly of Eut microcompartments. To address the role of encapsulated enzymes in microcompartment assembly, a more direct *in vivo* approach was employed by constructing a Form IA RubisCO deletion mutant of the model chemoautotrophic bacterium, *Halothiobacillus neapolitanus*.



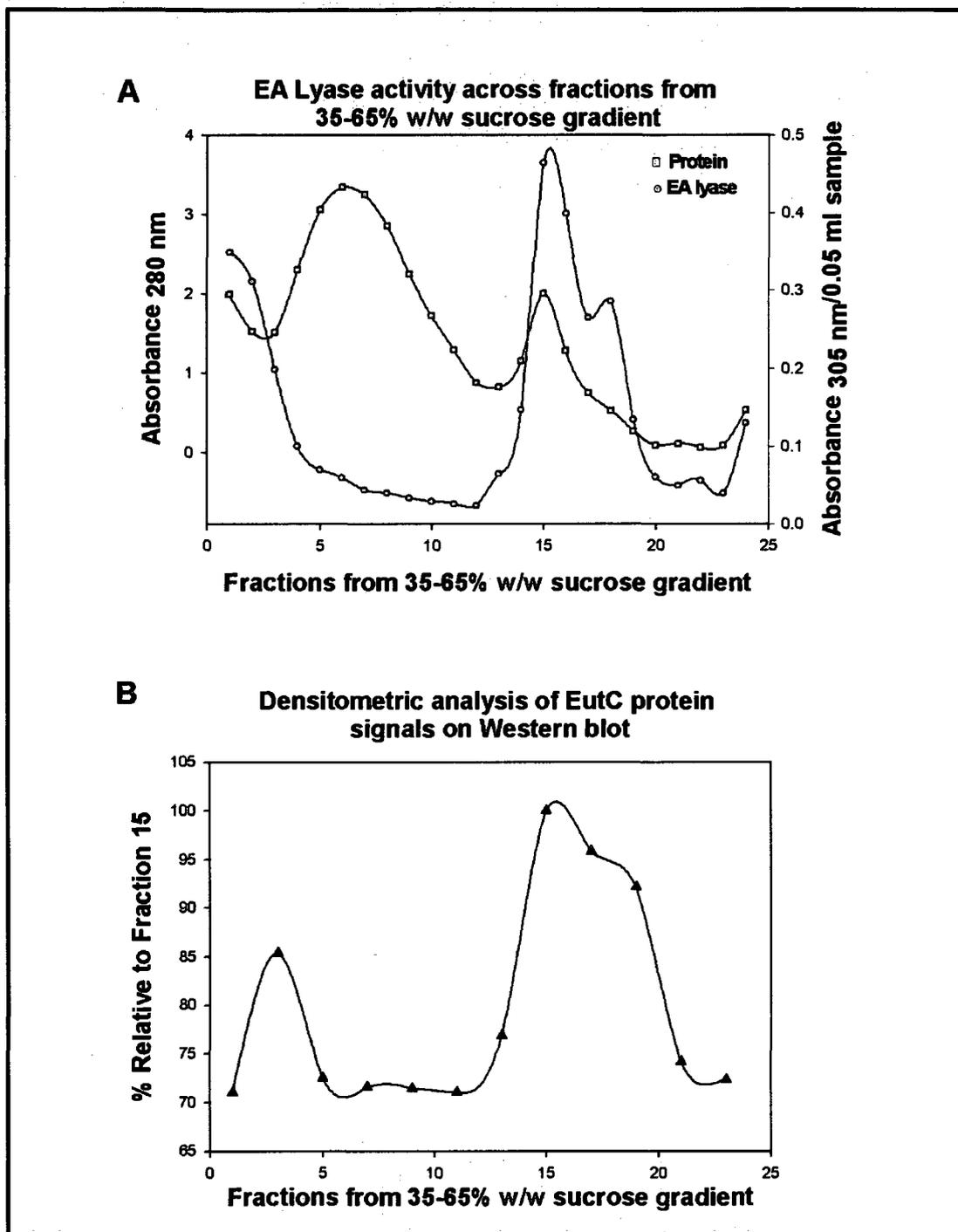
**Figure 15. Band enriched for EAL separated on a 35-65% w/w sucrose density gradient**

The pellet obtained from the 48,000 x g spin of the purification protocol was loaded onto a 35-65% w/w sucrose density gradient and spun at 104,000 x g for 12 hours at 4°C. The gradient was divided into 1.5 ml fractions and subjected to activity assays and SDS-PAGE and anti-EutC western blotting analyses. The band labeled with an asterisk was thought to be enriched for the Eut microcompartments, based on the level of EAL activity and matching immunoblot signal of the EutC protein in this fraction.



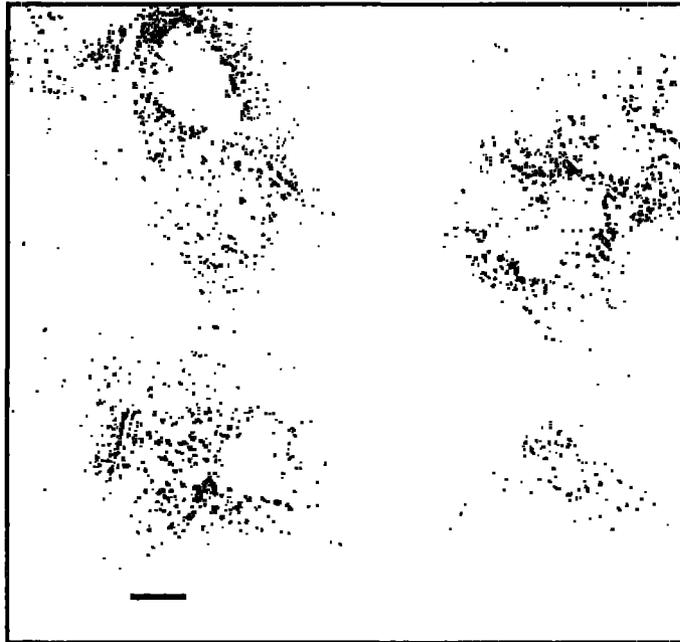
**Figure 16. SDS-PAGE and western blot analyses of fractions from Eut microcompartment purification steps**

Samples from different fractions obtained during the purification of Eut microcompartments were separated on a 4-20 % SDS-polyacrylamide gradient gel and stained with GelCode Blue [A]. A blot of an identical gel was probed with anti-EutC antibodies [B]. The sample in lane 15 represents the band observed in the sucrose gradient shown in **Figure 15**. Protein molecular weight standards in kDa are indicated on the left of the gel. (CE = cell extract, S1 and P1 = supernatant and pellet fractions from 12,000 x g spin, respectively, S2 and P2 = supernatant and pellet fractions from 48,000 x g spin, respectively, 1 to 23 = fractions from the sucrose gradient)



**Figure 17. EAL activity and densitometry of fractions from 35-65% w/w sucrose gradient**

Fractions 1-25 from the 35-65% w/w sucrose gradient were assayed for EAL activity and total protein [A]. These fractions were separated by SDS-PAGE (Figure 16A), electroblotted and probed with anti-EutC antibodies (Figure 16B). Signals corresponding to the EutC protein on the immunoblot (Figure 16B) were subjected to densitometric analysis using Biorad's Quantity One v 4.6.3 software [B].



**Figure 18. Electron microscopy of the fraction enriched for EAL**

Fraction 15 from the sucrose gradient that showed enrichment for ethanolamine ammonia lyase activity was harvested, centrifuged, and resuspended in TEME buffer. A sample was stained with 1% w/v ammonium molybdate and visualized by electron microscopy at a magnification of 25,000 X. The black bar represents a scale of 50 nm. Polyhedral microcompartments, indicated by arrows, could be seen in this fraction, along with other contaminants such as membrane fragments and flagella-like structures.

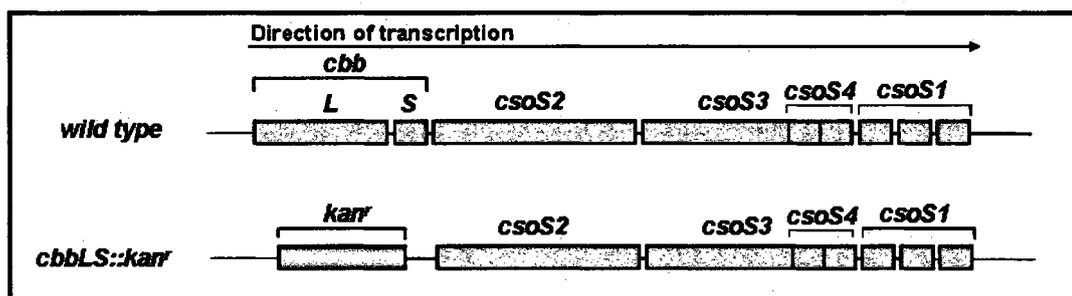
## II. Evaluation of the role of Form IA RubisCO in carboxysome biogenesis

The involvement of Form IA RubisCO in carboxysome biogenesis was assessed in the chemoautotrophic bacterium, *Halothiobacillus neapolitanus*, which is known to possess a robust homologous recombination system that facilitates mutant generation [46]. Most importantly, carboxysomes have been purified to homogeneity and characterized from this bacterium [36, 42].

### Generation of a *Halothiobacillus neapolitanus* Form IA RubisCO deletion mutant

A Form IA RubisCO deletion mutant, *cbbLS::kan<sup>r</sup>*, was constructed by electroporating exponentially growing wild type *H. neapolitanus* cells with a pUC18 vector-based plasmid that harbored a kanamycin (*kan<sup>r</sup>*) cassette designed to replace the enzyme's large and small subunit-encoding *cbbL* and *cbbS* genes in the *cso* operon, respectively (**Figure 19**). For selection purposes, electroporated cells were passaged several times in liquid medium containing kanamycin and incubated at 30°C in air supplemented with 5% CO<sub>2</sub>. The rationale for using pUC18 vector-based plasmids for electroporation was a previous study that had reported the inability of *H. neapolitanus* cells to stably maintain non-thiobacillus replicon-based vectors beyond fifty generations [46]. Genomic integration of the *kan<sup>r</sup>* cassette within the *cso* operon occurred via homologous recombination at the 5' upstream and 3' downstream regions of the wild type *H. neapolitanus cbbL* and *cbbS* sequences, respectively. To ensure that the deletion mutant had lost the pUC18 plasmid following homologous recombination, transformants were initially plated on solid medium containing

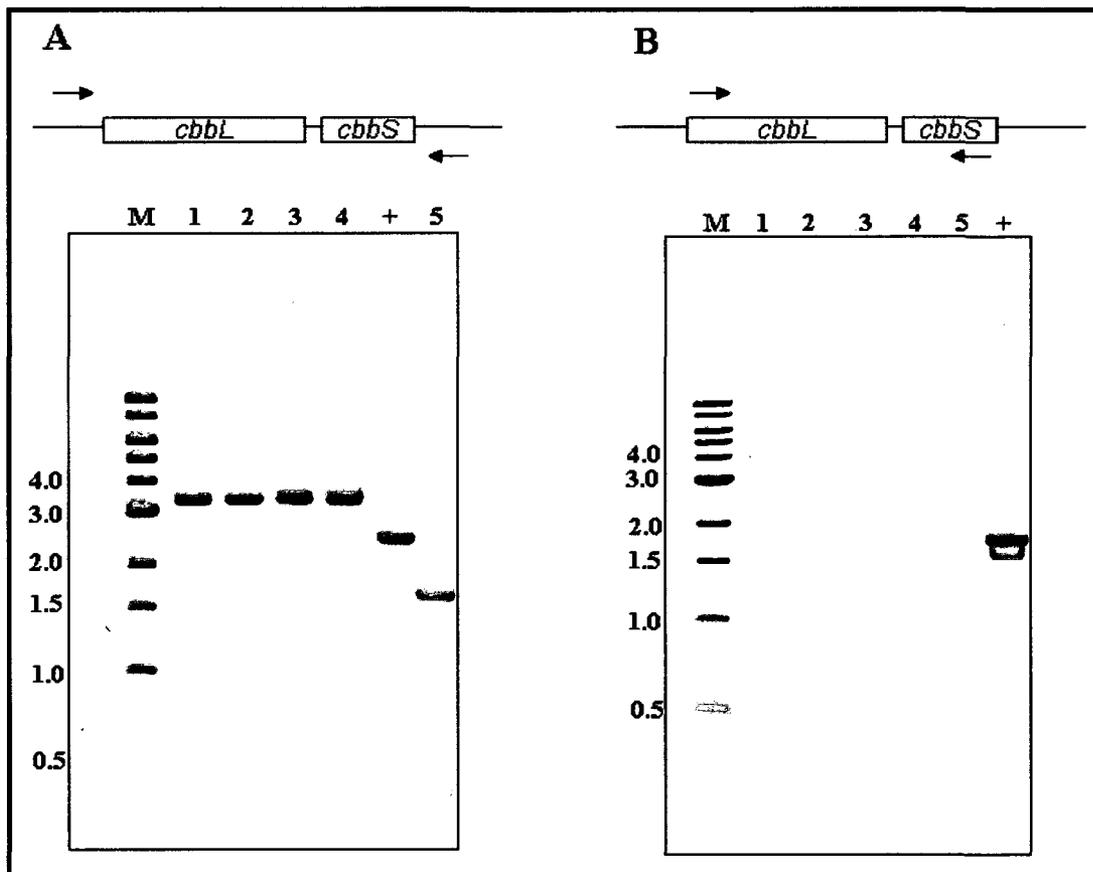
kanamycin. On average, twenty five isolated colonies were patched onto kanamycin-containing and ampicillin-containing solid medium and their ability to grow on these plates was monitored. Growth on ampicillin-containing solid medium would indicate that the transformants still retained the pUC18 plasmid. After 72 h of incubation, no growth was observed on the ampicillin-containing plates suggesting that the mutant had lost the pUC18 plasmid. However, the mutant could grow on kanamycin-containing medium due to the *kan'* gene used for selection purposes (Figure 19). To verify the absence of the plasmid-borne *amp'* marker in the mutants, genomic DNA from these mutants was used to transform Top10 *E. coli* competent cells. The transformation mixture was plated on LB + ampicillin, LB + kanamycin, and LB + ampicillin + kanamycin plates. No transformants were observed on any of these plates, strongly indicating that the *H. neapolitanus* mutants had no longer retained any plasmid DNA capable of replicating in *E. coli* cells.



**Figure 19. Organization of genes within the *cso* operon of the Form IA RubisCO deletion mutant**

The deletion mutant, *cbbLS::kan'*, was constructed by replacing the *H. neapolitanus* Form IA RubisCO large (*cbbL*) and small (*cbbS*) subunit-encoding genes with a kanamycin (*kan'*) cassette (white box).

To confirm the deletion of the RubisCO-encoding genes in the *cbbLS::kan<sup>r</sup>* mutant, polymerase chain reactions (PCR) were performed. Genomic DNA isolated from wild type and mutant cells was used as template. Primer pairs that would anneal approximately 300 bp upstream and downstream of the *cbbL* and *cbbS* regions, respectively, within the *cso* operon of *H. neapolitanus* were chosen (**Figure 20A**). Upon amplification and separation by DNA gel electrophoresis, the wild type reaction showed an approximately 2500 bp band, which represents the endogenous *cbbLS* genes. The reaction from the deletion mutant showed an approximately 1600 bp band, which indicated that the *cbbLS* genes were replaced with the *kan<sup>r</sup>* cassette (**Figure 20A**, lane 5). To ensure that the mutant did not harbor any endogenous Form IA RubisCO-encoding genes in its genome, another set of amplification reactions was performed using primer pairs that would anneal within the coding regions of the endogenous *cbbL* and *cbbS* sequences (**Figure 20B**). The wild type control reaction showed a band of approximately 1750 bp band, while the reaction from the deletion mutant showed no product (**Figure 20B**, lane 5). In addition to PCR amplification, integration of the *kan<sup>r</sup>* cassette at the endogenous *cbbL-cbbS* locus within the *cso* operon was confirmed by genomic DNA sequencing (University of Maine DNA Sequencing Facility).



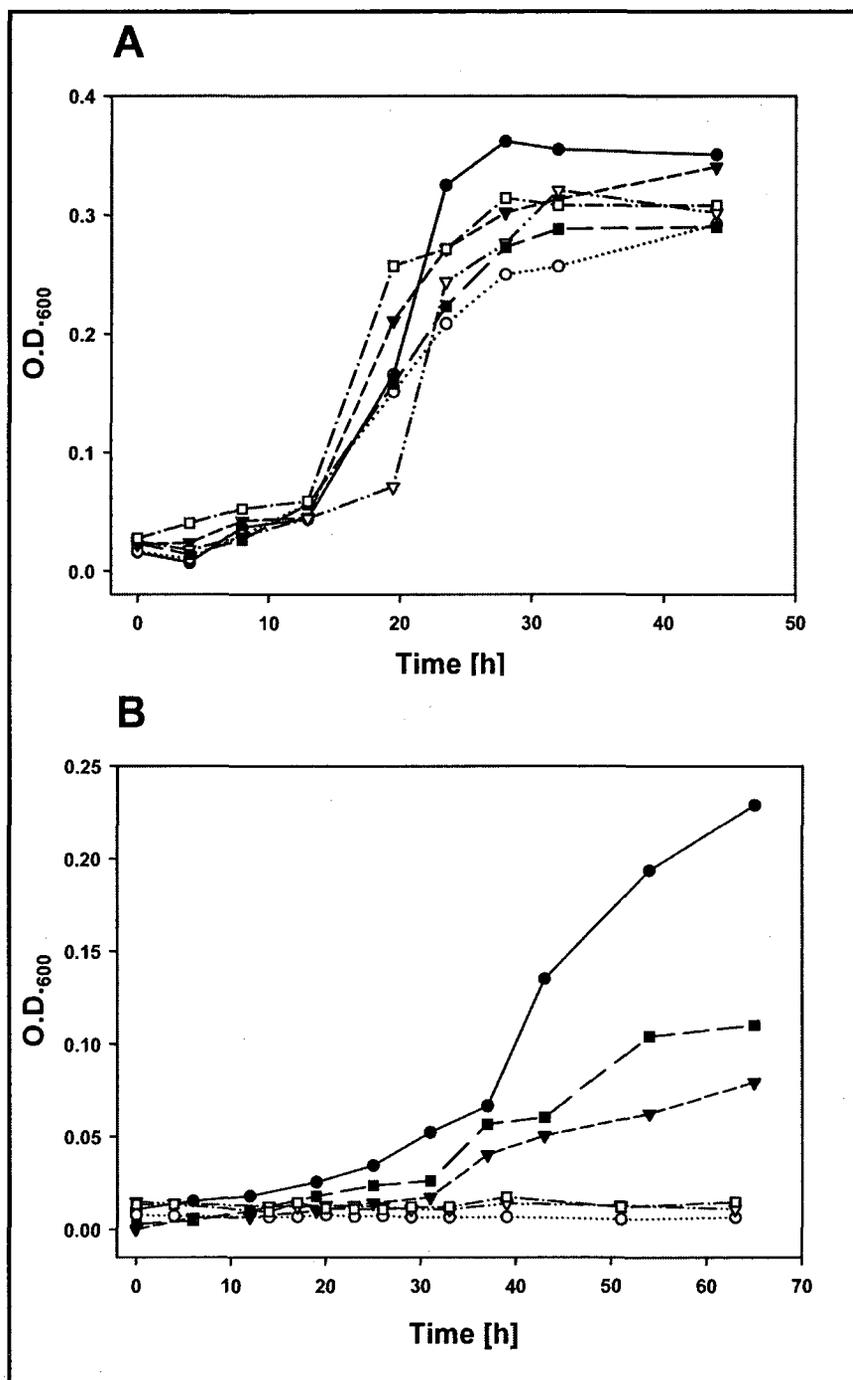
**Figure 20. Confirmation of *H. neapolitanus* Form IA RubisCO mutants via PCR amplification of genomic DNA**

Primer pairs annealing upstream and downstream [A] and within [B] the *cbbL-cbbS* region of the *cso* operon of *H. neapolitanus* were chosen to confirm the genotypes of the RubisCO replacement and deletion mutants. DNA molecular weight standards in kbp are indicated on the left of each gel. M = marker, 1 = *cbbL::Tc NC cbbL*, 2 = *cbbS::Tc NC cbbS*, 3 = *cbbLS::Tc NC cbbLS*, 4 = *cbbL::Tc C cbbL*, 5 = *cbbL::kan<sup>r</sup>*.

### Phenotypic characterization of the Form IA RubisCO deletion mutant

Inactivation of genes encoding carboxysomal proteins have been reported to increase the CO<sub>2</sub> requirement for growth by mutant *H. neapolitanus* cells [15, 46]. For instance, Baker *et al.* [15] reported that an *H. neapolitanus* mutant, wherein the *cbbL* gene was inactivated by the insertion of a kanamycin cassette, required elevated levels of CO<sub>2</sub> for survival. Similarly, disruption of the *csoS1A* gene that encodes a major carboxysomal shell component resulted in the generation of a high CO<sub>2</sub>-requiring (*hcr*) *H. neapolitanus* mutant [46]. In light of these findings, it became important to follow the growth patterns of the Form IA RubisCO mutants generated in this study in ambient and enriched CO<sub>2</sub> environments.

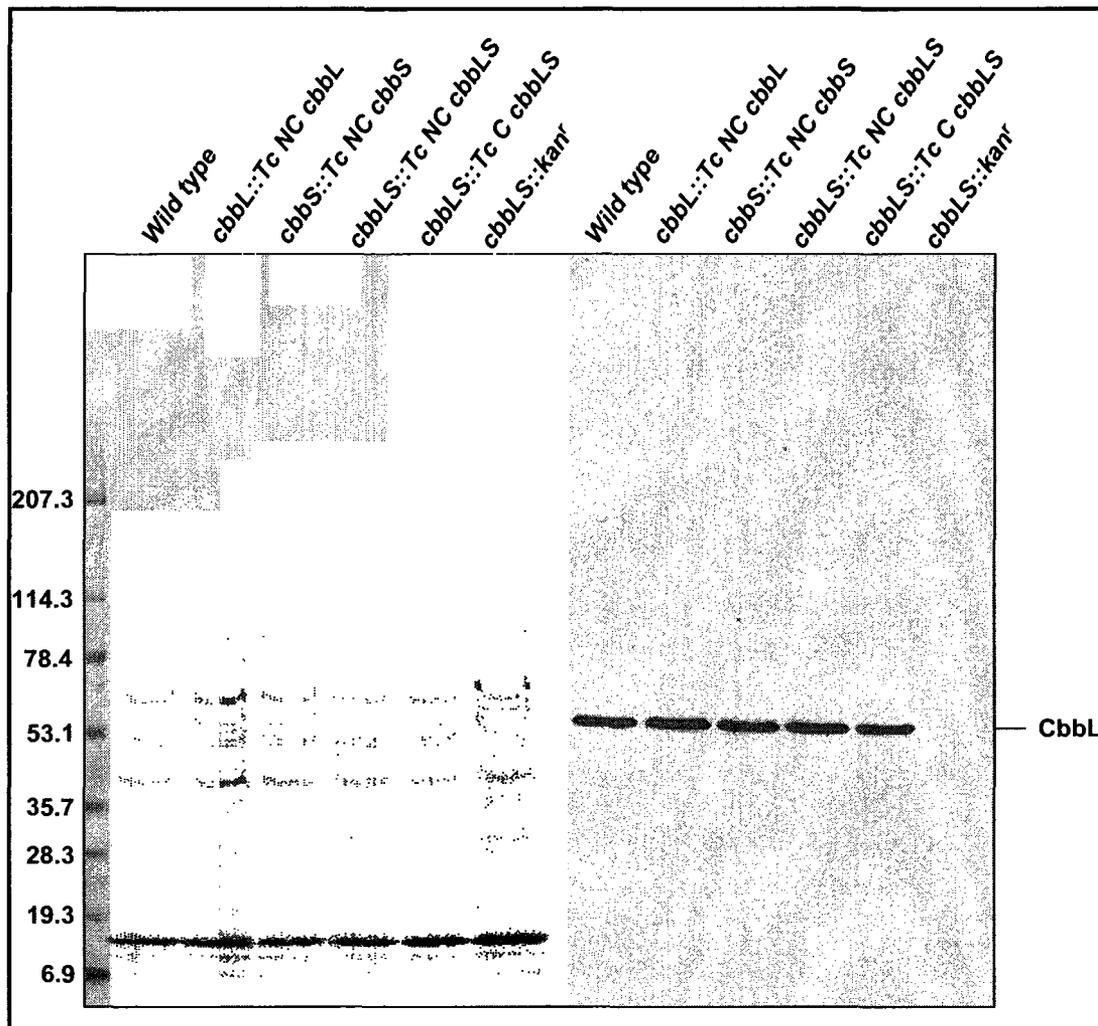
Under elevated CO<sub>2</sub> conditions, the *cbbLS::kan'* mutant was able to grow at a rate and to a maximum density similar to those of wild type (**Figure 21A**). However, under ambient CO<sub>2</sub> conditions, the mutant failed to grow even beyond a period of more than 60 hours (**Figure 21B**). The failure of the mutant to grow in air could be attributed to the lack of Form IA RubisCO and hence inability to efficiently utilize the intracellular inorganic carbon (C<sub>i</sub>) pool. To determine if the mutant synthesized functional Form II enzyme under high CO<sub>2</sub> conditions, RubisCO activity assays and immunoblotting analysis using anti CbbM (Form II RubisCO) were performed after centrifugation of a clarified extract of the mutant on a 0.2-0.8 M sucrose gradient. The RubisCO activity profile of the deletion mutant was found to be similar to that of the *cbbL::Km* mutant reported by Baker *et al.* [15]. A peak of activity with matching immunoblot signals corresponding to Form II RubisCO was observed in fractions 19 and 20 (**Figure 22**).



**Figure 21. Growth patterns of wild type and Form IA RubisCO mutants**  
 Growth of wild type (●) and Form IA RubisCO mutants *cbbL::Tc NC cbbL* (○), *cbbS::Tc NC cbbS* (▼), *cbbLS::Tc NC cbbLS* (▽), *cbbL::Tc C cbbLS* (■), *cbbL::kan<sup>r</sup>* (□) in [A] ambient CO<sub>2</sub>, and [B] air supplemented with 5% CO<sub>2</sub> was monitored by measuring the optical density of batch cultures at 600 nm. All mutants grow analogous to wild type in air supplemented with 5% CO<sub>2</sub>. However, only a few mutants are able to grow in air.



That the *cbbLS::kan<sup>r</sup>* mutant did not express Form IA enzyme was confirmed by subjecting crude cell extract of the mutant that was grown under elevated CO<sub>2</sub> conditions to immunoblotting analysis using Form IA CbbL-specific antibodies. As expected, no CbbL protein was detected in the mutant's extract (**Figure 23**).



**Figure 23. Expression of the large subunit (CbbL) in wild type and Form IA RubisCO mutants**

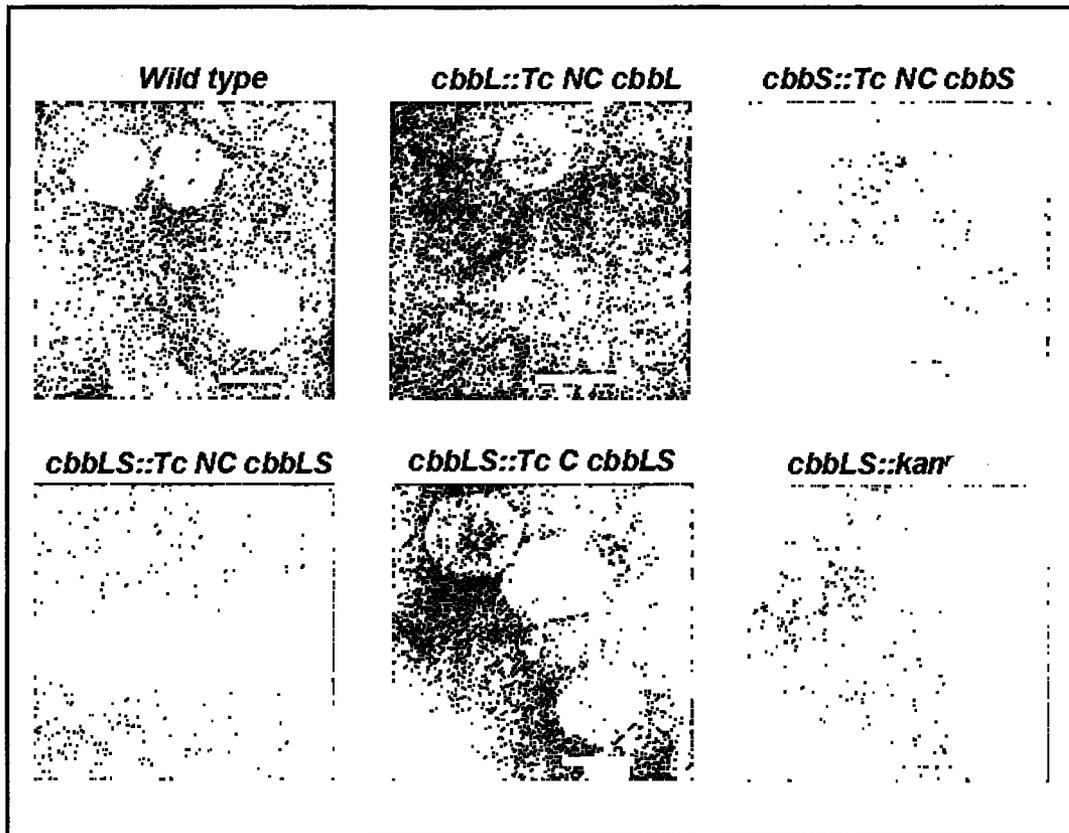
Crude cell extracts (10  $\mu$ g) of wild type and RubisCO mutants were resolved by SDS-PAGE and stained with GelCode Blue stain (left). A blot of an identical gel was probed with anti-CbbL antibodies that specifically recognize the large subunit of Form IA RubisCO. All mutants except *cbbLS::kanr* express the large subunit at levels similar to that of wild type. Protein molecular weight standards in kDa are shown on the left of the gel.

### Microcompartment assembly in the Form IA RubisCO deletion mutant

Because the *cbbL-cbbS* deletion in the *cbbLS::kan<sup>r</sup>* mutant left downstream genes under the control of the endogenous *cso* operon promoter or, more likely, the *kan<sup>r</sup>* promoter, it was expected that the carboxysome shell proteins would be expressed. Baker *et al.* [15] reported that a *cbbL::Km* mutant lacking the large subunit of Form IA RubisCO contained empty polyhedral inclusions that resembled wild type carboxysome shells but were smaller in size. That the polyhedral structures observed were those of carboxysome shells was not proven. To assess the ability of the *cbbLS::kan<sup>r</sup>* mutant to assemble carboxysomes, the well-established cell fractionation protocol was employed that has been routinely used for the purification of carboxysomes from wild type *H. neapolitanus* cells [36, 42, 53]. In the *cbbLS::kan<sup>r</sup>* mutant, the final 48,000 x g differential centrifugation step yielded the typical pellet that in wild type cells is significantly enriched for carboxysomes. Further, purification of this fraction by sucrose density centrifugation resulted in the formation of a milky white band that occupied a position in the gradient similar to those observed for wild type carboxysomes. Electron microscopy studies revealed that the *cbbLS::kan<sup>r</sup>* mutant assembled polyhedral microcompartments possessing the regular size and shape of wild type carboxysomes (**Figure 24**). To substantiate these observations, the purified microcompartments were subjected to SDS-PAGE and immunoblotting analyses.

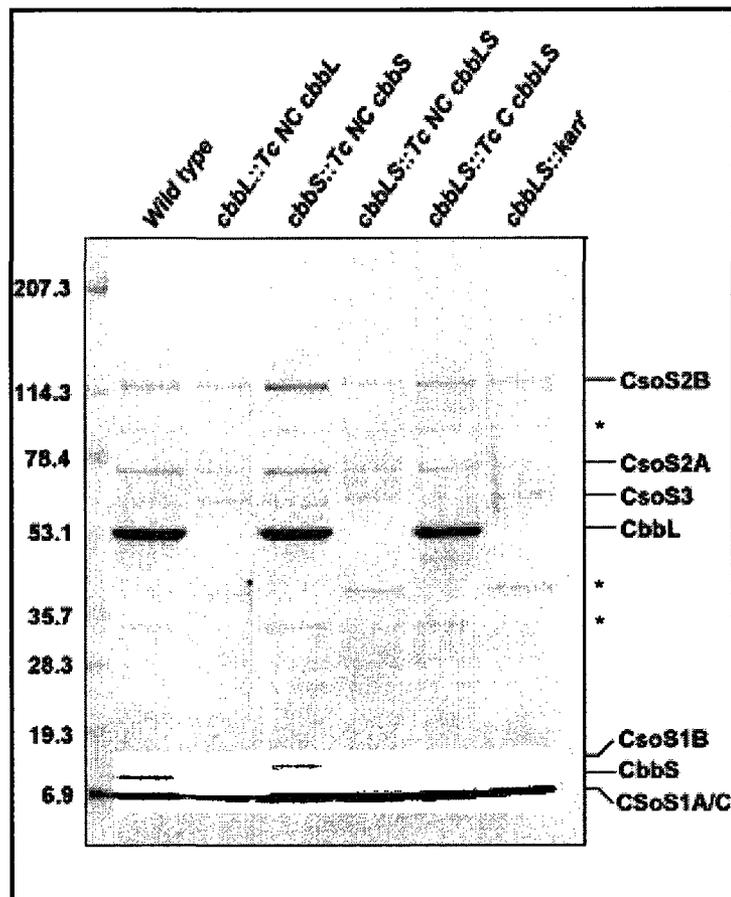
When resolved by SDS-PAGE, microcompartments purified from the *cbbLS::kan<sup>r</sup>* mutant appeared to possess the typical assortment of carboxysomal shell components (**Figure 25**). However, clearly lacking were the Form IA RubisCO large

and small subunits, CbbL and CbbS, respectively. This observation was consistent with the electron microscopy findings that the *cbbLS::kan<sup>r</sup>* mutant's microcompartments were devoid of any Form IA RubisCO molecules. To further substantiate these findings, the mutant polyhedra were subjected to immunoblotting using anti-CsoS1 and anti-CbbL antibodies, which recognize the CsoS1 shell proteins and the large subunit (CbbL) of Form IA RubisCO, respectively. It was seen that the anti-CbbL antibody failed to recognize any CbbL protein in the mutant polyhedra (**Figure 26A**). On the other hand, the anti-CsoS1 immunoblot identified the CsoS1C, -1A, and -1B shell proteins in the mutant polyhedra at levels similar to those of wild type carboxysomes (**Figure 26B**). These results confirmed that the microcompartments observed in the *cbbLS::kan<sup>r</sup>* mutant represented empty carboxysome shells.



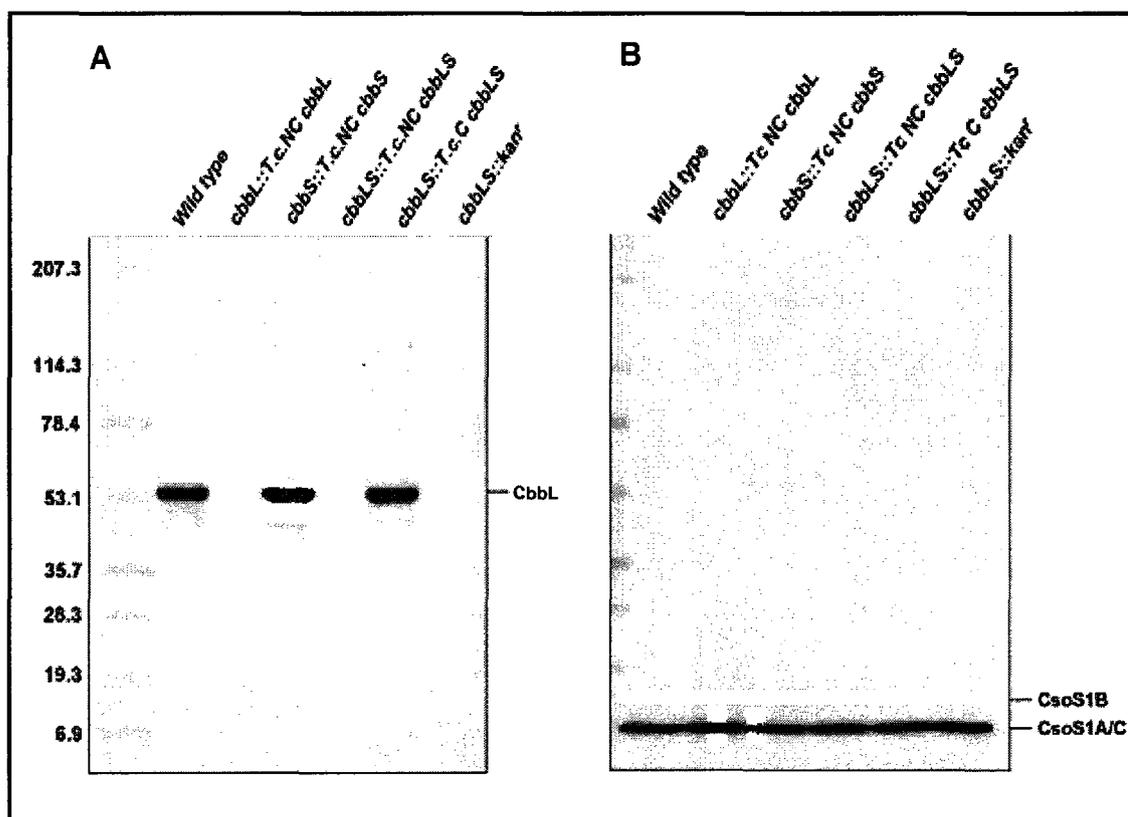
**Figure 24. Electron micrographs of wild type and mutant carboxysomes**

Purified carboxysomes were stained with 1% w/v ammonium molybdate and visualized by electron microscopy at a magnification of 50,000 X. White bars represent a scale of 100 nm. Carboxysomes purified from the *cbbS::Tc NC cbbS* mutant appear to lack the characteristic donut-shaped RubisCO molecules in comparison to those observed in carboxysomes purified from *wild type* and the *cbbLS::Tc C cbbLS* mutant. The *cbbL::Tc NC cbbL*, *cbbLS::Tc NC cbbLS*, and *cbbLS::kan<sup>r</sup>* mutants assemble empty carboxysome shells.



**Figure 25. Polypeptide composition of wild type and mutant carboxysomes**

Wild type and mutant carboxysomes were separated on a denaturing 4-20% SDS-polyacrylamide gradient gel and stained with GelCode Blue. Amount of total protein in mutant carboxysomes was normalized to that of wild type such that all lanes contained equal number of carboxysomes. Protein molecular weight standards in kDa are represented on the left of the gel. Bands labeled with an asterisk might represent SDS-resistant oligomeric forms of shell proteins, as yet unidentified carboxysomal components, or membrane protein contaminants. All mutant carboxysomes are composed of the typical set of shell proteins in near wild type stoichiometric ratios. Carboxysomes of the *cbbL::Tc NC cbbL*, *cbbLS::Tc NC cbbLS*, and *cbbLS::kan<sup>r</sup>* mutants lack the large and small subunit of Form IA RubisCO.



**Figure 26. Western blot analysis of wild type and mutant carboxysomes probed with anti-CbbL and anti-CsoS1 antibodies**

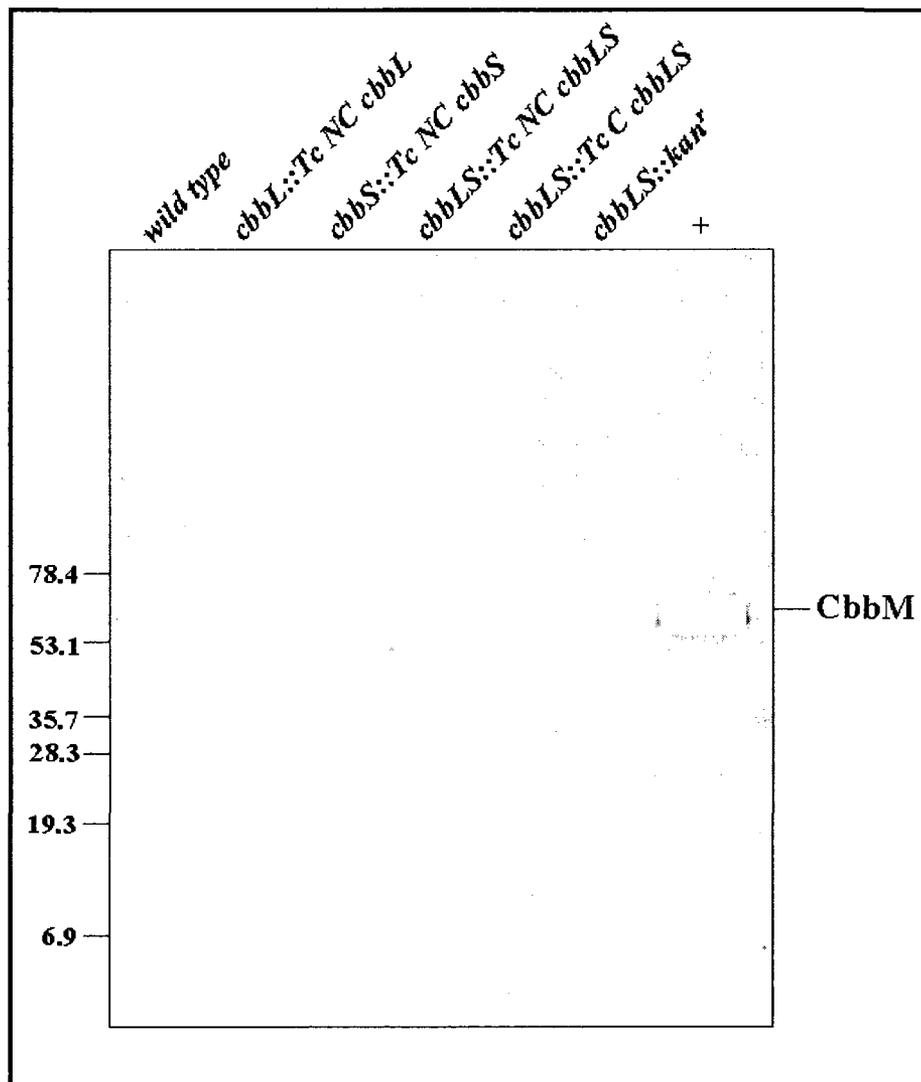
Blots of an identical gel to that shown in **Figure 25** were probed with [A] anti-CbbL (Form IA specific) antibodies and [B] anti-CsoS1 antibodies. Protein mass of carboxysomes lacking RubisCO was normalized to that of wild type carboxysomes such that each lane contained an equal number of carboxysomes. Protein molecular weight standards in kDa are represented on the left of blot shown in [A]. No CbbL protein is detected in carboxysomes of the *cbbL::Tc NC cbbL*, *cbbLS::Tc NC cbbLS*, and *cbbLS::kan<sup>r</sup>* mutants. All mutant carboxysomes appear to contain near wild type levels of the CsoS1 shell proteins.

## Form II RubisCO is not associated with carboxysomes

Previous studies in *Hydrogenovibrio marinus* and *Halothiobacillus neapolitanus* have shown that the expression of Forms I and II of RubisCO can be correlated with CO<sub>2</sub> availability [15, 16]. Under high CO<sub>2</sub> conditions, Form II RubisCO is the predominant species while under low CO<sub>2</sub> conditions Form IA RubisCO predominates. In addition, thin sections of wild type *Hy. marinus* cells grown under high CO<sub>2</sub> conditions were found to lack carboxysomes [16]. In the present study it was observed that the Form IA RubisCO deletion mutant, when grown under high CO<sub>2</sub> conditions, was able to assemble empty carboxysome shells. Although results from electron microscopy and CO<sub>2</sub> fixation assays of the empty carboxysomes suggested that they did not contain a Form IA RubisCO, the possibility remained that in this mutant Form II RubisCO could be physically tethered to the carboxysome shell without being sequestered into the core. If this were the case, SDS-PAGE analysis and electron microscopy may lack the resolution needed to detect low levels of Form II enzyme. To analyze wild type and mutant carboxysomes for the presence of Form II RubisCO, an anti-CbbM western blot was performed. Results revealed that, at least within detection limits, wild type and mutant carboxysomes had no Form II RubisCO associated with them (**Figure 27**).

The finding that empty carboxysome shells assemble in the absence of both forms of RubisCO also raises an important question in chemoautotrophic bacteria possessing duplicate copies of the *cbbL* and *cbbS* genes. The genomes of these bacteria have one Form IA RubisCO-encoding gene set associated with their carboxysome (*csa*) operon, while the other is located elsewhere. It remains unknown

whether microcompartments of these bacteria sequester the *csO* or the non-*csO* associated Form IA RubisCO species.



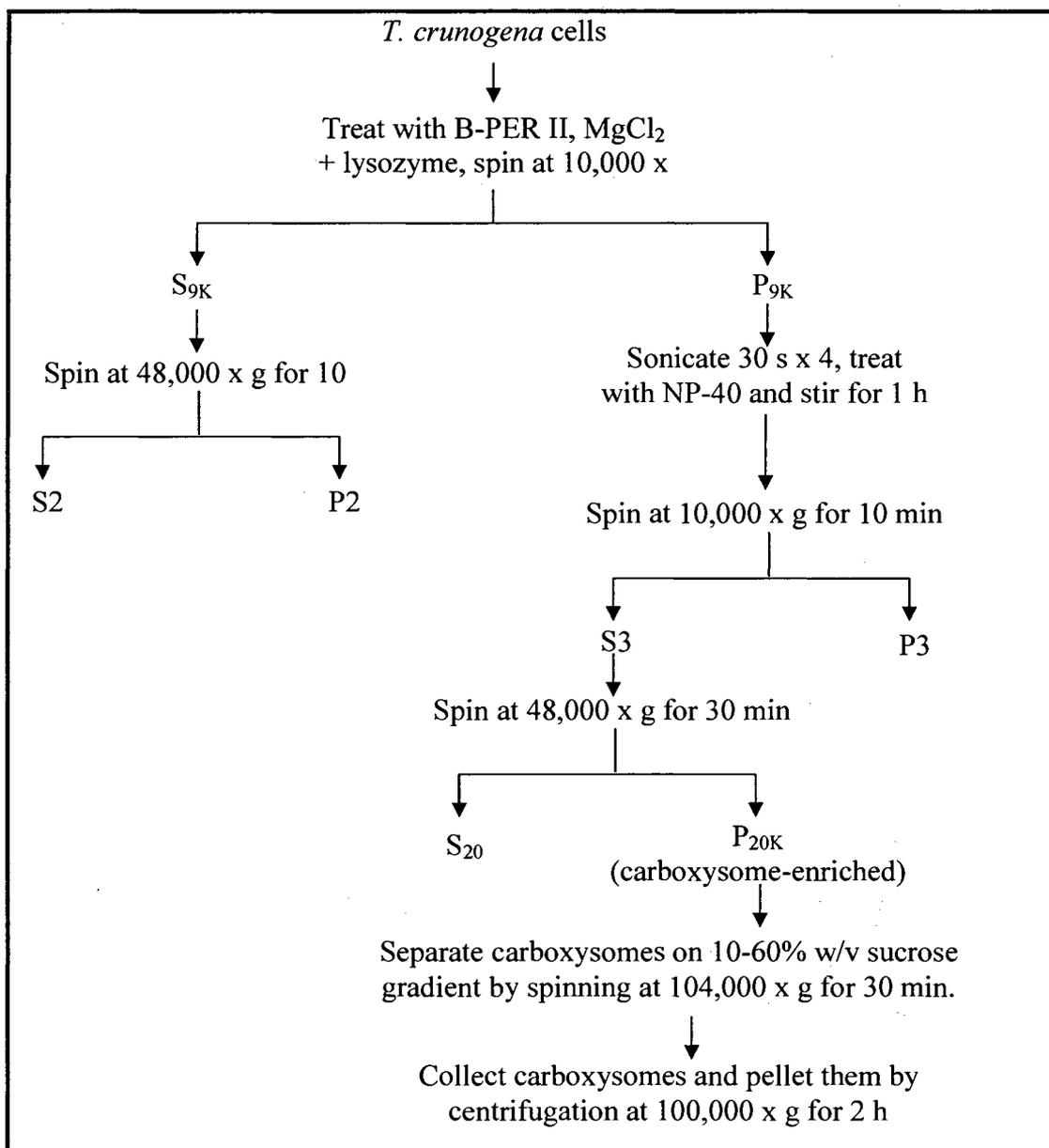
**Figure 27. Western blot analysis of wild type and mutant carboxysomes probed with anti-CbbM (Form II RubisCO) antibodies**

A blot of an identical gel to that shown in **Figure 25** was probed with anti-CbbM antibodies. All lanes except the one labeled '+' contained equal number of carboxysomes. The lane labeled '+' is a positive control for the immunoblot containing 25  $\mu$ g crude cell extract of wild type cells grown in air supplemented with 5% CO<sub>2</sub>. A doublet band is observed in the positive control lane, which may represent the carbamylated and unmodified forms of the Form II RubisCO (CbbM) that is generated by over-heating of samples prior to loading on a denaturing gel, or a degradation product of the enzyme.

### Characterization of carboxysomes from *Thiomicrospira crunogena*

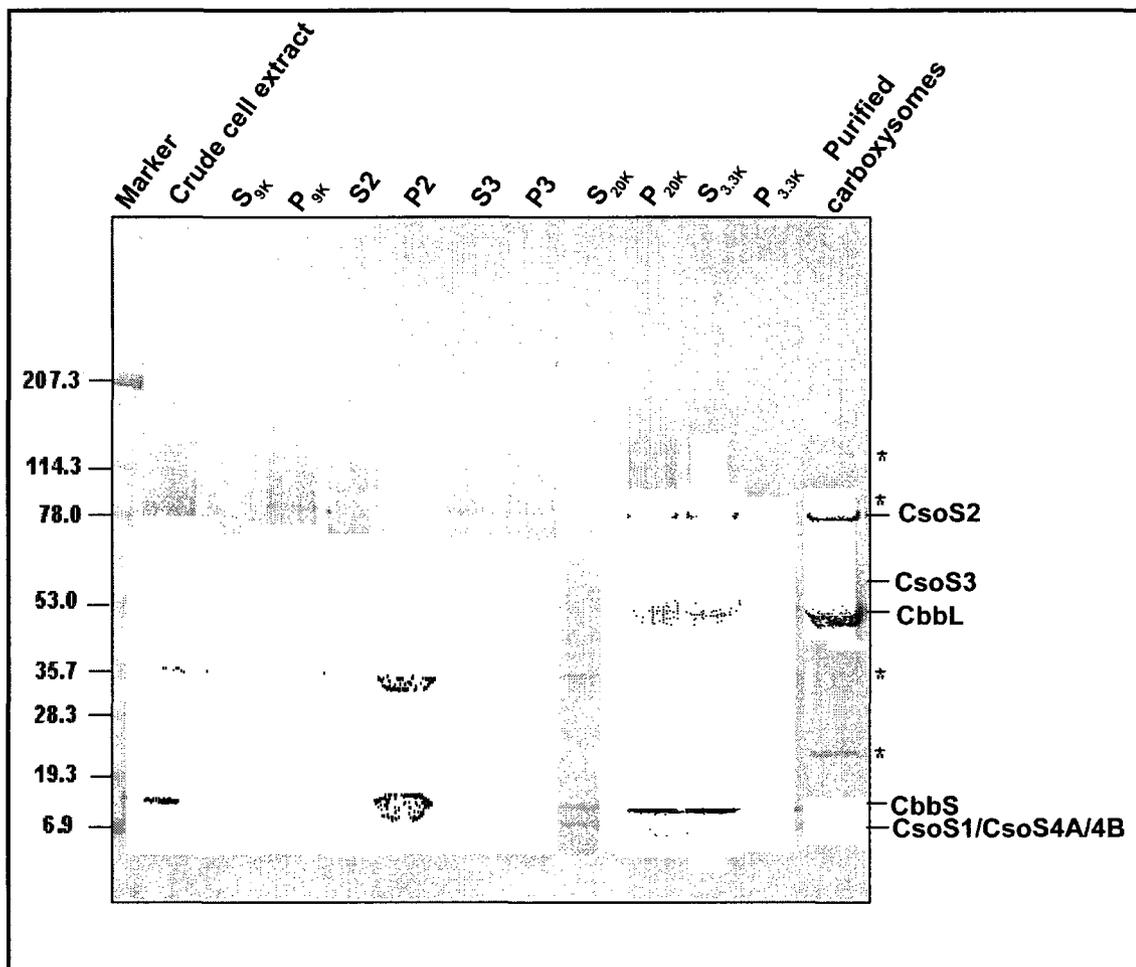
Carboxysomes from the hydrothermal vent  $\gamma$ -proteobacterium, *T. crunogena*, were purified and characterized to determine whether the non *csa* or *csb* operon-encoded copy of the Form IA RubisCO was sequestered within its microcompartments.

Purification of carboxysomes from the model sulfur-oxidizing bacterium, *H. neapolitanus* includes detergent treatment of cells, sonication, and a series of differential centrifugation steps that have been well documented [36, 42]. The use of B-PER II detergent treatment and sonication ensures maximal cell lysis and results in the release of cytosolic contents along with carboxysomes, some of which might be membrane-bound [50]. When the method of So *et al.* [36] was employed initially to isolate carboxysomes from *T. crunogena*, no enrichment of carboxysomes was achieved. To address this issue, the P<sub>9K</sub> fraction, which typically contains most of the disrupted membrane fragments and vesicles that result from sonication and detergent treatment of cells, was sonicated for four 30 s bursts, with 1 min cooling intervals between each cycle. In addition, the sonicated P<sub>9K</sub> fraction was treated with Nonidet P-40 (NP-40) for 1 h to enhance solubilization of the membrane debris (**Figure 28**). Although the harsher sonication conditions employed could possibly result in the disruption of a few intact carboxysomes, it was felt that there was a better chance of liberating the majority of the trapped microcompartments. Using the protocol depicted in **Figure 28**, a near-homogenous preparation of carboxysomes was obtained as judged by SDS-PAGE analysis (**Figure 29**).



**Figure 28. Flowchart for purification of carboxysomes from *T. crunogena***

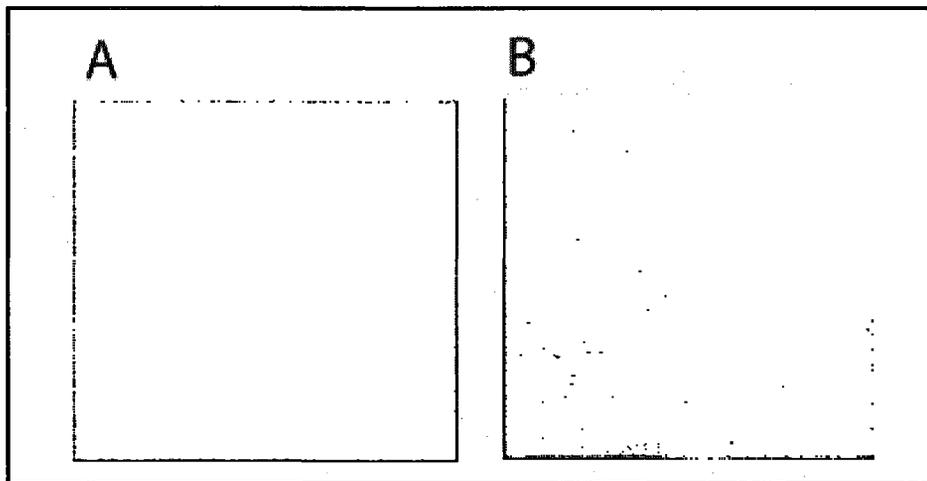
The P<sub>9K</sub> fraction was used as starting material for the purification of *T. crunogena* carboxysomes. Upon subjecting the carboxysome-enriched P<sub>20K</sub> pellet to sucrose density gradient centrifugation, a thick milky white band was observed towards the middle of the gradient. The position of this band corresponded to that occupied by wild type *H. neapolitanus* carboxysomes on a sucrose gradient. Subsequent electron microscopy, SDS-PAGE and western blotting analyses revealed that a near-homogeneous preparation of carboxysomes was obtained from *T. crunogena* cells.



**Figure 29. SDS-PAGE analysis of fractions obtained during the purification of *T. crunogena* carboxysomes**

Fractions depicted in **Figure 28** were separated on a denaturing 4-20% SDS-polyacrylamide gradient gel and stained with GelCode Blue stain. Protein molecular weight standards in kDa are represented on the left of the gel. The lane on the extreme right represents 10  $\mu$ g of purified *T. crunogena* carboxysomes. Lanes labeled S<sub>3.3K</sub> and P<sub>3.3K</sub> represent clarified supernatant and pellet fractions of the P<sub>20K</sub> pellet, obtained after spinning at 1,000  $\times$  g for 4 min. Bands labeled with an asterisk may represent SDS-resistant oligomeric forms of carboxysomal shells, as yet unidentified carboxysomal components or membrane protein contaminants.

Visual examination by electron microscopy revealed that carboxysomes purified from *T. crunogena* cells had an identical size and shape to those purified from *H. neapolitanus* wild type cells (**Figure 30**). Also seen was the typical 3-4 nm thick outer shell encapsulating many molecules of the donut-shaped RubisCO enzyme.



**Figure 30. Transmission electron micrographs of negatively stained carboxysomes from *H. neapolitanus* and *T. crunogena***

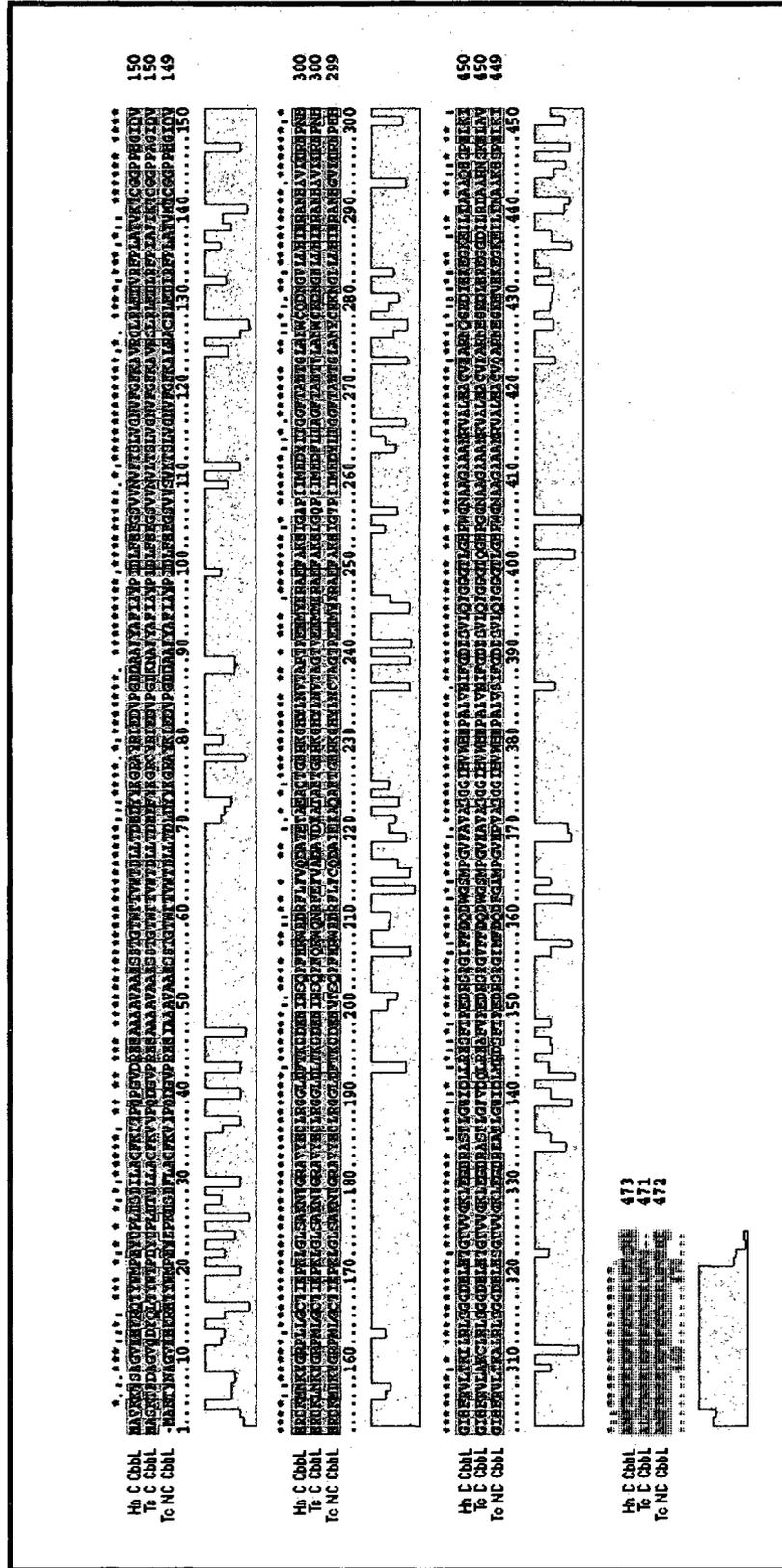
Carboxysomes were purified from *H. neapolitanus* [A] using the method of So *et al.* [36] and from *T. crunogena* [B] using the protocol outlined in **Figure 28**. Samples were stained with 1% w/v ammonium molybdate and observed under a transmission electron microscope at 30,000X magnification. The white scale bars represent 100 nm.

### **Identification of Form IA RubisCO species associated with carboxysomes of *T. crunogena***

Many CO<sub>2</sub>-fixing autotrophic bacteria possess multiple copies of the RubisCO-encoding genes [16, 18, 91, 92]. Genomes of autotrophs such as *Ralstonia eutropha* and *Chromatium vinosum* harbor a duplicate pair of the Form IA RubisCO-encoding genes [91, 92] while the genomes of others, such as *H. neapolitanus*, *Thiomonas intermedia*, and *Rhodobacter sphaeroides*, possess genes encoding one Form IA and a Form II RubisCO [93]. Intriguingly, three sets of RubisCO-encoding genes have been observed in *Acidithiobacillus ferrooxidans* and *Hydrogenovibrio marinus* [16, 18]. Two of them encode Form IA enzymes while the third encodes a Form II RubisCO.

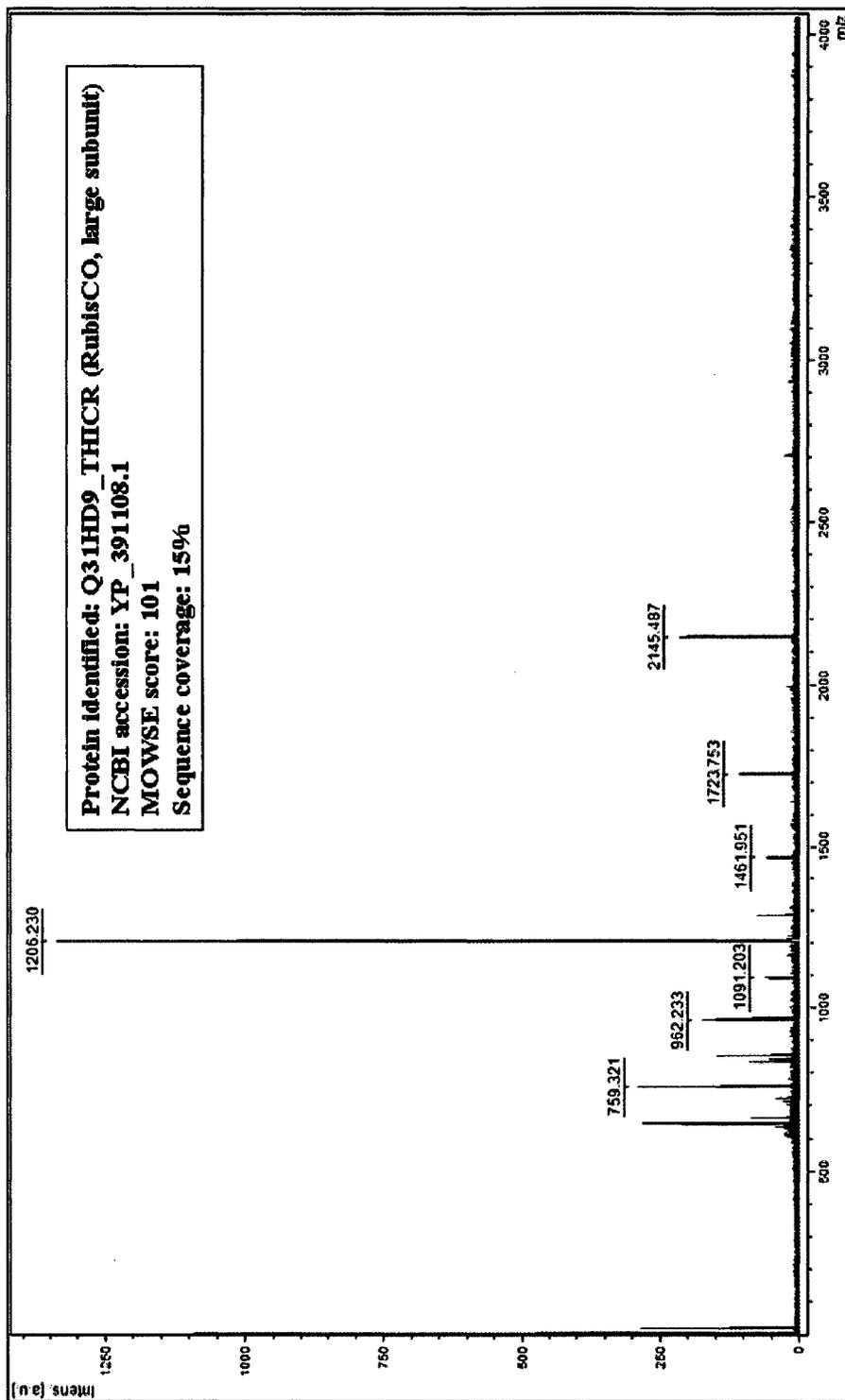
The recent sequencing of the *T. crunogena* genome also revealed the presence of three RubisCO-encoding gene sets, the organization of which was found to be similar to those of *A. ferrooxidans* and *Hy. marinus* [17]. The large subunits (CbbL) of Form IA RubisCO found in *T. crunogena* and other autotrophic bacteria exhibit close to 80% identity (**Figure 31**). Their small subunits (CbbS), however, show considerable sequence divergence. The presence of a six amino acid insertion at the N-terminus of the non *csa* operon encoded CbbS sequences has been used to subclassify the RubisCO enzymes as Form IAq (non *csa* or noncarboxysomal) and Form IAc (*csa* or carboxysomal) (**Figure 32**) [56]. It has been speculated that the N-terminal motif found in the small subunits of noncarboxysomal RubisCO may interfere with sequestration of the Form IAq species within carboxysomes [56].

To identify the species of Form IA RubisCO associated with carboxysomes of *T. crunogena*, the bands corresponding to the large and small subunit of the carboxysomal that were recovered from an SDS-polyacrylamide gel were subjected to mass spectrometric analysis. Clearly, the *csa* operon-encoded large and small subunit of Form IA RubisCO with molecular weight search (MOWSE) scores of 101 and 77, respectively, were identified as the carboxysomal species (**Figures 33 and 34**). MOWSE scores of 65 and up were considered as significant hits. There was no indication of the non *csa* operon-encoded Form I enzyme or Form II RubisCO being associated with carboxysomes of *T. crunogena* based on the mass spectrometric analysis.



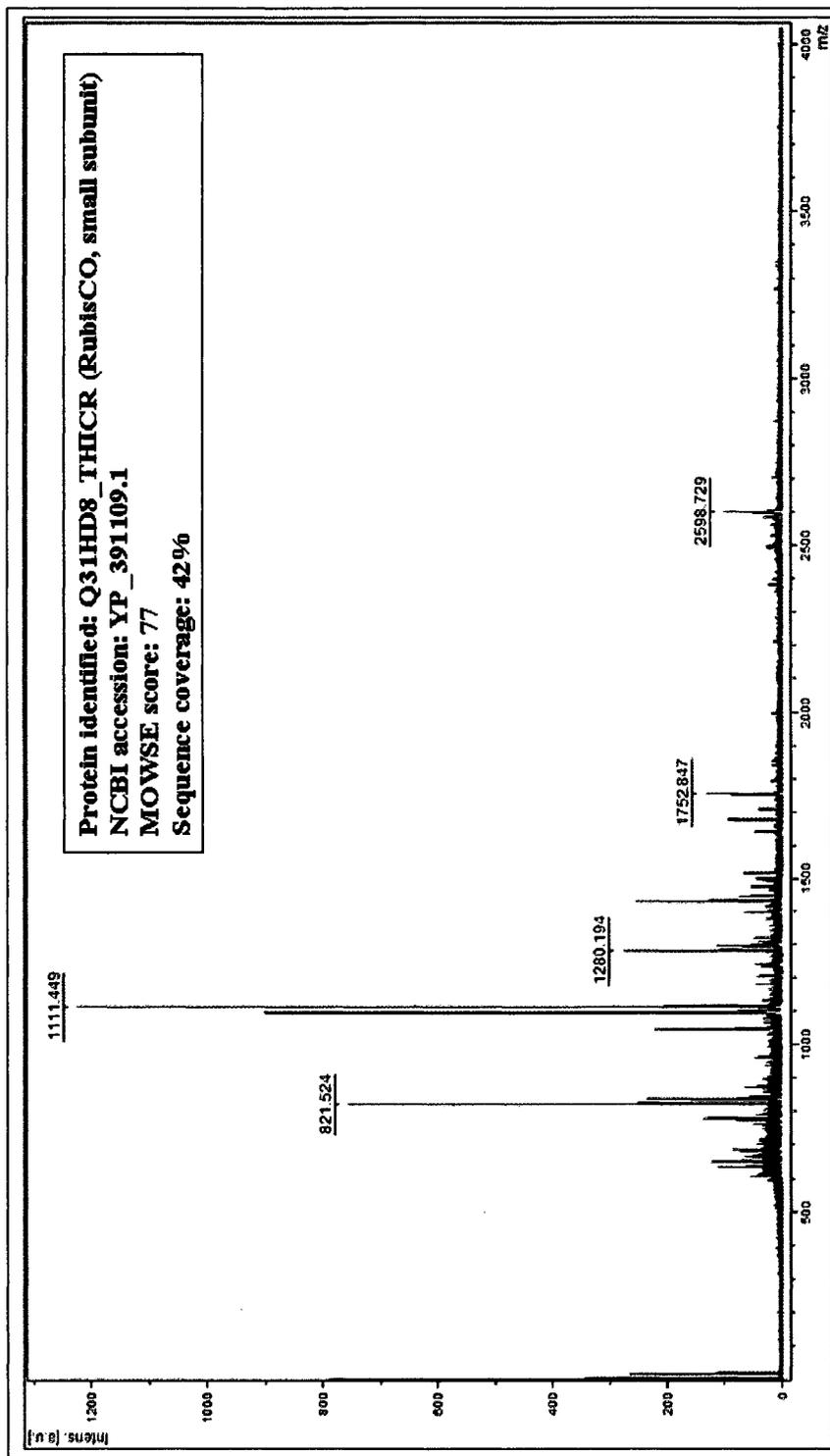
**Figure 31. Primary sequence alignment of Form IA RubisCO large subunits**  
 Amino acid sequences of carboxysomal (C) and noncarboxysomal (NC) large subunits (CbbL) of RubisCO from *H. neapolitanus* (Hn) and *T. crunogena* (Tc) were aligned using ClustalX 1.83. These CbbL sequences exhibit more than 80% identity.





**Figure 33. Identification of the *csa* operon-encoded RubisCO large subunit in carboxysomes of *T. crunogena***

M/z ratios shown in the above spectrum were entered into the MASCOT Peptide Mass Fingerprint search engine, part of Bruker Daltonics' FlexAnalysis software, to identify proteins from primary sequence databases. The *T. crunogena* carboxysomal large subunit, CbbL, was identified with a MOWSE score of 101 and a sequence coverage of 29%. MOWSE scores 65 and above were considered to be significant hits.



**Figure 34. Identification of the *csO* operon-encoded RubisCO small subunit in carboxysomes of *T. crunogena***  
 M/z ratios were searched against protein database as described in Figure 33. The *T. crunogena* carboxysomal small subunit, CbbS, was identified with a MOWSE score of 77 and a sequence coverage of 46%. MOWSE scores 65 and above were considered to be significant hits.

### ***Thiomicrospira crunogena* CsoS2 is expressed as a single polypeptide**

SDS-PAGE and western blotting analyses of *T. crunogena* carboxysomes revealed some interesting findings. Although the overall polypeptide composition and relative abundance of *T. crunogena* carboxysomal components were found to be similar to those of *H. neapolitanus* (**Figure 35, Table 1**), a few variations in the migration patterns of the shell proteins could be detected. The CsoS1 shell proteins appeared to co-migrate as a single band of 7 kDa, unlike the CsoS1B and CsoS1A/C proteins of *H. neapolitanus*, which are known to migrate as two distinct bands of apparent molecular weight (MW) 15 and 6.5 kDa, respectively (**Figure 35**). In addition to SDS-PAGE, this difference in migration patterns was confirmed by western blotting analysis, using antibodies that recognize all three CsoS1 proteins of *H. neapolitanus* (**Figures 35 and 36**). These antibodies recognized a heavily represented band of 7 kDa on the blot containing carboxysomes of *T. crunogena*. A faint band of 25 kDa was also observed on the blot, which may represent oligomeric forms of the CsoS1 proteins. Alternatively, the faint band may represent a BMC domain-containing protein encoded by the gene Tcr\_0851 that cross reacts with the anti-CsoS1 antibody (**Figure 36**). The *csoS1B*, *csoS1A*, and *csoS1C* genes of *H. neapolitanus* encode proteins that are highly homologous to one another [9, 94]. CsoS1B differs slightly from CsoS1C and CsoS1A in having a twelve amino acid extension at its C-terminus. In contrast, sequence alignments in *T. crunogena* revealed that the longest CsoS1 protein has a six amino acid extension at its C-terminus, which does not possess any homology to the C-terminal twelve amino acid extension found in CsoS1B. This extension apparently caused no difference in the

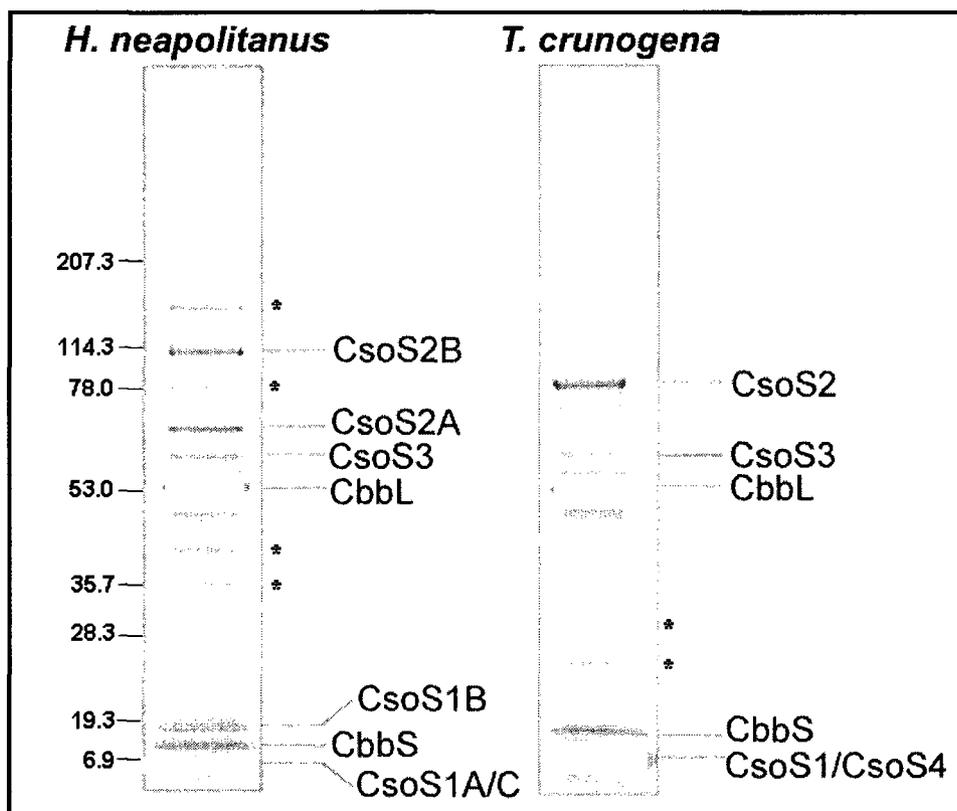
protein's migration pattern based on observations from SDS-PAGE and anti-CsoS1 western blotting analyses. The contribution of such CsoS1 C-terminal extensions in carboxysome assembly and architecture remains unknown.

The most striking difference observed in the polypeptide composition of *T. crunogena* carboxysomes, as judged by SDS-PAGE analysis, was the absence of the characteristic 130 kDa CsoS2B and 85 kDa CsoS2A polypeptides found in carboxysomes of *H. neapolitanus* [52]. Instead, a single band that migrated at 80 kDa was observed. Western blotting analysis using antibodies raised against *H. neapolitanus* CsoS2B polypeptide was used to identify the 80 kDa band as CsoS2. Surprisingly, the CsoS2 proteins of these two organisms are only ~35% identical. A previous report suggested that the *H. neapolitanus* *csoS2* gene products might be post-translationally modified via O-linked glycosylation [52]. It was reasoned that differential glycosylation results in the aberrant migration patterns of the CsoS2B and CsoS2A polypeptides on denaturing polyacrylamide gels [52]. Although *T. crunogena* CsoS2 migrated as a single band of 80 kDa, which was slightly higher than its predicted MW of 68 kDa, it remains undetermined if the *T. crunogena* CsoS2 protein undergoes glycosylation or some other kind of post-translational modification as speculated in its *H. neapolitanus* counterpart.

From SDS-PAGE analysis, a band migrating at about 60 kDa was thought to be carbonic anhydrase (CsoS3 or CsoSCA) (**Figure 35**). The observed MW weight of this protein correlated well with the predicted MW of the *csoS3* gene product, which was calculated to be 59.2 kDa. However, whether this band represents CsoS3 could not be ascertained, since no immune reactive bands within that size range were

detected using antibodies raised against the *H. neapolitanus* CsoS3 protein. To confirm the identity of this band, mass spectrometry will have to be performed.

The earlier finding that carboxysomes of *T. crunogena* sequester the *cso* operon-encoded copy of Form IA RubisCO had raised an important question. Could the noncarboxysomal copy of Form IA RubisCO be incorporated within carboxysomes of *T. crunogena* in the absence of the carboxysomal copy? To answer this question, a series of *H. neapolitanus* Form IA RubisCO replacement mutants was constructed wherein the endogenous *cbbL* and/or *cbbS* genes were replaced with orthologs from *T. crunogena*.



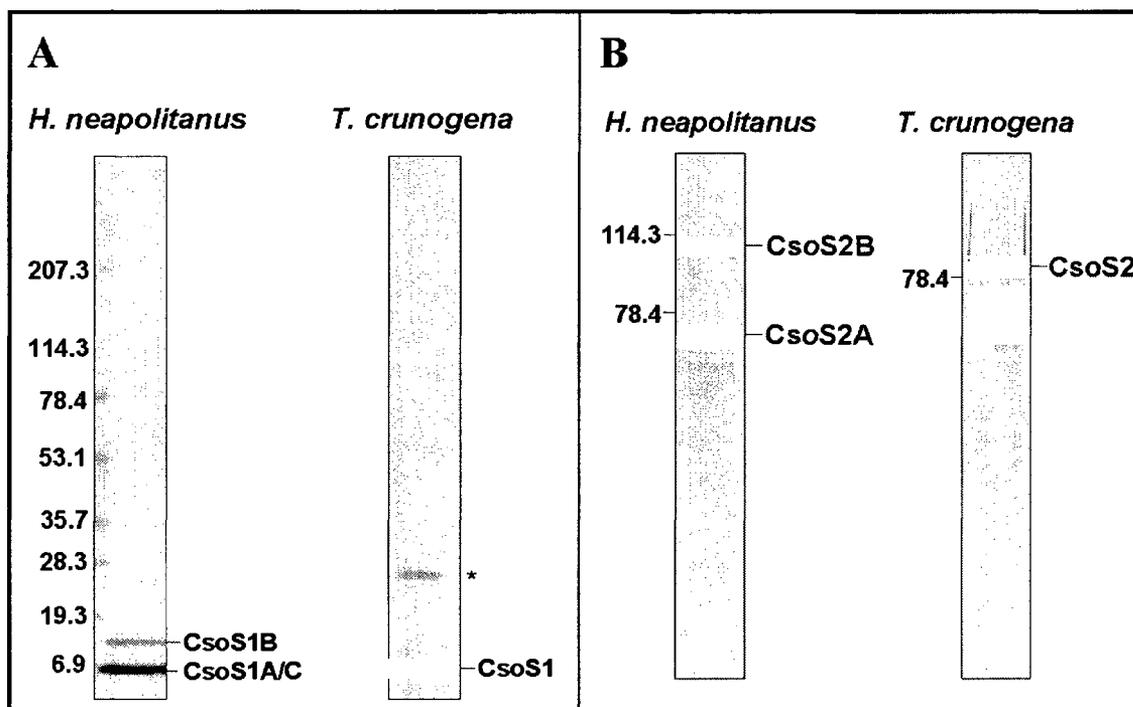
**Figure 35. Separation of *H. neapolitanus* and *T. crunogena* carboxysome proteins via SDS-PAGE**

Purified intact carboxysomes (20  $\mu$ g) from *H. neapolitanus* and *T. crunogena* were resolved on a denaturing 4-20% SDS-polyacrylamide gradient gel and stained with GelCode Blue. Protein molecular weight standards in kDa are indicated on the left. Bands labeled with an asterisk may represent SDS-resistant oligomeric forms of the shell proteins, as yet unidentified carboxysomal components, or membrane protein contaminants that co-fractionated with carboxysomes.

Carboxysomal Components	Abundance in <i>H. neapolitanus</i> carboxysomes	Abundance in <i>T. crunogena</i> carboxysomes
CsoS2B	6.2%	16.5%
CsoS2A	5.8%	
CsoS3	4.3%	5.3%
CbbL	50%	45%
CbbS	12%	10%
CsoS1B	8.7%	23.2%
CsoS1A/C	13%	
CsoS4A	-	
CsoS4B	-	

**Table 1. Stoichiometric distribution of carboxysomal components in *H. neapolitanus* and *T. crunogena***

Carboxysomes purified from *H. neapolitanus* and *T. crunogena* were separated on a denaturing 4-20% SDS-polyacrylamide gradient gel and stained with GelCode Blue (Figure 35). To estimate the relative abundance of carboxysomal components, densitometry was performed on the separated polypeptides using BioRad's Quantity One v 4.6.3 software. The fraction of CsoS4A and 4B proteins in carboxysomes of *H. neapolitanus* could not be estimated as they represent very low abundance proteins. Likewise, the CsoS1 proteins of *T. crunogena* could not be resolved on a one-dimensional gel as they appeared to co-migrate as one band of apparent MW 7 kDa.



**Figure 36. Western blotting analyses of *H. neapolitanus* and *T. crunogena* carboxysomes**

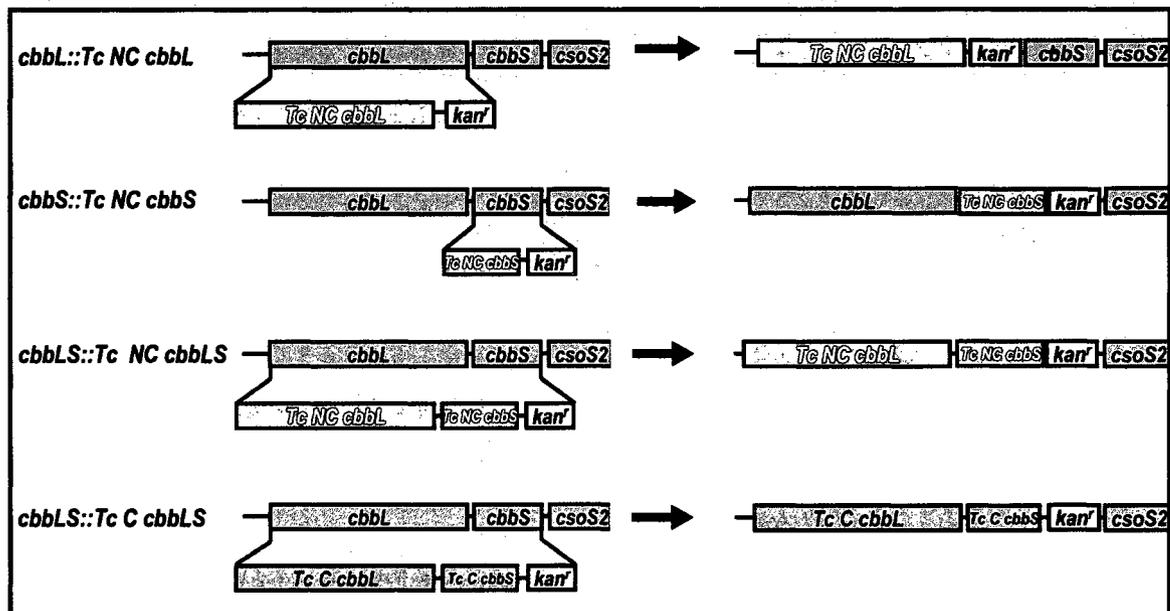
Purified carboxysomes (5  $\mu$ g) of *H. neapolitanus* and *T. crunogena* were separated by SDS-PAGE, electroblotted, and probed using antibodies raised against the [A] CsoS1 and [B] CsoS2 proteins of *H. neapolitanus*. The anti-CsoS1 and -CsoS2 antibodies recognized single bands of apparent MW 7 kDa and 80 kDa respectively. Protein molecular weight standards in kDa are indicated on the left of both blots. The band labeled with an asterisk in [A] may represent an SDS-resistant oligomeric form of the CsoS1 shell proteins or a BMC domain-containing protein encoded by the gene, Tcr\_0851 that cross reacts with the anti-CsoS1 antibody.

### III. Sequestration of foreign Form IA RubisCO within carboxysomes of *Halothiobacillus neapolitanus*

*H. neapolitanus* was chosen for the construction and phenotypic characterization of Form IA RubisCO replacement mutants because of several reasons. Previous studies have shown *H. neapolitanus* to possess a robust homologous recombination system for constructing carboxysome mutants [15, 46]. Similar studies indicating the amenability of *T. crunogena* to genetic manipulations have not been reported. Furthermore, the genome of *H. neapolitanus* does not harbor the noncarboxysomal Form IA RubisCO-encoding genes [18], thereby eliminating their possible interference in constructing the replacement or deletion mutants and interpreting their phenotypes. Finally, the protocol for carboxysome isolation and purification in *H. neapolitanus* is well established [36, 42]. Hence, the ability of the Form IA RubisCO mutants to assemble and incorporate foreign Form IA RubisCO can be evaluated.

Form IA RubisCO replacement mutants (**Figure 37**) were constructed using a strategy similar to that used in generating the *cbbLS::kan<sup>r</sup>* mutant. Genomic integration of the replacement fragments within the *cso* operon of *H. neapolitanus* was verified by polymerase chain reactions. Genomic DNA extracted from wild type and replacement mutants served as template for the reactions. The same set of primers used to confirm the *cbbLS::kan<sup>r</sup>* mutant was used to verify the replacement mutants (**Figure 20A, B**). When the primer pair designed to anneal 300 bp upstream and downstream of the *cbbL* and *cbbS* regions, respectively, were used, an approximately 2500 bp band was observed in the wild type positive control reaction, which

represents the endogenous *cbbLS* genes. Reactions from Form IA RubisCO replacement mutants showed the presence of approximately 3500 bp bands (**Figure 20A**). The difference in size is due to the presence of the 1000 bp *kan<sup>r</sup>* cassette flanking the *cbbL* or *cbbS* gene in the replacement mutants. When the primer pair designed to anneal within the coding regions of the endogenous *cbbL-cbbS* genes was used, only the wild type positive control reaction showed an approximately 1750 bp band that corresponded to endogenous *cbbL-cbbS*. None of the reactions involving genomic DNA of the replacement mutants showed a product (**Figure 20B**). In addition to PCR amplification, integration of the replacement fragments at the endogenous *cbbLS* loci within the *cs0* operon was confirmed by genomic DNA sequencing (University of Maine DNA Sequencing Facility).



**Figure 37. Organization of genes within the *cso* operon of the Form IA RubisCO replacement mutants**

Mutants were constructed by replacing the *H. neapolitanus* genes for large (*cbbL*) and/or small (*cbbS*) subunit (green boxes) with noncarboxysomal (*Tc NC*; red boxes) or carboxysomal (*Tc C*; orange boxes) genes from *T. crunogena*. All mutants carry a kanamycin (*kan'*) cassette for selection purposes.

### **Growth phenotypes of the Form IA RubisCO replacement mutants**

Perturbations of the Form IA RubisCO-encoding genes in *H. neapolitanus* and *Synechococcus* PCC 6803 have been reported to result in the generation of mutants displaying a high CO<sub>2</sub>-requiring (*hcr*) phenotype [15, 95, 96]. Considering these findings, it became important to monitor the growth patterns of the Form IA RubisCO replacement mutants in ambient and enriched CO<sub>2</sub> environments.

At an elevated CO<sub>2</sub> level, the replacement mutants were able to grow at rates and to maximum densities similar to those of wild type (**Figure 21A**). However, under ambient CO<sub>2</sub> conditions, the *cbbS::Tc NC cbbS* and *cbbLS::Tc C cbbLS* mutants grew considerably more slowly than wild type cells. On the other hand, the *cbbL::Tc NC cbbL* and *cbbLS::Tc NC cbbLS* mutants failed to grow even beyond a period of more than 60 hours (**Figure 21B**).

### **Expression analysis of the large subunit (CbbL) in crude cell extracts of Form IA RubisCO mutants**

The reason why the *cbbL::Tc NC cbbL* and *cbbLS::Tc NC cbbLS* mutants could not grow at ambient CO<sub>2</sub> levels was unclear, especially since in these mutants the replaced *cbbL* and/or *cbbS* genes were placed under the transcriptional control of the endogenous *csa* promoter or of the *kan<sup>r</sup>* promoter, which should have resulted in the expression of the foreign CbbL and CbbS proteins at appreciable levels.

To examine the expression of Form IA RubisCO in these mutants, crude extracts of wild type and mutant cells, grown under elevated CO<sub>2</sub> conditions, were subjected to SDS-PAGE analysis and immunoblotting with Form IA CbbL-specific

antibodies. Results revealed that the large subunit in all mutants was being expressed at levels comparable to those of wild type cells (**Figure 23**). Thus, the inability of the *cbbL::Tc NC cbbL* and *cbbLS::Tc NC cbbLS* mutants to grow in ambient CO<sub>2</sub> did not appear to be the result of reduced CbbL expression.

To determine if the *cbbL::Tc NC cbbL* and *cbbLS::Tc NC cbbLS* mutants were expressing enzymatically active holoenzyme, RubisCO activity assays and immunoblotting analyses using anti-CbbL and anti CbbM (Form II RubisCO) antibodies were performed after centrifugation of clarified extracts on 0.2-0.8 M sucrose gradients at 72,000 x g for 24 h at 4°C [15, 97]. Under these conditions, the activity profiles of the L<sub>8</sub>S<sub>8</sub> Form IA RubisCO holoenzyme and of the smaller Form II RubisCO can be clearly distinguished based on the enzymes' different sedimentation rates through the gradients. It was observed that the *cbbLS::Tc NC cbbLS* mutant, in which the endogenous Form IA RubisCO-encoding genes were replaced with the noncarboxysomal orthologs from *T. crunogena*, did contain heterologous Form IA RubisCO holoenzyme, as indicated by a peak of activity and matching immunoblot signals centering around fractions 24-26. Form II RubisCO activity was found to be concentrated in fractions 19 and 20 (**Figure 38**). In the *cbbL::Tc NC cbbL* mutant, a peak of similar magnitude was observed for Form IA RubisCO, but it was less pronounced because of masking by the higher Form II RubisCO activity (**Figure 39**). Collectively, these results indicate that enzymatically active Form I RubisCO holoenzymes are assembled in the *cbbL::Tc NC cbbL* and *cbbLS::Tc NC cbbLS* mutant; however these mutants fail to grow in ambient CO<sub>2</sub>.





## **Isolation and characterization of carboxysomes from Form IA RubisCO mutants**

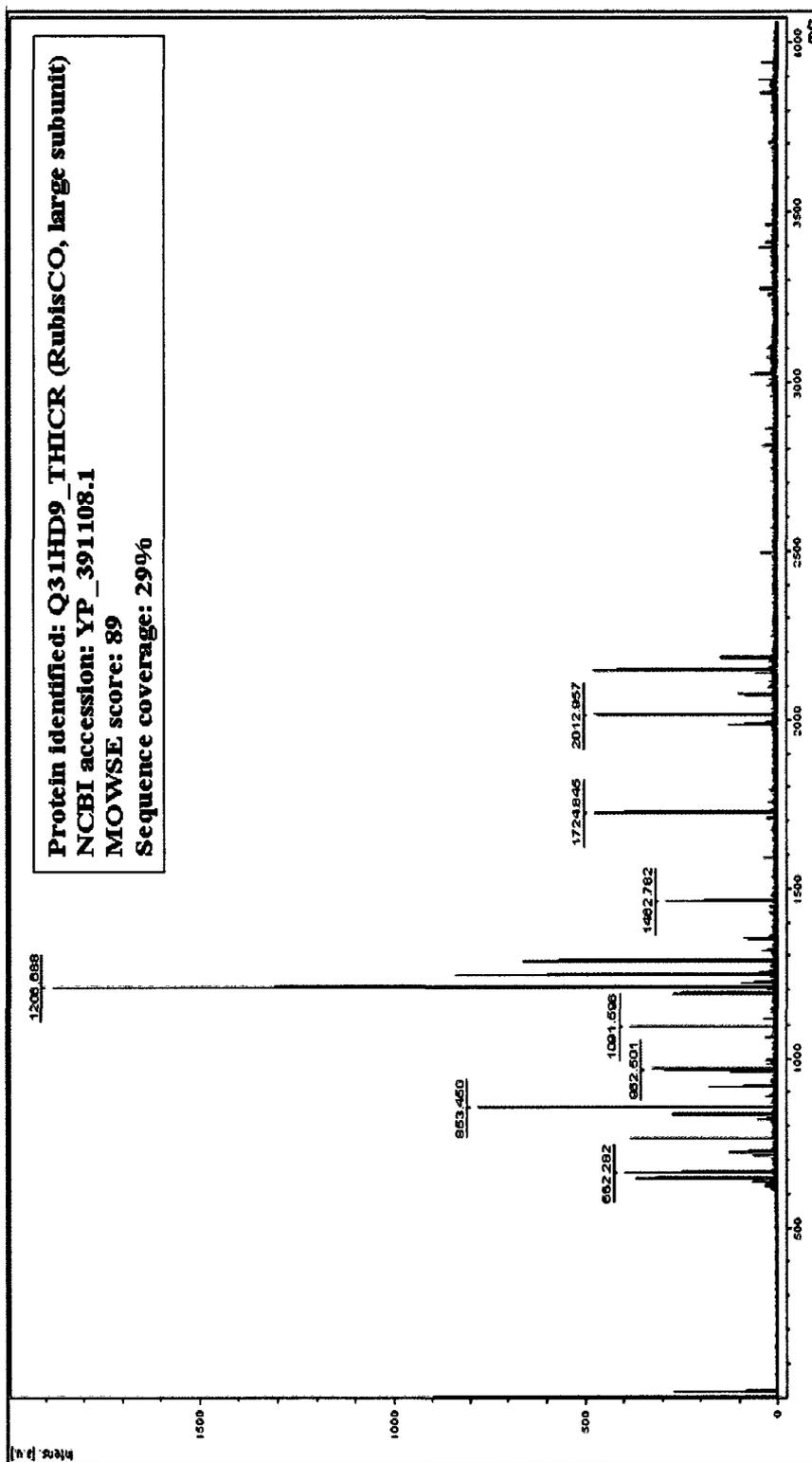
All Form IA RubisCO replacement mutants expressed the large subunit at wild type levels. Since the *cbbL* and/or *cbbS* substitutions in all of the mutants left downstream genes unperturbed and under the control of the endogenous *csa* operon promoter or, more likely, the *kan<sup>r</sup>* promoter, it was expected that the carboxysome shell proteins would be expressed. The ability of the replacement mutants to assemble carboxysome shells and sequester the heterologous RubisCO enzymes within their microcompartments needed to be evaluated, especially in light of the findings by Pierce *et al.* [95] and Amichay *et al.* [96], who reported that replacement of the Form IA RubisCO-encoding genes in *Synechocystis* 6803 with a Form II RubisCO-encoding gene from *Rhodospirillum rubrum* resulted in the generation of mutants lacking microscopically observable carboxysomes.

The well-established cell fractionation protocol used routinely for purifying carboxysomes from wild type *H. neapolitanus* cells [10, 36, 42] was used to assess the ability of replacement mutants to assemble microcompartments. In all replacement mutants, the final step of the fractionation protocol yielded the typical pellets that in wild type cells correspond to pure carboxysomes. Electron microscopy studies revealed that all replacement mutants assembled polyhedral microcompartments that possessed the regular size and shape of wild type carboxysomes (**Figure 24**). However, one striking difference observed in the polyhedra isolated from the *cbbL::Tc NC cbbL* and *cbbLS::Tc NC cbbLS* mutants was the absence of RubisCO molecules sequestered within them. Rather, they appeared as

intact yet empty carboxysome shells. Surprisingly, carboxysomes purified from the *cbbS::Tc NC cbbS* and *cbbLS::Tc C cbbLS* mutants were filled with RubisCO molecules (**Figure 24**). This finding suggests that the *cbbS::Tc NC cbbS* mutant assembles carboxysomes containing chimeric RubisCO molecules that are comprised of *H. neapolitanus* large subunit and *T. crunogena* noncarboxysomal small subunit and that the *cbbLS::Tc C cbbLS* mutant contains heterologous RubisCO molecules comprised of *T. crunogena* carboxysomal large and small subunits. To substantiate these observations, microcompartments purified from the replacement mutants were subjected to SDS-PAGE and immunoblotting analyses.

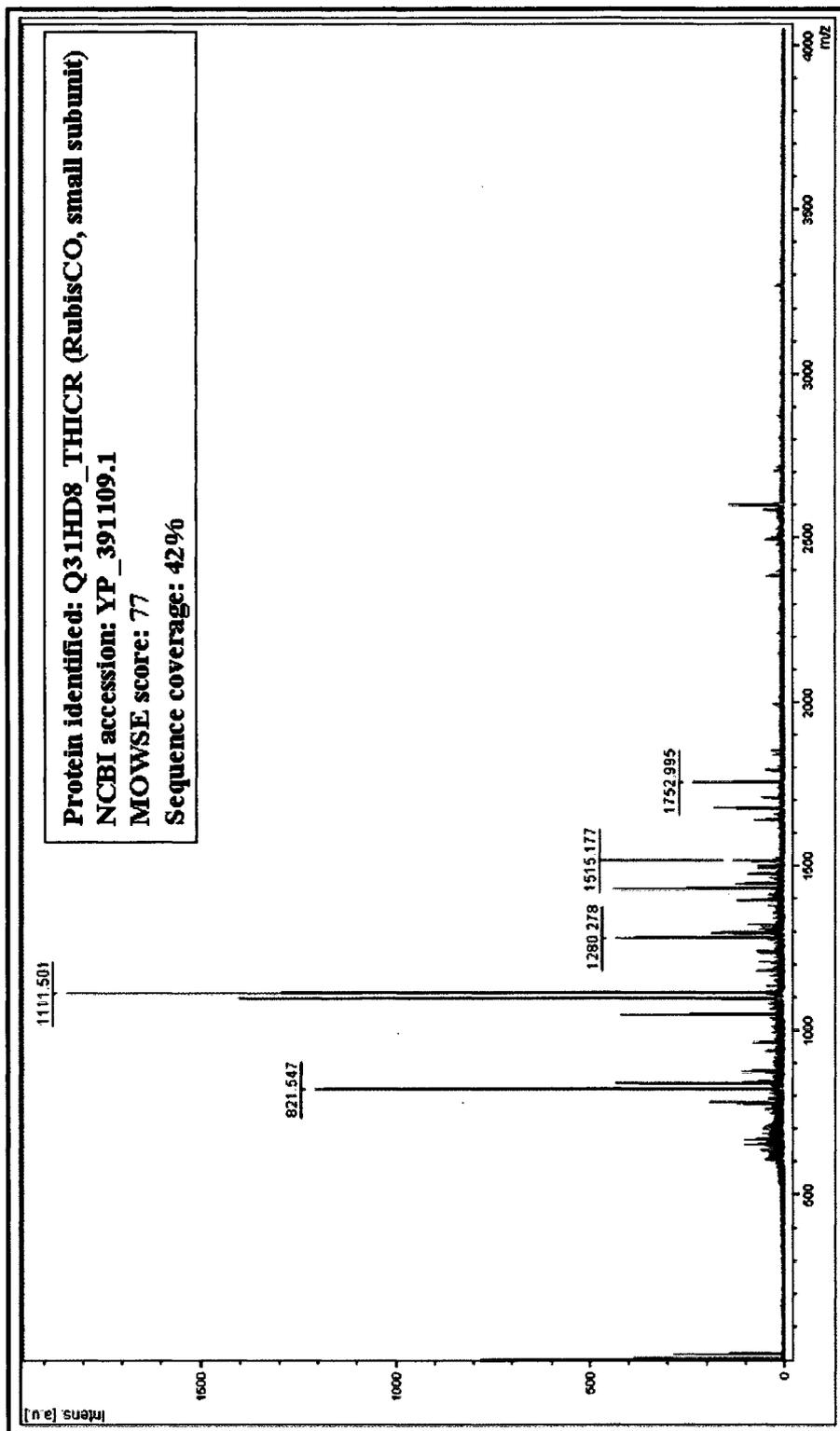
When resolved by SDS-PAGE, purified carboxysomes of the *cbbS::Tc NC cbbS* and *cbbLS::Tc C cbbLS* mutants were found to possess the typical stoichiometric distribution and polypeptide composition of wild type carboxysomes (**Figure 25**). An important observation made was the difference in migration pattern of the *cbbS::Tc NC cbbS* mutant's small subunit on a denaturing 4-20% SDS-polyacrylamide gradient gel. The CbbS band was located above *H. neapolitanus* and *T. crunogena* carboxysomal counterparts. The slower migration of CbbS was attributed to its higher molecular weight in comparison to the other small subunits from both bacteria. The presence of noncarboxysomal small subunit in carboxysomes of the *cbbS::Tc NC cbbS* mutant and of heterologous carboxysomal RubisCO in carboxysomes of the *cbbLS::Tc C cbbLS* mutant was confirmed by mass spectrometric analysis, which identified the *T. crunogena* non-*cso*-associated CbbS and *cso*-associated CbbL and CbbS proteins with MOWSE scores of 89, 77, and 65,

respectively (**Figures 40-42**). MOWSE scores of 65 and higher were considered as significant hits.



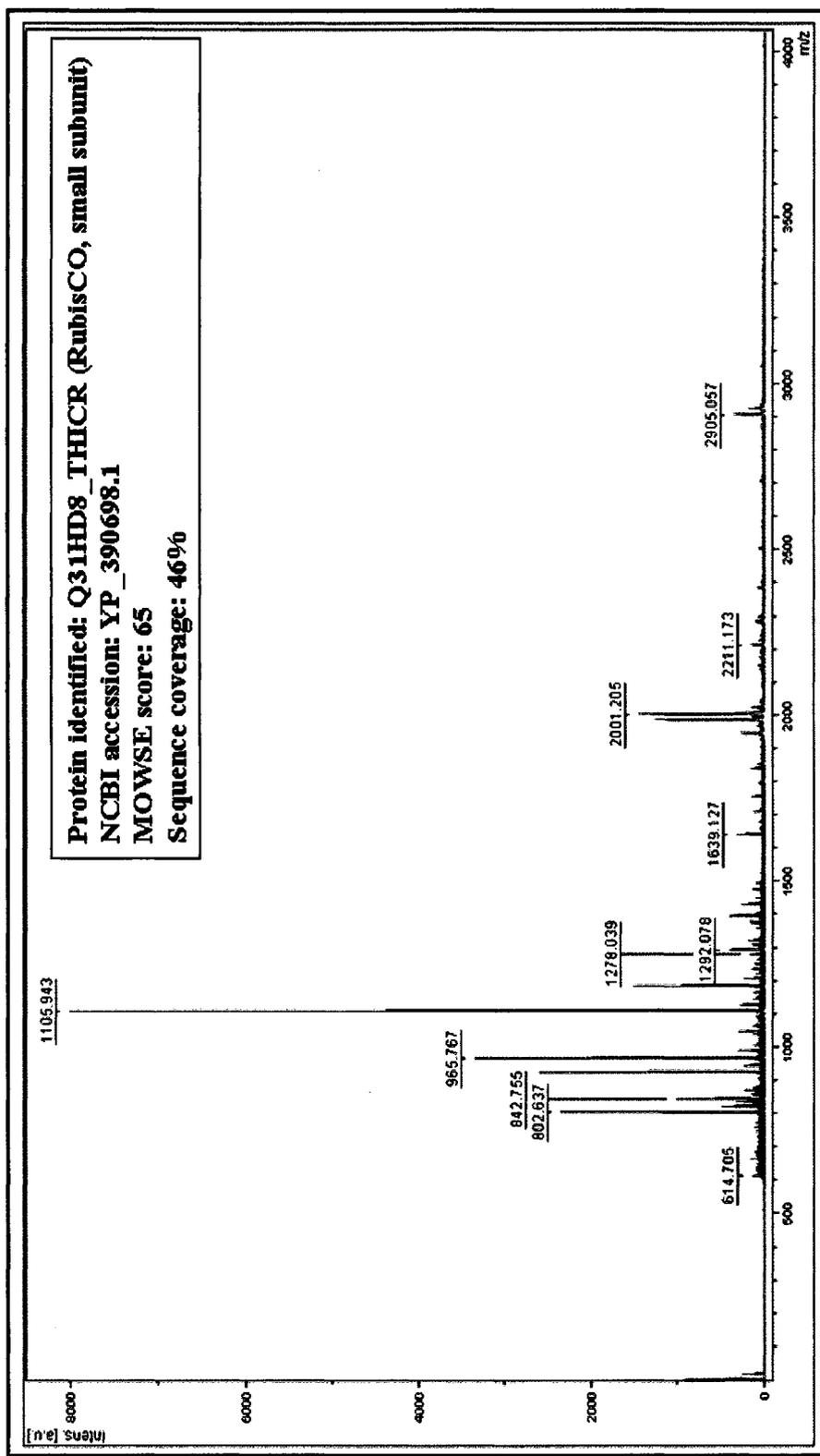
**Figure 40. Identification of *T. crunogena* carboxysomal large subunit (CbbL) in carboxysomes of the *cbbLS::Tc CbbLS* mutant via mass spectrometry**

M/z ratios shown in the above spectrum were entered into the MASCOT Peptide Mass Fingerprint search engine, part of Bruker Daltonics' FlexAnalysis software, to identify proteins from primary sequence databases. The *T. crunogena* carboxysomal large subunit, CbbL, was identified with a molecular weight search (MOWSE) score of 89 and a sequence coverage of 29%. MOWSE scores 65 and above are considered to be significant hits.



**Figure 41. Identification of *T. crunogena* carboxysomal small subunit (CbbS) in carboxysomes of the *cbbLS::Tc CbbLS* mutant via mass spectrometry**

M/z ratios were searched against protein databases as described in Figure 40. The *T. crunogena* carboxysomal small subunit, CbbS, was identified with a molecular weight search (MOWSE) score of 77 and a sequence coverage of 42%. MOWSE scores 65 and above are considered to be significant hits.



**Figure 42. Identification of *T. crunogena* noncarboxysomal small subunit (Cbbs) in carboxysomes of the *cbbs::Tc NC cbbs* mutant via mass spectrometry**

M/z ratios were searched against protein databases as described in Figure 40. The *T. crunogena* noncarboxysomal small subunit, Cbbs, was identified with a molecular weight search (MOWSE) score of 65 and a sequence coverage of 46%. MOWSE (MW search) scores 65 and above are considered to be significant hits.

In contrast, SDS-PAGE analysis of polyhedra purified from the *cbbL::Tc NC cbbL*, and *cbbLS::Tc NC cbbLS* mutants revealed that they lacked the CbbL and CbbS proteins. However, they did appear to contain the remaining carboxysomal shell components (**Figure 25**). This observation was in accordance with the electron microscopy findings that their polyhedra lacked Form IA RubisCO molecules.

To validate observations made from the SDS-PAGE and electron microscopy analysis, carboxysomes purified from the replacement mutants were subjected to immunoblotting using anti-CsoS1 and anti-CbbL antibodies, which recognize the CsoS1 shell proteins and the large subunit (CbbL) of Form IA RubisCO, respectively. It was seen that the anti-CsoS1 immunoblot identified the CsoS1C, -1A, and -1B shell proteins in microcompartments purified from all replacement mutants (**Figure 26B**). Moreover, mutant carboxysomes were found to contain approximately the same level of the CsoS1 shell proteins as their wild type counterparts (**Figure 26B**). On the other hand, the anti-CbbL antibody failed to recognize any CbbL protein in polyhedra purified from the *cbbL::Tc NC cbbL* and *cbbLS::Tc NC cbbL* mutants; however, CbbL was detected in purified carboxysomes of the *cbbS::Tc NC cbbS* and *cbbLS::Tc NC cbbLS* mutants at levels similar to those in wild type carboxysomes (**Figure 26A**). These results were consistent with those from the SDS-PAGE analysis and with the electron microscopy findings. Furthermore, an anti CbbM (Form II RubisCO) western blot detected no CbbM protein in carboxysomes of replacement mutants (**Figure 27**).

### CO<sub>2</sub> fixation by carboxysomes

To relate the growth phenotypes in air of the *cbbS::Tc NC cbbS* and *cbbLS::Tc C cbbLS* mutants to the activity of the chimeric and heterologous RubisCO species sequestered within their microcompartments, CO<sub>2</sub> fixation activity of mutant carboxysomes was determined using a radiometric assay that measures the conversion of radioactive bicarbonate to acid stable products [10]. These assays were performed by Zhicheng Dou. Results obtained revealed that carboxysomes isolated from the *cbbS::Tc NC cbbS* that contained chimeric Form IA RubisCO sequestered within them, exhibited only 10% of wild type carboxysome activity, while carboxysomes isolated from the *cbbLS::Tc C cbbLS* mutants that contained heterologous Form IA RubisCO sequestered within them, were 25% as active as wild type carboxysomes. To determine the activity of free RubisCO molecules, a freeze-thaw protocol was used to mechanically break intact carboxysomes and release the encapsulated RubisCO [98]. Results showed that the free RubisCO molecules released from carboxysomes of the *cbbS::Tc NC cbbS* and *cbbLS::Tc C cbbLS* mutants were catalytically compromised in comparison to their wild type counterparts (**Table 2**). Since carboxysome shells of the *cbbL::Tc NC cbbL* and *cbbLS::Tc NC cbbLS* mutants did not contain any Form IA RubisCO, it was no surprise that these polyhedra showed no CO<sub>2</sub> fixation activity (**Table 2**), which agreed well with the other results.

Genotype	Carboxysomes $\mu\text{mol min}^{-1} \text{mg}^{-1}$	Free RubisCO $\mu\text{mol min}^{-1} \text{mg}^{-1}$
<i>wild type</i>	$0.966 \pm 0.020$	$2.173 \pm 0.174$
<i>cbbL:: Tc NC cbbL</i>	0	– *
<i>cbbS:: Tc NC cbbS</i>	$0.096 \pm 0.003$	$0.146 \pm 0.009$
<i>cbbLS:: Tc NC cbbLS</i>	0	– *
<i>cbbLS:: Tc C cbbLS</i>	$0.234 \pm 0.003$	$0.295 \pm 0.010$
<i>cbbLS::kan<sup>r</sup></i>	0	– *

**Table 2. CO<sub>2</sub> fixation activities of carboxysomes and free RubisCO released from carboxysomes**

\*Only empty carboxysomes shells were assembled in these mutants.

## CHAPTER V

### DISCUSSION

Whether the biogenesis of carboxysomes and related microcompartments requires the participation of enzymes they normally sequester remains poorly understood. In this study, yeast two-hybrid screens were used to identify potential interactions between putative structural proteins and the ethanolamine ammonia lyase (EAL) enzyme of the EA utilization (Eut) organelles of *Salmonella enterica*. Strong interactions were observed between the large subunit of EAL and certain shell proteins, suggesting that these components may be essential in Eut organelle assembly. However, since the Eut organelles from *S. enterica* have not been purified and characterized to date, any conclusions concerning the role of encapsulated enzymes in microcompartment biogenesis that are based solely on the data obtained from yeast two-hybrid interaction studies, remain inconclusive.

To address the role of encapsulated enzymes in microcompartment biogenesis *in vivo* the model chemoautotrophic bacterium *Halothiobacillus neapolitanus*, for which a robust genetic manipulation system and a protocol for carboxysome purification have been well established [36, 42, 46], was used. A series of *H. neapolitanus* Form IA RubisCO mutants was generated and their phenotypes were characterized. Intriguingly, a mutant lacking the genes encoding Form IA RubisCO was found to assemble empty carboxysome shells of normal shape and composition [99]. The empty shells of this mutant did not sequester Form II RubisCO to compensate for the lack of the Form IA enzyme. Further, carboxysomes of *H.*

*neapolitanus* were found to sequester chimeric and heterologous species of RubisCO, a process that is determined by the enzyme's large subunit [99].

### **Interactions between EutB, EutN, and EutL may be central to Eut organelle assembly**

Not much is known about the Eut organelles of *S. enterica* that are believed to mediate the catabolism of ethanolamine (EA). By analogy to the 1,2-propanediol utilization (Pdu) organelles of *S. enterica* [72, 75, 77], it is speculated that the Eut microcompartments are filled with the enzyme EAL, which is responsible for the breakdown of EA [3, 21, 73]. Whether these enzymes contribute to the assembly or overall polyhedral shape of these organelles remains unclear. With regard to the *pdu* organelles, it was previously reported that deletion of genes encoding encapsulated enzymes did not seem to affect the assembly of microcompartment shells [2, 77]. Similar genetic studies with respect to the Eut organelles have not been performed.

To gain insight into the molecular interactions that might be important for the assembly of Eut organelles, yeast two-hybrid screens involving the presumed structural components of the shell and the EAL enzyme were performed. Strong interactions were observed between the enzyme's large subunit, EutB, and the putative shell proteins, EutN and EutL. A similar level of interaction was also observed between the EutL and EutN proteins. These findings are interesting since the EutL and EutN proteins have been grouped with the CcmK/CsoS1 bacterial microcompartment (BMC) family of proteins and CcmL/CsoS4 family of proteins, respectively, based on primary sequence similarities [3, 4, 7, 94]. An identical yeast

two-hybrid screen had previously shown strong interactions between the *H. neapolitanus* CsoS1B and CsoS4A structural proteins and the small subunit of the carboxysomal Form IA RubisCO [100]. Members of the CcmL/CsoS4 family of proteins typically form pentamers and are thought to occupy the vertices of the carboxysome shell; however, in crystals, the EutN protein from *E. coli* was recently shown to form hexamers [101]. It was speculated that either the Eut organelles do not require a pentameric protein to contribute to their geometry, which would explain their less regular shape compared to carboxysomes, or some protein other than EutN may substitute as a pentameric shell component [13].

Based on results obtained from the yeast two-hybrid screens, it may be tempting to assume that interactions between the EutB, EutN, and EutL proteins provide the basic framework for Eut organelle assembly. However, to validate such assumptions, the polypeptide composition of the Eut organelle will have to be identified. The lack of a purification protocol that would allow characterization of the structural components of the Eut organelles has hindered progress in testing results from the yeast two-hybrid screens by biochemical approaches. In this study, attempts to purify the Eut microcompartments were met with little success. When the method of Havemann and Bobik [19] was employed, a fraction that seemed to be partially enriched for the Eut organelles was identified based on EAL activity assays, anti-EutC western blotting, and electron microscopy; however there appeared to be considerable contamination, possibly by membrane and flagellar proteins, as judged by SDS-PAGE and electron microscopy analyses. It may be that the majority of the Eut organelles remained membrane bound and could not be released during the initial

purification steps involving detergent treatment and sonication. This would explain the recovery of only a small fraction of the Eut organelles in the final step of the purification protocol. Alternatively, the conditions employed for isolating the Eut microcompartments could have affected their stability. In either case, a modified purification strategy for isolating the Eut polyhedra will have to be devised. Once purified, the ultimate test to identifying core components of the Eut organelles would be to study the effect of microcompartment biogenesis in wild type and mutant *S. enterica* cells lacking a presumed structural component of the shell or the EAL enzyme. In fact, such mutants have been constructed and studied [21], but none addressed the role of individual components in Eut organelle biogenesis.

Therefore, to assess the *in vivo* role of encapsulated enzymes in microcompartment assembly, genetic and biochemical studies were performed in the chemoautotrophic bacterium *Halothiobacillus neapolitanus*, which has served as an excellent model for understanding various aspects of carboxysome structure and function [1].

### ***H. neapolitanus* assembles empty carboxysome shells in the absence of Form IA RubisCO**

Although the function of Form I RubisCO in CO<sub>2</sub> fixation has been well established [9, 32, 50, 102], its role in carboxysome biogenesis and as a determinant of overall architecture remains poorly understood. Based on observations in the filamentous cyanobacterium, *Anabaena variabilis* M3, Price and Badger proposed that the carboxysome shell is assembled prior to the insertion of RubisCO molecules

[32]. The empty polyhedral structures observed by Baker *et al.* in the *cbbL::Km* mutant [15] and cryo-electron tomograms showing the spatial arrangement of RubisCO molecules inside carboxysomes of *Synechococcus* strain WH8102 [43] suggest an assembly pathway for carboxysomes that does not rely on the presence of a RubisCO core. On the other hand, Orús *et al.* observed ring-shaped structures of different sizes in thin sections of *Synechococcus* PCC 7942 and proposed that a core of RubisCO molecules was necessary for the assembly of  $\beta$ -carboxysome shells [68]. The authors speculated that one of the earliest stages in the maturation process of  $\beta$ -carboxysomes involves an orderly arrangement of RubisCO molecules and that shells are added as the final step in the biogenesis pathway [68]. This view is also supported by observations of RubisCO molecules arranged in paracrystalline arrays in electron micrographs of isolated carboxysomes [40]. Moreover, results from recent studies involving *in vitro* affinity-based pull-down assays and *E. coli* co-expression studies of certain  $\beta$ -carboxysomal components revealed that the large subunit of Form IB cyanobacterial RubisCO, RbcL, interacts with CcmM [69, 70], a putative shell protein that harbors three RbcS-like repeats at its C-terminus and a domain similar to that of carbonic anhydrases and acetyltransferases at its N-terminus. The authors speculated that CcmM may serve as a nucleation point for recruiting RbcL. The CcmM-RbcL complex is subsequently thought to recruit carbonic anhydrase, and this subcomplex is envisaged to serve as a scaffold for  $\beta$ -carboxysome shells to assemble [69, 70]. However, since  $\beta$ -carboxysomes have not yet been purified to homogeneity, the stoichiometric ratios of individual carboxysomal components remain speculative.

Whether predictions made based on *in vitro* studies represent prerequisite cascades for  $\beta$ -carboxysome biogenesis *in vivo* remains to be seen.

To evaluate if Form IA RubisCO is necessary for carboxysome biogenesis, a *H. neapolitanus cbbLS::kan<sup>r</sup>* mutant was generated, in which the holoenzyme-encoding *cbbL* and *cbbS* genes were deleted and replaced with a kanamycin cassette. Biochemical and ultrastructural analyses of this mutant provided, for the first time, direct evidence that empty carboxysome shells of apparently normal size, shape, and composition were formed in the absence of Form IA RubisCO [99]. These carboxysome shells were relatively stable under the routine conditions employed to purify them. No Form II RubisCO was found to be associated with these empty polyhedra, suggesting that the dimeric form does not compensate for the lack of Form IA RubisCO in microcompartment assembly. Thus, it can be concluded from analysis of the *cbbLS::kan<sup>r</sup>* mutant that Form IA RubisCO does not participate in the assembly process of carboxysome shells and that the polyhedral shape of the carboxysome is not the result of the spatial distribution and filling of RubisCO molecules within the interior of these microcompartments. This inference is consistent with the empty polyhedral structures observed in the *cbbL::Km* mutant, wherein the coding region of the Form IA RubisCO large subunit-encoding *cbbL* gene was disrupted by the insertion of a kanamycin resistance cassette [15]. Likewise, it was observed in the chemoheterotrophic bacterium *Salmonella enterica* that deletion of genes encoding sequestered enzymes involved in the catabolism of 1,2-PD does not appear to affect the assembly of Pdu shell structures [2]. Recent cryo-electron tomography studies of

isolated carboxysomes from *Synechococcus* sp. WH8102 also suggest no contacts between the encapsulated RubisCO and the microcompartment shell [43].

The inability of the *cbbLS::kan<sup>r</sup>* mutant to grow in ambient CO<sub>2</sub> was reminiscent of the high CO<sub>2</sub>-requiring (*hcr*) phenotype reported by Baker *et al.* for the *cbbL::Km* mutant [15]. In another study, Pierce *et al.* [95] reported that replacement of the RubisCO *rbcL* gene in the cyanobacterium *Synechocystis* PCC 6803 by the *rbcM* gene encoding Form II RubisCO from the photosynthetic anaerobe *Rhodospirillum rubrum* resulted in a ‘*cyanorubrum*’ mutant that could not grow in air [95]. Likewise, a ‘*tobacco-rubrum*’ mutant, in which the endogenous hexadecameric Form I RubisCO of *Nicotiana tabacum* was replaced with the dimeric Form II RubisCO requires elevated levels of CO<sub>2</sub> for survival [103]. The lack of growth of the *cbbLS::kan<sup>r</sup>* mutant in air may be attributed to the absence of Form IA RubisCO and the lower efficiency of the Form II enzyme at ambient CO<sub>2</sub> levels [97].

**The *cso*-operon encoded copy of Form IA RubisCO is packaged within carboxysomes of *Thiomicrospira crunogena***

Genomes of certain chemolithotrophic bacteria such as *Thiomicrospira crunogena*, *Hydrogenovibrio marinus*, and *Acidithiobacillus ferrooxidans* are found to harbor an additional copy of Form IA RubisCO-encoding genes that is not part of the *cso* operon [16-18, 104]. The duplicate copy in these bacteria, also referred to as the noncarboxysomal or non *cso*-associated copy, is believed to be the result of gene duplication and reorganization during the evolutionary process [16]. In these bacteria, it remains unknown which copy of Form IA RubisCO is packaged within their

carboxysomes, especially since the large subunits of both enzymes are found to be close to 80% identical in their primary sequence. Nevertheless, it may seem logical to assume that the *csa* operon-associated copy is sequestered within their carboxysomes.

Studies in *Hy. marinus* reported that when cells were grown in an atmosphere enriched with CO<sub>2</sub>, only the *csa*-associated *cbbL* and *cbbS* genes are expressed, which also correlates with the presence of carboxysomes [16]. It was thus speculated that the product of the *csa*-associated copy is most likely sequestered within their polyhedra [16]. However, that study failed to investigate the formation of carboxysomes at ambient CO<sub>2</sub> levels, where both RubisCO copies are expressed and one would expect to observe microcompartments. Thus, which Form IA RubisCO copy is accommodated within carboxysomes remained inconclusive. To address this issue, carboxysomes from the deep sea vent  $\gamma$ -proteobacterium, *T. crunogena*, were purified and characterized. The large and small subunits of the RubisCO species associated with *T. crunogena* carboxysomes were identified as being *csa*-operon encoded by mass spectrometric analysis. This result suggests that between the two copies of Form IA RubisCO, carboxysomes of *T. crunogena* prefer the incorporation of the *csa* operon-encoded Form IA<sub>c</sub> copy. Whether the Form IA<sub>q</sub> enzyme in *T. crunogena* is functional is not known. It is possible that Form IA<sub>q</sub> RubisCO is not functional, which may be why carboxysomes do not favor this enzyme's sequestration. Results from the mass spectrometric analysis raised an important question regarding the ability of *T. crunogena* carboxysomes to sequester noncarboxysomal RubisCO in the absence of the carboxysomal copy. To test this possibility, one would have to delete the *csa* operon-associated *cbbL* and *cbbS* genes in *T. crunogena* and evaluate the

ability of its carboxysomes to accommodate the noncarboxysomal Form IAq RubisCO. However, a genetic system for generating mutants in *T. crunogena* has not been established yet.

**Carboxysomes of *T. crunogena* and *H. neapolitanus* are indistinguishable in morphology but display variations in shell protein composition**

Although microscopic observations suggested that carboxysomes of *T. crunogena* are similar in size and shape to those of *H. neapolitanus*, SDS-PAGE and western blotting analyses revealed a striking difference in their shell protein composition. It appeared that the *T. crunogena* CsoS2 shell protein may have been expressed as a full length polypeptide as opposed to the CsoS2A polypeptide found in *H. neapolitanus*. The observed molecular weight (MW) of the putative full length CsoS2 protein was found to be around 80 kDa, which exceeds its calculated MW by about 12 kDa. Such differences between the observed and predicted MW of proteins may be associated with post-translational modification like phosphorylation and glycosylation, or with inherent properties such as high isoelectric points (pIs) and resistance to complete denaturation of proteins.

An earlier study in *H. neapolitanus* had suggested that the CsoS2 protein likely undergoes post-translational O-linked glycosylation event, which causes it to migrate as two polypeptides of apparent MW 130 kDa and 85 kDa, respectively, upon separation by SDS-PAGE [52]. However, if this were to be the case, expression of the *H. neapolitanus* *csoS2* gene in *E. coli* should have resulted in the synthesis of a single, full length protein, since glycosylation is a rare event in *E. coli*. On the

contrary, CsoS2 is expressed as CsoS2B and CsoS2A polypeptides in the heterologous host [53], suggesting that differential glycosylation might not be the explanation for the existence of these two polypeptides.

Two alternative explanations for the expression of CsoS2B and CsoS2A polypeptides in *H. neapolitanus* have been proposed [53]. A programmed ribosome frame-shifting event may cause premature translational termination that results in the synthesis of a smaller CsoS2A polypeptide. This scenario does not seem to apply to *T. crunogena* CsoS2, because no smaller CsoS2A-like protein was detected in carboxysomes of this bacterium by SDS-PAGE and western blotting analyses. Alternatively, the CsoS2 protein might be subject to proteolytic cleavage by a protease found in *H. neapolitanus* and *E. coli*, which results in the formation of a full length CsoS2B polypeptide and a truncated CsoS2A form. That CsoS2A is a C-terminally truncated form of the full length CsoS2B polypeptide was recently shown by characterizing a *H. neapolitanus* mutant engineered to express a sequential peptide affinity (SPA) tag at the C-terminus of the CsoS2 protein, and by mass spectrometry analysis [53]. However, no such protease-mediated cleavage of CsoS2 appeared to be occurring in *T. crunogena* as judged by SDS-PAGE and western blotting analyses. Either *T. crunogena* does not synthesize such a protease or its CsoS2 protein does not contain a specific cleavage site. The latter scenario is appealing given that *T. crunogena* CsoS2 lacks stretches of amino acid residues towards the C-terminal half of the protein that are present in *H. neapolitanus* CsoS2.

The reason for the aberrant migration behavior of *T. crunogena* CsoS2 remains unknown. Although glycosylation of CsoS2 cannot be completely ruled out,

a simpler explanation may be incomplete denaturation of the protein when subjected to SDS-PAGE. This is true especially considering the protein's high isoelectric point (calculated  $pI = 9.7$  for *T. crunogena* CsoS2), which might influence the net charge of CsoS2 under denaturing conditions and result in slower electrophoresis, mobility and a higher than expected apparent MW. The N-terminus of *T. crunogena* CsoS2 is rich in proline residues, and previous studies have shown that this feature often leads to anomalous migration patterns during SDS-PAGE [105, 106]. Intriguingly, *T. crunogena* CsoS2 contains a  $(PXX)_n$  motif [105] ( $^{32}PAAPRKPAAPVAEPVVAAAPAPSQRSRRKVSAPTATP^{71}$ ), which is typically found in pancreatic polypeptide hormones [107] and neuropeptides [108] that assume the polyproline II conformation and are known to possess anomalous migration patterns on denaturing polyacrylamide gels [106]. This motif is absent in the predicted primary structure of *H. neapolitanus* CsoS2 polypeptides and the reason for the aberrant migration pattern of its CsoS2B and CsoS2A forms remains puzzling.

Based on these results, it seems reasonable to infer that carboxysome shell assembly proceeds irrespective of any cleavage events of the CsoS2 protein in *T. crunogena*. To investigate this possibility further, *csoS2* mutants are being constructed in *H. neapolitanus*, wherein the endogenous *csoS2* gene is either deleted or replaced with its counterpart from *T. crunogena*. Phenotypic characterization of these mutants and functional studies of their carboxysomes should also provide more insight into the requirement for CsoS2 in carboxysome shell assembly and the effect of heterologous CsoS2 expression on structure and function of *H. neapolitanus* carboxysomes.

## ***H. neapolitanus* carboxysomes accommodate chimeric and heterologous Form IA RubisCO species**

The finding that carboxysomes of *T. crunogena* were filled with the *cso*-associated copy of Form IA RubisCO raised an important question. Considering the high degree of homology between the large subunits of Form IAc and IAq enzymes [17], could carboxysomes of *T. crunogena* accommodate the noncarboxysomal Form IA RubisCO in the absence of the carboxysomal copy? The simplest way to test this idea would be to construct a *T. crunogena* mutant lacking the *cso*-associated *cbbL* and *cbbS* genes and to characterize the RubisCO species in its carboxysomes. Unfortunately, a genetic system for mutant construction in *T. crunogena* has not yet been established [Scott, KM, personal communication]. Hence, this possibility was tested in *H. neapolitanus* by constructing a series of Form IA RubisCO mutants, wherein the *cbbL* and *cbbS* genes were replaced with the carboxysomal and noncarboxysomal orthologs from *T. crunogena*.

It would seem reasonable to assume that the assembly of  $\alpha$ -carboxysomes in different autotrophic bacteria relies on species-specific interactions of their individual protein components, especially since primary sequence comparisons of the shell components have indicated that not all are particularly well conserved [55] and carboxysomes of *H. neapolitanus* and *T. crunogena* were shown to differ in their shell protein composition. However, it was surprising to find that carboxysomes of *H. neapolitanus* can incorporate chimeric and heterologous species of Form IA RubisCO [99]. The chimeric version was composed of *H. neapolitanus* large subunit and *T. crunogena* noncarboxysomal small subunit while the heterologous RubisCO was

composed of the *T. crunogena* carboxysomal large and small subunits. In both cases, mutant carboxysomes were found to be nearly identical in size, shape, and protein composition to their wild type counterparts.

Carboxysomes containing chimeric or heterologous Form IA RubisCO were enzymatically active, as judged by CO<sub>2</sub> fixation assays, but at reduced levels that corresponded well with the slower growth rates of these mutants in air compared to the wild type. The mild *hcr* phenotypes observed in the *cbbS::Tc NC cbbS* and *cbbLS::Tc C cbbLS* mutants is likely a reflection of the compromised carboxylation activity exhibited by the chimeric and heterologous RubisCO species. In the *cbbLS::Tc C cbbLS* mutant that harbored heterologous Form IA RubisCO, the poor performance of the enzyme may have been a manifestation of a lower intrinsic carboxylation activity of the carboxysomal *T. crunogena* enzyme compared to that of its *H. neapolitanus* orthologs. This possibility needs to be confirmed by determining the kinetic constants for the *T. crunogena* carboxysomal Form IA RubisCO.

Alternatively, holoenzyme assembly may be compromised in the heterologous host.

The reduced specific activity of carboxysomes possessing chimeric Form IA RubisCO in the *cbbS::Tc NC cbbS* mutant was not surprising. Previous studies, which mainly focused on evaluating the contribution of the small subunit to the structure and function of RubisCO, have shown that chimeras of Form I holoenzyme that are composed of subunits from different species have compromised carboxylation activities [109-114]. Co-expression of the *Synechococcus* sp. PCC 6301 large subunit (RbcL) and several higher plant small subunit (RbcS) in *E. coli* resulted in the assembly of hexadecameric hybrid enzymes with reduced activity [111, 112]. The

interaction between the large and small subunits in these hybrid enzymes was weak, and the small subunits easily dissociated from the L<sub>8</sub> core when the hybrid holoenzymes were passed through gel filtration or DEAE ion exchange columns [111]. A similar scenario may be encountered in the *cbbS::Tc NC cbbS* mutant if the molecular interactions between the *H. neapolitanus* large subunit (CbbL) and *T. crunogena* noncarboxysomal small subunit (CbbS) are less favorable than in the respective wild type holoenzyme. In fact, electron micrographs of negatively stained carboxysomes purified from the *cbbS::Tc NC cbbS* mutant did not show the characteristic donut-shaped RubisCO molecules that are clearly visible in wild type carboxysomes. This mutant's carboxysomes appeared to be filled with irregularly shaped clusters, which may represent protein aggregates.

**The large subunit, CbbL, determines whether Form IA RubisCO is incorporated into carboxysomes**

Based on results obtained with the different Form IA RubisCO replacement mutants, it could be concluded that some structural feature of the carboxysomal RubisCO protein that is not present in the noncarboxysomal species is required for the incorporation of Form IAc RubisCO into carboxysomes of *H. neapolitanus*. The primary sequences of the carboxysomal and noncarboxysomal CbbL proteins found in different bacteria exhibit a high degree of identity, in agreement with the highly conserved structure and function of the free enzyme [56, 60]. The CbbS proteins, on the other hand, display considerably more sequence divergence [56]. It would thus seem reasonable to predict that the small subunit determines whether a particular

RubisCO species can be accommodated within carboxysomes. Badger and Bek [56] had proposed that a six amino acid insertion at the N-terminus of the small subunit of Form IAq RubisCO could possibly disrupt protein interactions that may be necessary for holoenzyme sequestration within carboxysomes [56]. However, results from this study showed that the *T. crunogena* noncarboxysomal CbbS protein, which harbors the additional N-terminal six amino acid residues, associated with the *H. neapolitanus* large subunit, CbbL, and was incorporated into carboxysomes as part of a chimeric RubisCO holoenzyme [99]. Furthermore, only those mutants expressing endogenous or heterologous carboxysomal CbbL protein were able to sequester the resulting hybrid RubisCO species into their carboxysomes regardless of the CbbS species. On the other hand, mutants expressing RubisCO species with noncarboxysomal CbbL failed to sequester the resulting hybrid enzyme within their carboxysomes, even if it was paired with endogenous *H. neapolitanus* carboxysomal CbbS protein. These mutants assembled empty polyhedra like those observed in the *cbbLS::kan<sup>r</sup>* mutant [99]. From these results, it could be concluded that the large and not the small subunit of Form IA RubisCO dictates holoenzyme sequestration within carboxysomes.

The growth phenotypes of the mutants that failed to sequester RubisCO species within their carboxysome shells were similar to that of the *cbbLS::kan<sup>r</sup>* mutant. However, these mutants did assemble functional holoenzyme as judged by CO<sub>2</sub>-fixing assays of clarified cell extracts [99]. Presumably, in these mutants the hybrid RubisCO species, rather than being sequestered within carboxysome shells, were most likely dispersed in the cytosol. Thus, the absence of a localized environment for RubisCO to ensure efficient CO<sub>2</sub> fixation in these mutants likely

resulted in their *hcr* phenotype. Alternatively, the noncarboxysomal holoenzyme expressed within these mutants may be less equipped to distinguish between its competing substrates, CO<sub>2</sub> and O<sub>2</sub>. A low specificity factor has been previously reported for the noncarboxysomal enzyme in *Hy. marinus* in comparison to its carboxysomal counterpart [16]. Purification and a thorough kinetic characterization of those RubisCO species that fail to be incorporated within carboxysomes will be a prerequisite for providing a more informed explanation about their ability to function in low CO<sub>2</sub> environments.

In light of the finding that cells expressing RubisCO species that contain noncarboxysomal CbbL could still assemble empty polyhedra, it is difficult to explain why carboxysomes have not been observed in the chemoautotrophic bacterium, *Thiobacillus denitrificans*. The *csa* operon of this organism contains an entire set of carboxysome shell genes but lacks those encoding the large and small subunits of Form IA RubisCO [55]. Instead, a non-*csa* operon associated set of *rbcL* and *rbcS* genes has been located adjacent to the *cbbO/cbbQ* gene cluster that is believed to encode structural proteins aiding in RubisCO assembly [56]. It was previously reported that expression of this *rbcL-rbcS* gene set in a RubisCO null mutant of the purple bacterium *Rhodobacter sphaeroides* was able to yield an active holoenzyme and complement the mutant phenotype [97]. Since *R. sphaeroides* does not assemble carboxysomes, this study did not address the issue of enzyme sequestration.

Sequence comparisons identified the six amino acid insertion motif at the N-terminus of the *T. denitrificans* RbcS and led to the speculation that the presence of this noncarboxysomal RubisCO species is the reason why microcompartments have

not yet been observed in this bacterium [56]. In *T. denitrificans*, the noncarboxysomal enzyme may have evolved in an environment containing a high CO<sub>2</sub>/O<sub>2</sub> concentration ratio, thereby circumventing the organism's need for a highly efficient enzyme adept in discriminating between CO<sub>2</sub> and O<sub>2</sub> [97]. Moreover, since *T. denitrificans* can grow under anaerobic conditions, where the concentration of CO<sub>2</sub> is high and the Form II enzyme is presumably expressed, the organism may not require carboxysomes for efficient CO<sub>2</sub> fixation [97]. However, based on results from this study, one would expect to find empty carboxysome shells in cells of *T. denitrificans*. A careful expression analysis of the *cso* shell proteins in *T. denitrificans*, an attempt to purify the potentially empty shells, and an assessment of its ability to assemble carboxysomes or empty polyhedra under aerobic and anaerobic conditions are thus warranted.

Primary sequence alignment of the *cso* and non-*cso* encoded CbbL proteins, showed no striking differences that would explain the selective large-subunit dependent sequestration of RubisCO into carboxysomes of *H. neapolitanus*. Because all CbbL proteins share a high degree of identity [56], it seems likely that any structural differences between the carboxysomal and noncarboxysomal large subunits are subtle. Crystal structure determination followed by a careful comparison with the known structure of the *H. neapolitanus* wild type carboxysomal RubisCO (PDB ID: 1SVD, Kerfeld CA *et al.*, 2005) should be able to highlight these differences. Subsequent site-directed mutagenesis of strategic amino acids may help in identifying residues that contribute to the selective large subunit-mediated sequestration of Form IA RubisCO.

The finding that the assembly of carboxysomal shells does not require Form IA RubisCO is consistent with the biogenesis model proposed by Price and Badger [32]. However, several questions about the RubisCO-independent shell biogenesis model still remain unanswered. Whether the empty shells that assemble in the absence of Form IA RubisCO retain the icosahedral geometry of wild type carboxysomes remains unclear. Besides, the nature of the contents packaged within the empty polyhedra remains unknown. It is possible that the interior of the empty polyhedra is filled with either cytosolic contents or CO<sub>2</sub>, assuming that the shell-associated carbonic anhydrase in these polyhedra are functional. Finally, the surprising discovery that carboxysomes of *H. neapolitanus* incorporate chimeric and heterologous Form IA RubisCO opens possibilities for designing engineered microcompartments that could accommodate foreign proteins. Eventually, such engineered microcompartments could be used in nanotechnology-based applications such as biomedicine.

## CHAPTER VI

### CONCLUSIONS AND FUTURE WORK

By far the most important conclusion drawn from this study is that carboxysome shell assembly and sequestration of Form IA RubisCO are two functionally linked yet independent processes. This conclusion is based on the finding that all *H. neapolitanus* mutants characterized in this study assemble carboxysome shells of apparently normal size and shape, irrespective of whether Form IA RubisCO is packaged. Contrary to the sequence-based prediction made by Badger and Bek [56] that the small subunit is responsible for determining which species of Form IA RubisCO can be sequestered, this study clearly identifies the large subunit as an important determinant of enzyme sequestration within carboxysomes. Finally, the discovery that carboxysomes of *H. neapolitanus* can accommodate chimeric and heterologous species of Form IA RubisCO has paved the way for designing engineered protein microcompartments for use in nanotechnology-based applications such as biomedicine.

Although the polyhedral shape of carboxysomes does not seem to be the result of filling by RubisCO molecules, whether the empty polyhedra assembled in the absence of Form IA RubisCO retain the icosahedral geometry of wild type carboxysomes remains unclear. To gain better insights into the three-dimensional architecture of the empty polyhedra, cryo-electron tomography studies of the Form IA RubisCO deletion mutant are currently being performed in collaboration with Dr. Grant Jensen at the California Institute of Technology. The fine structure data

obtained from the tomograms should be able to highlight structural differences between the empty polyhedra and wild type carboxysomes *in vivo*. In addition, the tomograms may also provide information whether the carboxysomal shells assembled in the Form IA RubisCO deletion mutant are truly empty or contain random cytosolic proteins.

Primary sequence comparisons of the highly conserved carboxysomal and noncarboxysomal CbbL proteins found in several chemoautotrophic bacteria [56] have not revealed any obvious residues or signature that would help to explain the selective large subunit-specific sequestration of RubisCO within carboxysomes. Nevertheless, sequence alignments of all the CbbL proteins found in the chemoautotrophic bacteria *H. neapolitanus*, *Hy. marinus*, *T. crunogena*, and *A. ferrooxidans* have identified certain amino acid residues that are exclusive to either the carboxysomal or noncarboxysomal large subunits. The residues in the carboxysomal large subunit <sup>25</sup>Leu, <sup>333</sup>Ala, <sup>334</sup>Ser, <sup>342</sup>Leu, and <sup>344</sup>Glu are substituted with Lys, Glu, Ala, Met, and Asp amino acids, respectively, in the noncarboxysomal CbbL proteins. To test if these residues hold the key to selective Form IA RubisCO sequestration within carboxysomes, site-directed mutagenesis within the *H. neapolitanus cbbL* gene will have to be performed. Next, the ability of the mutated *cbbL* genes to encode large subunits competent of forming functional holoenzymes and mediating sequestration of Form IA RubisCO within carboxysomes will have to be evaluated by generating and characterizing *H. neapolitanus* mutants like those described in this study. Aside from this approach, efforts are currently underway to purify and deduce the crystal structure of the noncarboxysomal Form IA RubisCO

species that are not sequestered within carboxysomes. Structure comparisons between the noncarboxysomal and carboxysomal RubisCO species (PDB ID: 1SVD, Kerfeld CA *et al.*, 2005) should be able to accentuate subtle differences that may be key to enzyme sequestration. Such studies will lay the groundwork for the design of engineered microcompartments for use in nanotechnology-based applications such as biomedicine.

Yeast two-hybrid screens identified strong protein interactions between the putative shell components EutN and EutL, and the large subunit of ethanolamine ammonia lyase, EutB, of *Salmonella enterica*. The possibility that these proteins comprise essential components that participate in the assembly of the ethanolamine utilization (Eut) organelles was raised. However, without information about the polypeptide composition of the Eut microcompartments, results from the interaction studies remain tentative. The purification protocol adopted in this study yielded a fraction that seemed to be partially enriched for the Eut microcompartments. In order to purify and characterize the Eut polyhedra, the existing protocol will have to be modified. Milder sonication conditions and/or alternative detergents like octyl glucoside, which is typically used for stripping integral membrane proteins, may have to be used for obtaining a homogenous preparation of the Eut organelles from *S. enterica* cells. To evaluate the progress of the revised purification strategy, the ethanolamine ammonia lyase assay, which was standardized in this study, will have to be used in conjunction with SDS-PAGE, anti-EutC western blotting, and electron microscopy studies. Once purified, Eut microcompartment biogenesis can be

evaluated in *S. enterica* mutants lacking structural components or the ethanolamine ammonia enzyme.

## APPENDIX

A1. Oligonucleotides used for constructing *Halothiobacillus neapolitanus* Form IA RubisCO mutants

Primer name	Primer sequence
NC <i>cbbL</i> F	5'- <u>GGATCC</u> ACAGGGGGTGGCGGAATCCCCCATCCTTTCAGGAGGAACTCATGGCTAAGACTTATAACGCCGGTG-3'
NC <i>cbbL</i> R	5'- <u>GGTACCTT</u> ACTTATGCTTAACATCTAGCTTGTCAACTGTGTCGAAATTCGAACTTGATTTCTTCCAAAGTTCCAT-3'
NC <i>cbbL-kan<sup>r</sup></i> F	5'- <u>GGTACCCCGGA</u> ATGCCAGCTGGGGCCCTCTGGTAAGGTTGGGAAGCCCTGC AAAGTAAACTGGATGGCTTTC-3'
NC <i>cbbL-kan<sup>r</sup></i> R	5'- <u>CTCGAGGC</u> ATTTTCAGCCATGGTTACTCACCTTAGTATGTTGTGGTACGAGGGATCAGAAAGAACTCGTCAAGAA-3'
NC <i>cbbS</i> F	5'- <u>GGATCCC</u> GGTTGATCCCTCGTACCACACAACATACTAAGGTGAGTAACCATGAGTATTCAAGATTACCCATCTC-3'
NC <i>cbbS</i> R	5'- <u>GGTACCTT</u> ACATGTCA CCAAGCTTAACAAGCATGTTAGCA CTTGAGACTGTGCA TAGTTATCATAAACCAATCAAAC-3'
NC <i>cbbS-kan<sup>r</sup></i> F	5'- <u>GGTACCCCGGA</u> ATTGCCAGCTGGGGCCCTCTGGTAAGGTTGGGAAGCCCTGC AAAGTAAACTGGATGGCTTTC-3'
NC <i>cbbS-kan<sup>r</sup></i> R	5'- <u>CTCGAGTGA</u> ATAAGTGTGCTGCTAGAGAAACGCACAGCCGAATGACAGACTTGA CTCAAGAAAGAACTCGTCAAGAA-3'

Underlined sequences are homologous with the 5' and 3' flanking regions of the *cbbL* and *cbbS* genes in *Halothiobacillus neapolitanus*

Italicized sequences represent restriction sites used for cloning purposes

**A1. Oligonucleotides used for constructing *Halothiobacillus neapolitanus* Form IA RubisCO mutants (Cont'd)**

Primer name	Primer sequence
NC <i>cbbLS F</i>	5'-GGATCCACAGGGGGTTCGGCGAATCCCCATCCCTTTCAGGAGGAAGCACTCATG GCTAAGACTTATAACGCCGGTG-3'
NC <i>cbbLS R</i>	5'-GGTACCTTACATGTACCCACGCTTAACAAGCATGTTAGCACCTTGAGACTGTGC ATAGTTATCAATAACCAATCAAAC-3'
NC <i>cbbLS-kan' F</i>	5'-GTACCCCGGAATTGCCAGCTGGGGCCCTCTGGTAAGGTTGGGAAGCCCTGC AAAGTAAACTGGATGGCTTTC-3'
NC <i>cbbLS-kan' R</i>	5'-TCGAGTGAATAAGTGTGTGCGTAGAGAAACGCACAGCGCAATGACAGACTTG ACTCAGAAAGAACTCGTCAAGAA-3'
C <i>cbbLS F</i>	5'-GGATCCCGTTGATCCCTCGTACCACACAAACATACTAAGGTGAGTAACCATGGC AAGTAAAACGTTTGATGCTG-3'
C <i>cbbLS R</i>	5'-GGTACCTTACATGCCACGTGGGGGATAACCACAAAGTTGTGACCTTGGCACTG AGTATAGTTGTCAATCCCAACCA-3'
C <i>cbbLS-kan' F</i>	5'-GGTACCCCGGAATTGCCAGCTGGGGCCCTCTGGTAAGGTTGGGAAGCCCTGC CAAAGTAAACTGGATGGCTTTC-3'
C <i>cbbLS-kan' R</i>	5'-CTCGAGTGAATAAGTGTGTGCGTAGAGAAACGCACAGCGCAATGACAGACTT GACTCAGAAAGAACTCGTCAAGAA-3'

Underlined sequences are homologous with the 5' and 3' flanking regions of the *cbbL* and *cbbS* genes in *Halothiobacillus neapolitanus*

Italicized sequences represent restriction sites used for cloning purposes

**A1. Oligonucleotides used for constructing *Halothiobacillus neapolitanus* Form IA RubisCO mutants**

Primer name	Primer sequence
<i>kar<sup>f</sup> F</i>	5'- <u>GGATCCCAAGCCACAGGGGGGTTGGCGGAATCCCCCATCCTTTCAGGGAGGAA</u> <u>CTCCCGGAATTGCCAGCTGGGGC</u> -3'
<i>kar<sup>f</sup> R</i>	5'- <u>CTCGAGTGAATAAGTGTGCTGCGTAGAGAAAACGCACAGCGCAATGACAGACT</u> <u>TGACTCAGAAAGAACTCGTCAAGAA</u> -3'

Underlined sequences are homologous with the 5' and 3' flanking regions of the *cbbL* and *cbbS* genes in *Halothiobacillus neapolitanus*

Italicized sequences represent restriction sites used for cloning purposes

**A2. Oligonucleotides used in generating *eut* gene constructs to be used in the  
Yeast Two Hybrid Screen**

EutBfNdeIY2H	5'-GGTGGT <i>CATATGAAACTAAAGACCACATTG</i> -3'
EutBrBamHIY2H	5'-GTGGTGGATCCTCAGAAGAACAGTGACGGA-3'
EutCfNdeIY2H	5'-GGTGGT <i>CATATGGATCAAAAACAGATTGAA</i> -3'
EutCrBamHIY2H	5'-GTGGTGGATCCTTAACGGGTCATGTTGATG-3'
EutKfNdeIY2H	5'-GGTGGT <i>CATATGATCAATGCCCTGGGATTA</i> -3'
EutKrBamHIY2H	5'-GTGGTGGATCCTTAATTTTTGATGCGATAG-3'
EutLfNdeIY2H	5'-GGTGGT <i>CATATGCCTGCATTAGATTTAATT</i> -3'
EutLrBamHIY2H	5'-GTGGTGGATCCTTACGCACGCTGGACAGGG-3'
EutMfNdeIY2H	5'-GGTGGT <i>CATATGGAAGCATTAGGAATGATT</i> -3'
EutMrBamHIY2H	5'-GTGGTGGATCCTCAAATGTTGCTGTGCCT-3'
EutNfNdeIY2H	5'-GGTGGT <i>CATATGGAGGCGGATATGAAACTG</i> -3'
EutNrBamHIY2H	5'-GTGGTGGATCCCTATTTATGGAAAACCACT-3'
EutSfNdeIY2H	5'-GGTGGT <i>CATATGAATAAAGAACGCATTATT</i> -3'
EutSrBamHIY2H	5'-GTGGTGGATCCTTAACTTTTTGTAACTCA-3'

Underlined sequences represent coding regions of the *eut* genes.

Italicized sequences represent restriction sites used for cloning purposes

**A3. Oligonucleotides used in generating *eut* gene constructs for protein expression**

EutKfBamH1	5'- <u>AGGATCC</u> <u>ATGATCAATGCCCTGGGA</u> -3'
EutKrXhoI	5'- <u>ACTCGAGTTAATTTTTGATGCGATA</u> -3'
EutLfBamH1	5'- <u>AGGATCC</u> <u>ATGCCTGCATTAGATTTA</u> -3'
EutLrXhoI	5'- <u>ACTCGAGTTACGCACGCTGGACAGG</u> -3'
EutMfBamH1	5'- <u>AGGATCC</u> <u>ATGGAAGCATTAGGAATG</u> -3'
EutMrXhoI	5'- <u>ACTCGAGTCAAATGTTGCTGTCGCC</u> -3'
EutNfBamH1	5'- <u>AGGATCC</u> <u>ATGGAGGCGGATATGAAA</u> -3'
EutNrXhoI	5'- <u>ACTCGAGCTATTTATGGAAAACCCAC</u> -3'
EutSfBamH1	5'- <u>AGGATCC</u> <u>ATGAATAAAGAACGCATT</u> -3'
EutSrXhoI	5'- <u>ACTCGAGTTAACTTTTTGTTAACTC</u> -3'
EutCfBamH1	5'- <u>AGGATCC</u> <u>ATGGATCAAA AACAGATT</u>
EutCrXhoI	5'- <u>ACTCGAGTTAACGGGTCATGTTGAT</u>

Underlined sequences represent coding regions of the *eut* genes.

Italicized sequences represent restriction sites used for cloning purposes

## REFERENCES

1. Yeates, T.O., et al., *Protein-based organelles in bacteria: carboxysomes and related microcompartments*. Nat Rev Micro, 2008. **6**(9): p. 681-691.
2. Bobik, T.A., et al., *The propanediol utilization (pdu) operon of Salmonella enterica serovar Typhimurium LT2 includes genes necessary for formation of polyhedral organelles involved in coenzyme B(12)-dependent 1, 2-propanediol degradation*. J Bacteriol, 1999. **181**(19): p. 5967-5975.
3. Kofoed, E., et al., *The 17-gene ethanolamine (eut) operon of Salmonella typhimurium encodes five homologues of carboxysome shell proteins*. J Bacteriol, 1999. **181**(17): p. 5317-29.
4. Stojiljkovic, I., A.J. Baumler, and F. Heffron, *Ethanolamine utilization in Salmonella typhimurium: nucleotide sequence, protein expression, and mutational analysis of the cchA cchB eutE eutJ eutG eutH gene cluster*. J Bacteriol, 1995. **177**(5): p. 1357-66.
5. Lewis, P.J., Thaker, S. D., Errington, J., *Compartmentalization of transcription and translation in Bacillus subtilis*. Embo J, 2000. **19**(4): p. 710-718.
6. Lewis, P.J., *Bacterial subcellular architecture: recent advances and future prospects*. Molecular Microbiology, 2004. **54**(5): p. 1135-1150.
7. Heinhorst, S., G.C. Cannon, and J.M. Shively, *Carboxysomes and carboxysome-like inclusions*, in *Microbial Monographs*, J.M. Shively, Editor. 2006, Springer-Verlag: Berlin Heidelberg. p. 141-164.

8. Gitai, Z., *The New Bacterial Cell Biology: Moving Parts and Subcellular Architecture*. Cell, 2005. **120**(5): p. 577-586.
9. Cannon, G.C., et al., *Microcompartments in prokaryotes: carboxysomes and related polyhedra*. Appl Environ Microbiol, 2001. **67**(12): p. 5351-5361.
10. Dou, Z., et al., *CO<sub>2</sub> fixation kinetics of Halothiobacillus neapolitanus mutant carboxysomes lacking carbonic anhydrase suggest the shell acts as a diffusional barrier for CO<sub>2</sub>*. J. Biol. Chem., 2008. **283**(16): p. 10377-84.
11. Tsai, Y., et al., *The structure of the shell protein CsoS1A from Halothiobacillus neapolitanus and its implications for carboxysome function*. PLoS Biology, 2007. **5**(6): p. 1345-1354.
12. Yeates, T.O., *Self-assembly in the carboxysome: a viral capsid-like protein shell in bacterial cells*. Biochem Soc Trans, 2007. **035**: p. 508-511.
13. Tanaka, S., et al., *Atomic-level models of the bacterial carboxysome shell*. Science, 2008. **319**: p. 1083-1086.
14. Williams, E.B., *Identification and Characterization of Protein Interactions in the Carboxysome of Halothiobacillus neapolitanus*. Doctoral Dissertation, 2006. The University of Southern Mississippi, MS.
15. Baker, S.H., et al., *Insertion mutation of the form I cbbL gene encoding ribulose biphosphate carboxylase/oxygenase (RuBisCO) in Thiobacillus neapolitanus results in expression of form II RuBisCO, loss of carboxysomes, and an increased CO<sub>2</sub> requirement for growth*. J Bacteriol, 1998. **180**(16): p. 4133-4139.

16. Yoshizawa, Y., et al., *CO<sub>2</sub>-responsive expression and gene organization of three ribulose-1,5-bisphosphate carboxylase/oxygenase enzymes and carboxysomes in Hydrogenovibrio marinus strain MH-110*. J Bacteriol, 2004. **186**(17): p. 5685-5691.
17. Scott, K.M., et al., *The genome of deep-sea vent chemolithoautotroph Thiomicrospira crunogena XCL-2*. PLoS Biology, 2006. **4**(12): p. 2196-2212.
18. Heinhorst, S., et al., *Two Copies of Form I RuBisCO Genes in Acidithiobacillus ferrooxidans ATCC 23270*. Current Microbiology, 2002. **45**(2): p. 115-117.
19. Havemann, G.D. and T.A. Bobik, *Protein content of polyhedral organelles involved in coenzyme B12-dependent degradation of 1,2-propanediol in Salmonella enterica serovar Typhimurium LT2*. J Bacteriol, 2003. **185**(17): p. 5086-5095.
20. Sriramulu, D.D., et al., *Lactobacillus reuteri DSM 20016 produces cobalamin-dependent diol dehydratase in metabolosomes and metabolises 1,2-propanediol by disproportionation*. J. Bacteriol., 2008: p. JB.01535-07.
21. Brinsmade, S.R., T. Paldon, and J.C. Escalante-Semerena, *Minimal Functions and Physiological Conditions Required for Growth of Salmonella enterica on Ethanolamine in the Absence of the Metabolosome*. J. Bacteriol., 2005. **187**(23): p. 8039-8046.
22. Morris, D.M. and G.J. Jensen, *Toward a Biomechanical Understanding of Whole Bacterial Cells*. Annual Review of Biochemistry, 2008. **77**(1): p. 583-613.

23. Jacobs, C. and L. Shapiro, *Bacterial cell division: A moveable feast*. Proceedings of the National Academy of Sciences of the United States of America, 1999. **96**(11): p. 5891-5893.
24. Badger, M.R. and G.D. Price, *CO<sub>2</sub> concentrating mechanisms in cyanobacteria: molecular components, their diversity and evolution*. J Exp Bot, 2003. **54**(383): p. 609-622.
25. Price, G.D., et al., *Advances in understanding the cyanobacterial CO<sub>2</sub>-concentrating-mechanism (CCM): functional components, Ci transporters, diversity, genetic regulation and prospects for engineering into plants*. J. Exp. Bot., 2008. **59**(7): p. 1441-1461.
26. Shibata, M., et al., *Distinct constitutive and low-CO<sub>2</sub>-induced CO<sub>2</sub> uptake systems in cyanobacteria: genes involved and their phylogenetic relationship with homologous genes in other organisms*. Proc Natl Acad Sci U S A, 2001. **98**(20): p. 11789-94.
27. Price, G.D., et al., *Identification of a SulP-type bicarbonate transporter in marine cyanobacteria*. Proc Natl Acad Sci U S A, 2004. **101**(52): p. 18228-18233.
28. Shibata, M., et al., *Genes essential to sodium-dependent bicarbonate transport in cyanobacteria: function and phylogenetic analysis*. J Biol Chem, 2002. **277**(21): p. 18658-64.
29. Dobrinski, K.P., D.L. Longo, and K.M. Scott, *The carbon-concentrating mechanism of the hydrothermal vent chemolithoautotroph Thiomicrospira crunogena*. J Bacteriol, 2005. **187**(16): p. 5761-5766.

30. Price, G.D. and M.R. Badger, *Isolation and characterization of high CO<sub>2</sub>-requiring-mutants of the cyanobacterium Synechococcus PCC 7942: Two phenotypes that accumulate inorganic carbon but are apparently unable to generate CO<sub>2</sub> within the carboxysome*. *Plant Physiol.*, 1989. **91**: p. 514-525.
31. Price, G.D. and M.R. Badger, *Expression of Human Carbonic Anhydrase in the Cyanobacterium Synechococcus PCC7942 Creates a High CO<sub>2</sub>-Requiring Phenotype : Evidence for a Central Role for Carboxysomes in the CO<sub>2</sub> Concentrating Mechanism*. *Plant Physiol.*, 1989. **91**(2): p. 505-513.
32. Price, G.D. and M.R. Badger, *Evidence for the role of carboxysomes in the cyanobacterial CO<sub>2</sub>-concentrating mechanism*. *Can. J. Bot.*, 1991. **69**: p. 963-973.
33. Badger, M.R. and G.D. Price, *The CO<sub>2</sub> concentrating mechanism in cyanobacteria and microalgae*. *Physiol Plant*, 1992. **90**: p. 529-536.
34. Badger, M.R., et al., *The environmental plasticity and ecological genomics of the cyanobacterial CO<sub>2</sub> concentrating mechanism*. *J Exp Bot*, 2006. **57**: p. 249-265.
35. Yu, J.-W., et al., *Isolation of a Putative Carboxysomal Carbonic Anhydrase Gene from the Cyanobacterium Synechococcus PCC7942*. *Plant Physiol*, 1992. **100**: p. 794-800.
36. So, A.K., et al., *A novel evolutionary lineage of carbonic anhydrase (epsilon class) is a component of the carboxysome shell*. *J Bacteriol*, 2004. **186**(3): p. 623-630.

37. Codd, G.A. and W.J.N. Marsden, *The carboxysomes (polyhedral bodies) of autotrophic prokaryotes*. Biol. Rev., 1984. **59**: p. 389-422.
38. Shively, J.M. and R.S. English, *The carboxysome, a prokaryotic organelle: a mini review*. Can. J. Bot., 1991. **69**: p. 957-962.
39. Shively, J.M., et al., *Functional organelles in prokaryotes: Polyhedral inclusions (carboxysomes) in Thiobacillus neapolitanus*. Science, 1973. **182**: p. 584-586.
40. Shively, J.M., F.L. Ball, and B.W. Kline, *Electron microscopy of the carboxysomes (polyhedral bodies) of Thiobacillus neapolitanus*. J Bacteriol, 1973. **116**: p. 1405-1411.
41. Holthuijzen, Y.A., et al., *Electron microscopic studies of carboxysomes of Thiobacillus neapolitanus*. Archives of Microbiology, 1986. **144**(3): p. 258-262.
42. Cannon, G.C. and J.M. Shively, *Characterization of a homogeneous preparation of carboxysomes from Thiobacillus neapolitanus*. Arch Microbiol, 1983. **134**: p. 52-59.
43. Iancu, C.V., et al., *The structure of isolated Synechococcus strain WH8102 carboxysomes as revealed by electron cryotomography*. J Mol Biol, 2007. **372**: p. 764-773.
44. Schmid, M.F., et al., *Structure of Halothiobacillus neapolitanus carboxysomes by cryo-electron tomography*. J Mol Biol, 2006. **364**: p. 526-535.

45. Badger, M.R., D. Hanson, and G.D. Price, *Evolution and diversity of CO<sub>2</sub> concentrating mechanisms in cyanobacteria*. *Funct Plant Biol*, 2002. **29**: p. 161-173.
46. English, R.S., S. Jin, and J.M. Shively, *Use of electroporation to generate a Thiobacillus neapolitanus carboxysome mutant*. *Appl Environ Microbiol*, 1995. **61**: p. 3256-3260.
47. Price, G.D., et al., *Analysis of a genomic DNA region from the cyanobacterium Synechococcus sp. strain PCC7942 involved in carboxysome assembly and function*. *J. Bacteriol.*, 1993. **175**: p. 2871-2879.
48. Kerfeld, C.A., et al., *Protein structures forming the shell of primitive bacterial organelles*. *Science*, 2005. **309**(5736): p. 936-938.
49. Shiho Tanaka, M.R.S.M.P.T.O.Y., *Insights from multiple structures of the shell proteins from the beta-carboxysome*. *Protein Science*, 2009. **18**(1): p. 108-120.
50. Cannon, G.C., *Carboxysomes and CO<sub>2</sub> fixation in Thiobacillus neapolitanus*. Doctoral Dissertation, 1982. Clemson University.
51. Cai, F., et al., *Transcript analysis of the Halothiobacillus neapolitanus cso operon*. *Archives of Microbiology*, 2008. **189**(2): p. 141-150.
52. Baker, S.H., et al., *The correlation of the gene csoS2 of the carboxysome operon with two polypeptides of the carboxysome in Thiobacillus neapolitanus*. *Arch Microbiol*, 1999. **172**(4): p. 233-239.

53. Dou, Z., *Functional characterization and assembly studies of carboxysomes in Halothiobacillus neapolitanus*. Doctoral Dissertation, 2009. The University of Southern Mississippi.
54. Sawaya, M.R., et al., *The structure of beta-carbonic anhydrase from the carboxysomal shell reveals a distinct subclass with one active site for the price of two*. J Biol Chem, 2006. **281**: p. 7546-7555.
55. Cannon, G.C., et al., *Organization of carboxysome genes in the Thiobacilli*. Curr Microbiol, 2003. **46**(2): p. 115-119.
56. Badger, M.R. and E.J. Bek, *Multiple Rubisco forms in proteobacteria: their functional significance in relation to CO<sub>2</sub> acquisition by the CBB cycle*. J. Exp. Bot., 2008. **59**(7): p. 1525-1541.
57. Berg, I.A., et al., *A 3-Hydroxypropionate/4-Hydroxybutyrate Autotrophic Carbon Dioxide Assimilation Pathway in Archaea*. Science, 2007. **318**(5857): p. 1782-1786.
58. Thauer, R.K., *A Fifth Pathway of Carbon Fixation*. Science, 2007. **318**(5857): p. 1732-1733.
59. Gutteridge, S. and A.A. Gatenby, *Rubisco Synthesis, Assembly, Mechanism, and Regulation*. Plant Cell, 1995. **7**(7): p. 809-819.
60. Tabita, F.R., et al., *Distinct form I, II, III, and IV Rubisco proteins from the three kingdoms of life provide clues about Rubisco evolution and structure/function relationships*. J. Exp. Bot., 2008. **59**(7): p. 1515-1524.
61. Sato, T., H. Atomi, and T. Imanaka, *Archaeal Type III RuBisCOs Function in a Pathway for AMP Metabolism*. Science, 2007. **315**(5814): p. 1003-1006.

62. Li, H., et al., *Crystal Structure of a RuBisCO-like Protein from the Green Sulfur Bacterium Chlorobium tepidum*. 2005. **13**(5): p. 779-789.
63. Tabita, F.R., et al., *Function, Structure, and Evolution of the RubisCO-Like Proteins and Their RubisCO Homologs*. Microbiol. Mol. Biol. Rev., 2007. **71**(4): p. 576-599.
64. Beudeker, R.F., et al., *Relations between D-ribulose-1, 5-bisphosphate carboxylase, carboxysomes, and CO<sub>2</sub> fixing capacity in the obligate chemolithotroph Thiobacillus neapolitanus grown under different limitations in the chemostat*. Arch Microbiol, 1980. **124**: p. 185-189.
65. Purohit, K., B.A. McFadden, and M.M. Shaykh, *D-Ribulose-1,5-bisphosphate carboxylase and polyhedral inclusion bodies in Thiobacillus intermedius*. J Bacteriol, 1976. **127**(1): p. 516-522.
66. Holthuijzen, Y.A., et al., *Protein composition of the carboxysomes of Thiobacillus neapolitanus*. Archives of Microbiology, 1986. **144**(4): p. 398-404.
67. Heinhorst, S., et al., *Characterization of the carboxysomal carbonic anhydrase CsoSCA from Halothiobacillus neapolitanus*. J Bacteriol, 2006. **188**(23): p. 8087-8094.
68. Orus, M.I., et al., *Biogenesis and ultrastructure of carboxysomes from wild type and mutants of Synechococcus sp. strain PCC 7942*. Plant Physiol, 1995. **107**(4): p. 1159-1166.

69. Long, B.M., et al., *Analysis of Carboxysomes from Synechococcus PCC7942 Reveals Multiple Rubisco Complexes with Carboxysomal Proteins CcmM and CcaA*. J. Biol. Chem., 2007. **282**(40): p. 29323-29335.
70. Cot, S.S.-W., A.K.-C. So, and G.S. Espie, *A Multiprotein Bicarbonate Dehydration Complex Essential to Carboxysome Function in Cyanobacteria* J. Bacteriol., 2008. **190**(3): p. 936-945.
71. Daniel Emlyn-Jones, F.J.W.T.J.A.G.D.P.S.M.W., *A Synechococcus PCC7942 ccmM (cyanophyceae) Mutant Pseudoreverts to Air Growth without Regaining Carboxysomes*. Journal of Phycology, 2006. **42**(4): p. 769-777.
72. Bobik, T.A., *Bacterial microcompartments: their properties and paradoxes*. Bioessays, 2008. **30**(11-12): p. 1084-95.
73. Penrod, J.T. and J.R. Roth, *Conserving a volatile metabolite: a role for carboxysome-like organelles in Salmonella enterica*. J Bacteriol, 2006. **188**: p. 2865-2874.
74. Chen, P., D.I. Andersson, and J.R. Roth, *The control region of the pdu/cob regulon in Salmonella typhimurium*. J. Bacteriol., 1994. **176**(17): p. 5474-5482.
75. Bobik, T.A., *Polyhedral organelles compartmenting bacterial metabolic processes*. Appl Environ Microbiol, 2006. **70**: p. 517-525.
76. Havemann, G.D., E.M. Sampson, and T.A. Bobik, *PduA is a shell protein of polyhedral organelles involved in coenzyme B(12)-dependent degradation of 1,2-propanediol in Salmonella enterica serovar typhimurium LT2*. J Bacteriol, 2002. **184**(5): p. 1253-1261.

77. Bobik, T.A., *Bacterial Microcompartments*. Microbe, 2007. **2**(1): p. 25-31.
78. Roof, D.M. and J.R. Roth, *Ethanolamine utilization in Salmonella typhimurium*. J. Bacteriol., 1988. **170**(9): p. 3855-3863.
79. Faust, L.R., et al., *Cloning, sequencing, and expression of the genes encoding the adenosylcobalamin-dependent ethanolamine ammonia-lyase of Salmonella typhimurium*. J. Biol. Chem., 1990. **265**(21): p. 12462-12466.
80. Mori, K., et al., *Identification of a Reactivating Factor for Adenosylcobalamin-Dependent Ethanolamine Ammonia Lyase*. J. Bacteriol., 2004. **186**(20): p. 6845-6854.
81. Faust, L.P., Babior, B. M., *Overexpression, purification, and some properties of the AdoCbl-dependent ethanolamine ammonia-lyase from Salmonella typhimurium*. Arch Biochem Biophys, 1992. **294**(1): p. 50-54.
82. Starai, V.J., J. Garrity, and J.C. Escalante-Semerena, *Acetate excretion during growth of Salmonella enterica on ethanolamine requires phosphotransacetylase (EutD) activity, and acetate recapture requires acetyl-CoA synthetase (Acs) and phosphotransacetylase (Pta) activities*. Microbiology, 2005. **151**(11): p. 3793-3801.
83. Wackett, L.P., et al., *Genomic and Biochemical Studies Demonstrating the Absence of an Alkane-Producing Phenotype in Vibrio furnissii M1*. Appl. Environ. Microbiol., 2007. **73**(22): p. 7192-7198.
84. Yu, D., et al., *An efficient recombination system for chromosome engineering in Escherichia coli*. Proceedings of the National Academy of Sciences of the United States of America, 2000. **97**(11): p. 5978-5983.

85. Vishniac, W. and M. Santer, *THE THIOBACILLI*. Microbiol. Mol. Biol. Rev., 1957. **21**(3): p. 195-213.
86. Gietz RD, S.R., Willems AR, Woods RA., *Studies on the transformation of intact yeast cells by the LiAc/SS-DNA/PEG procedure*. Yeast, 1995. **11**(4): p. 355-360.
87. Smith, P.K., et al., *Measurement of protein using bicinchoninic acid*. Analytical Biochemistry, 1985. **150**(1): p. 76-85.
88. Toraya, T., et al., *Studies on the mechanism of the adenosylcobalamin-dependent diol dehydrase reaction by the use of analogs of the coenzyme*. J. Biol. Chem., 1977. **252**(3): p. 963-970.
89. Fields, S. and O.-k. Song, *A novel genetic system to detect protein-protein interactions*. Nature, 1989. **340**(6230): p. 245-246.
90. Albrecht, G., et al., *Monitoring the Gcn4 Protein-mediated Response in the Yeast *Saccharomyces cerevisiae**. J. Biol. Chem., 1998. **273**(21): p. 12696-12702.
91. Viale, A.M., H. Kobayashi, and T. Akazawa, *Expressed genes for plant-type ribulose 1,5-bisphosphate carboxylase/oxygenase in the photosynthetic bacterium *Chromatium vinosum*, which possesses two complete sets of the genes*. J Bacteriol, 1989. **171**(5): p. 2391-400.
92. Kusian, B., et al., *Characterization of the duplicate ribulose-1,5-bisphosphate carboxylase genes and *cbb* promoters of *Alcaligenes eutrophus**. J Bacteriol, 1995. **177**(15): p. 4442-50.

93. Gibson, J.L. and F.R. Tabita, *Isolation and preliminary characterization of two forms of ribulose 1,5-bisphosphate carboxylase from Rhodospseudomonas capsulata*. J Bacteriol, 1977. **132**(3): p. 818-23.
94. Shively, J.M., et al., *Sequence homologues of the carboxysomal polypeptide CsoS1 of the thiobacilli are present in the cyanobacteria and enteric bacteria that form carboxysomes-polyhedral bodies*. Can. J. Bot., 1998. **76**: p. 906-916.
95. Pierce, J., T.J. Carlson, and J.G.K. Williams, *A Cyanobacterial Mutant Requiring the Expression of Ribulose Bisphosphate Carboxylase from a Photosynthetic Anaerobe*. Proceedings of the National Academy of Sciences, 1989. **86**(15): p. 5753-5757.
96. Amichay, D., R. Levitz, and M. Gurevitz, *Construction of a Synechocystis PCC6803 mutant suitable for the study of variant hexadecameric ribulose bisphosphate carboxylase/oxygenase enzymes*. Plant Mol Biol, 1993. **23**(3): p. 465-76.
97. Hernandez, J.M., et al., *Deduced amino acid sequence, functional expression, and unique enzymatic properties of the form I and form II ribulose bisphosphate carboxylase/oxygenase from the chemoautotrophic bacterium Thiobacillus denitrificans*. J Bacteriol, 1996. **178**(2): p. 347-56.
98. Dou, Z., et al., *CO<sub>2</sub> fixation kinetics of Halothiobacillus neapolitanus mutant carboxysomes lacking carbonic anhydrase suggest the shell acts as a diffusional barrier for CO<sub>2</sub>*. J Biol Chem, 2008. **in press**.

99. Menon, B.B., et al., *Halothiobacillus neapolitanus* Carboxysomes Sequester Heterologous and Chimeric RubisCO Species. PLoS ONE, 2008. **3**(10): p. e3570.
100. Williams, E.B., et al. *Protein interactions in carboxysomes from Halothiobacillus neapolitanus*. in *5th International Symposium on Inorganic Carbon Utilization by Aquatic Photosynthetic Organisms*. 2004. Saint Sauveur, CN.
101. Forouhar, F., et al., *Functional insights from structural genomics*. Journal of Structural and Functional Genomics, 2007. **8**(2): p. 37-44.
102. Cannon, G.C., R.S. English, and J.M. Shively, *In situ assay of ribulose-1,5-bisphosphate carboxylase/oxygenase in Thiobacillus neapolitanus*. J Bacteriol, 1991. **173**(4): p. 1565-1568.
103. Whitney, S.M. and T.J. Andrews, *Photosynthesis and growth of tobacco with a substituted bacterial Rubisco mirror the properties of the introduced enzyme*. Plant Physiol, 2003. **133**(1): p. 287-94.
104. Nishihara, H., et al., *Phylogenetic position of an obligately chemoautotrophic, marine hydrogen-oxidizing bacterium, Hydrogenovibrio marinus, on the basis of 16S rRNA gene sequences and two form I RuBisCO gene sequences*. Arch Microbiol, 1998. **169**(4): p. 364-8.
105. Williamson, M.P., *The structure and function of proline-rich regions in proteins*. Biochem J, 1994. **297**(2): p. 249-60.
106. Ziemer, M.A., A. Mason, and D.M. Carlson, *Cell-free translations of proline-rich protein mRNAs*. J. Biol. Chem., 1982. **257**(18): p. 11176-11180.

107. Blundell T. L., P.J.E., Tickle I. J., Wood S. P., and Wu C. W., *X-ray analysis (1. 4-Å resolution) of avian pancreatic polypeptide: Small globular protein hormone*. Proc Natl Acad Sci U S A, 1981. **78**(7): p. 4175-4179.
108. Darbon H, B.J., Deleuze C, Chenu J, Roussel A, Cambillau C., *Solution conformation of human neuropeptide Y by <sup>1</sup>H nuclear magnetic resonance and restrained molecular dynamics*. Eur J Biochem, 1992. **209**(2): p. 765-771.
109. Read, B.A. and F.R. Tabita, *Amino acid substitutions in the small subunit of ribulose-1,5-bisphosphate carboxylase/oxygenase that influence catalytic activity of the holoenzyme*. Biochemistry, 1992. **31**(2): p. 519-25.
110. Andrews, T.J. and G.H. Lorimer, *Catalytic properties of a hybrid between cyanobacterial large subunits and higher plant small subunits of ribulose bisphosphate carboxylase- oxygenase*. J. Biol. Chem., 1985. **260**(8): p. 4632-4636.
111. Wang, Y.-L., et al., *Properties of Hybrid Enzymes between Synechococcus Large Subunits and Higher Plant Small Subunits of Ribulose-1,5-bisphosphate Carboxylase/Oxygenase in Escherichia coli*. Archives of Biochemistry and Biophysics, 2001. **396**(1): p. 35-42.
112. Saskia M. van der Vies, D.B., Anthony A. Gatenby, *Assembly of cyanobacterial and higher plant ribulose bisphosphate carboxylase subunits into functional homologous and heterologous enzyme molecules in Escherichia coli*. EMBO J., 1986. **5**(10): p. 2439-2444.
113. Spreitzer, R.J., *Role of the small subunit in ribulose-1,5-bisphosphate carboxylase/oxygenase*. Arch Biochem Biophys, 2003. **414**: p. 141-149.

114. Kanevski, I., et al., *Plastome engineering of ribulose-1,5-bisphosphate carboxylase/oxygenase in tobacco to form a sunflower large subunit and tobacco small subunit hybrid*. *Plant Physiol*, 1999. **119**(1): p. 133-42.