THÈSE
Présentée devant
Université Pierre et Marie CURIE

Pour l'obtention du
Diplôme de DOCTORAT
Par
Sten König

The bacterial symbiont in the shallow water lucinids *Codakia orbicularis* and *C. orbiculata* analyzed by physiological proteogenomics

directeur:
Prof. Hervé Le Guyader (Université Pierre-et-Marie-Curie (UPMC))

co-directeur:
Prof. Dr. Olivier Gros (Université Antilles Guyane/UPMC)
Prof. Dr. Thomas Schweder (University of Greifswald, Institute of Pharmacy)

rapporteur:
Prof. Dr. Horst Felbeck (University of California, Scripps Institution of Oceanography)
Prof. Dr. Katharina Riedel (University of Greifswald, Institute of Microbiology)

Date de soutenance: 17. 12. 2014

Résumé:

Les bivalves côtiers *Codakia orbicularis* et *C. orbiculata,* de la famille des *Lucinidae*, abritent des Gammaprotéobactéries endosymbiotiques sulfo-oxydantes dans leurs branchies. Ces deux bivalves vivent dans les herbiers à *Thalassia testudinum* et hébergent la même bactérie symbiotique selon les analyses effectuées à partir des séquences d'ADNr 16S. Lors de période de stabulation, la population bactérienne symbiotique décroit alors qu'il n'y a pas, dans le même temps, de relargage des symbiotes observé. Des analyses en cytochimie ont montré une forte activité d'enzymes lysosomales lors de ces épisodes de privation de nourriture et de soufre. Il a ainsi été montré que les symbiotes peuvent servir directement de source de nourriture aux bivalves pour survivre lors de ces périodes de crise. Le transfert de carbone des symbiotes vers l'hôte peut être flexible et pourrait consister en un simple transfert de matière organique ou "milking", dans des conditions normales de nutrition et de digestion des symbiotes et devenir du "farming", dans des conditions de stabulation.

Jusqu'à ce jour, le symbiote reste non cultivable. De ce fait, l'utilisation de techniques indépendantes de la culture comme les approches –omics ont été mises en place pour étudier la physiologie de cette bactérie symbiotique. Le génome du symbiote a été analysé par Next Generation Sequencing (NGS) permettant ainsi d'obtenir les bases du protéome et ainsi de pouvoir analyser la physiologie du symbiote. Dans ce travail, le protéome des bactéries a été analysé sous différentes conditions.

L'oxydation des sulfures est une des voies métaboliques clés du symbiote de *Codakia.* Cette voie fait très probablement appel au système Sox périplasmique, ainsi qu'à une sulfite réductase cytoplasmique (DsrAB), une APS réductase (AprAB) et une ATP sulfurylase (SopT). De plus, deux autres enzymes additionnelles d'oxydation des sulfures ont pu être mises en évidence dans l'espace périplasmique du symbiote telle que la quinone réductase (Sqr) et la sulfide déhydrogénase (FccAB). Les gènes des enzymes du cycle de Calvin Benson Bassham (CBB) ne semblent pas être tous présents dans le génome du symbiote. Les protéines de la RuBisCO sont abondamment exprimées. Il semblerait que la régénération du ribulose-1,5-bisphosphate soit effectuée de façon non conventionnelle via une phosphofructokinase PPi-dépendante. Une autre caractéristique du CBB est qu'il y a deux formes différentes de RuBisCO codées dans le génome du symbiote. Les deux formes sont exprimées en même temps, mais la forme I de la RuBisCO est 50 fois plus exprimée que la forme II. En plus de la vie autotrophique, plusieurs gènes nécessaires à une vie hétérotrophique sont présents dans le génome. Dans le protéome, les enzymes de la glycolyse et du cycle TCA sont faiblement exprimées. Les protéines du métabolisme du glycogène ont également été identifiées dans le protéome. De plus, plusieurs types de transporteurs comme ABC, TRAP et PTS sont présents dans le génome et certaines formes d'expression de ces transporteurs ont pu être suspectées, y compris lors de la vie intracellulaire du symbiote.

De façon inattendue, un groupe de gènes *nif* est présent dans le génome permettant la fixation de l'azote atmosphérique par le symbiote. Les protéines codées par les gènes clés, comme la nitrogénase NifH/K/D, ont été abondamment trouvées dans le

protéome. De plus, l'analyse du protéome montre une régulation forte de ces protéines dans des conditions de stabulation du bivalve hôte. La rubrerythrine est fortement exprimée et servirait à protéger la nitrogénase de l'oxygène au sein des bacteriocytes.
L'endosymbiote bactérien code également pour un système de sécrétion de type 6 (T6SS) pour le transport de molécules bactériennes effectrices à travers les membranes du cytoplasme de la cellule hôte et jouerait un rôle possible de communication directe avec l'hôte.
En résumé, les analyses génomiques et protéomiques du symbiote bactérien des *Codakia* améliorent notre connaissance des voies métaboliques majeures des bivalves Lucinidae. L'ensemble des données de cette étude pourra servir à de futures analyses plus poussées de la physiologie et des interactions métaboliques entre l'hôte et ses symbiotes.

mot-clé: symbiose, *Codakia orbicularis, Codakia orbiculata,* soufre metabolisme, proteomics, genomics

key words: symbiosis, *Codakia orbicularis, Codakia orbiculata*, sulfur metabolism, proteomics, genomics

This PhD thesis includes results of a paper this was accepted 10. November. 2014 in the journal Microscopy Research and Technique.

Thioautotrophic bacterial endosymbionts are degraded by enzymatic digestion during starvation: case study of two lucinids - *Codakia orbicularis* and *C. orbiculata.*

Sten König[1], Hervé Le Guyader[2] and Olivier Gros[1,3]

[1] *Institut de Biologie Paris-Seine, UMR 7138 - Evolution Paris-Seine, Equipe Biologie de la Mangrove. Université des Antilles et de la Guyane, UFR des Sciences Exactes et Naturelles, Département de Biologie, BP 592. 97159 Pointe-à-Pitre cedex, Guadeloupe, France.*

[2] *Sorbonne Universités Paris VI, Institut de Biologie Paris-Seine, UMR 7138 - Evolution Paris-Seine, Equipe Phylogénie, Anatomie, Evolution. C.N.R.S, Institut de Biologie Paris-Seine, UMR 7138 - Evolution Paris-Seine, Equipe Biologie de la Mangrove.*

[3] *C_3MAG, UFR des Sciences Exactes et Naturelles, Université des Antilles et de la Guyane, BP 592 - 97159 Pointe-à-Pitre, Guadeloupe, French West Indies.*

Table of contents

1. Introduction 1

1.1 An overview of symbiosis 1

1.1.1 Marine chemosynthesis 3

1.2 Chemoautotrophic symbiosis 4

1.3 *Codakia orbicularis* and *C. orbiculata* and its chemoautotrophic symbiont ... 6

1.4 Ecosystem *Thalassia testudinum* - habitat of *Codakia* 12

1.5 Objectives - aim of this thesis 13

2. Material and Methods 14

2.1 Material 14

2.1.1 Software and databases 14

2.1.2 Chemicals 14

2.2 Methods 14

2.2.1 Sampling of bivalves 14

2.2.2 Bivalve incubation - Induction of symbiont's loss 15

2.2.3 Cytochemistry - arylsulfatase and acid phosphatase detection 15

2.2.4 Purification of the bacterial fraction - density centrifugation 16

2.2.5 Sequencing of the *Codakia* symbionts 17

2.2.6 Gene annotation 17

2.2.7 Database creation for proteomics 18

2.2.8 Metabolic Pathways 18

2.2.9 Proteomics 18

2.2.9.1 Extraction of soluble proteins 18

2.2.9.2 Enrichment of the membrane proteins 18

2.2.9.3 1D gel SDS PAGE 19

2.2.9.4 In-gel protein digestion and ESI MS/MS analyses 20

2.2.9.5 Protein identification and data analysis with Scaffold 20

2.2.9.6 2D SDS PAGE 21

2.2.9.7 2D gel protein spot digestion and identification with MALDI-ToF 21

2.2.9.8 Gel free proteomics - Synapt G2 22

2.2.9.9 Gene expression comparison 23

3. Results 24

3.1 Genomics 24

3.1.1 General genome characteristics 24

3.1.2 Genome comparison 24

3.1.3 Genome annotation 26

3.1.4 Database creation 27

3.2 Proteomics 28

3.2.1 2D - PAGE 29

3.2.1.1 Theoretical 2D gel protein distribution 29

3.2.1.2 2D gels of freshly *C. orbicularis* and *C. orbiculata* symbionts 30

3.2.1.3 *C. orbicularis* incubation and 2D gels of the symbiont 34

3.2.2 Proteome of freshly *C. orbicularis* and *C. orbiculata* symbionts 37

3.2.3 Protein functions in freshly collected *Codakia* symbiont 42

3.2.3.1 Carbon metabolism 42

3.2.3.1.1 Carbohydrate metabolism 42

3.2.3.1.2 Calvin Benson Bassham Cycle (CBB) 49

3.2.3.2 Energy metabolism ... 52
3.2.3.2.1 Oxidative phosphorylation ... 52
3.2.3.2.2 Sulfur metabolism ... 52
3.2.3.3 Nitrogen metabolism ... 55
3.2.3.3.1 Nitrate reduction ... 55
3.2.3.3.2 Nitrogen fixation ... 58
3.2.3.4 Amino acid ... 63
3.2.3.5 Stress-related proteins ... 63
3.2.3.6 Motility / Chemotaxis ... 64
3.2.3.7 Environmental information processing ... 64
3.2.4 Symbiont proteome after one week of *C. orbicularis* incubation ... 68
3.2.5 Protein functions in *C. orbicularis* incubation ... 72
3.2.5.1 Carbon metabolism ... 72
3.2.5.1.1 Carbohydrate metabolism ... 72
3.2.5.1.2 Calvin Benson Bassham Cycle (CBB) ... 74
3.2.5.2 Energy metabolism ... 76
3.2.5.2.1 Oxidative phosphorylation ... 76
3.2.5.2.2 Sulfur metabolism ... 76
3.2.5.3 Nitrogen fixation ... 79
3.2.5.4 Stress-related proteins ... 80
3.2.6 Proteome of only symbiont incubation for 24 h ... 81
3.2.7 Supplementary material from the gel free analysis ... 84
3.3 Endosymbiont digestion in the bivalves *Codakia* - Cytochemistry ... 98
4. Discussion ... 102
4.1 Habitat characterization of the seagrass bed *T. testudinum* ... 102
4.2 Genomics ... 104
4.3 Proteomics - Physiology of the endoymbiont ... 106
4.3.1 Carbon metabolism ... 108
4.3.1.1 Carbohydrate metabolism ... 108
4.3.1.2 Calvin Benson Bassham Cycle (CBB) ... 112
4.3.2 Sulfur metabolism ... 115
4.3.2.1 Sox system - Sulfur oxidation ... 119
4.3.2.2 Dsr-complex - dissimilatory sulfide reductase complex ... 121
4.3.3 Oxidative phosphorylation ... 123
4.3.4 Nitrogen metabolism ... 124
4.3.4.1 Gaseous nitrogen fixation ... 124
4.3.4.2 Nitrate reduction ... 130
4.4 Amino acid and lipid metabolism ... 131
4.5 Cytochemistry ... 132
4.6 Symbioses - interaction between host and symbiont ... 133
4.7 Conclusion ... 135
4.8 Outlook ... 137
5. References ... 138
6. Supplementary material ... 156

List of tables

Table 1.1. Overview about symbiosis relevant redox reactions and their energetic yield. 4
Table 2.1. Software and *in silico* tools. 14
Table 2.2. Composition of 1D SDS gel. 19
Table 2.3. Chemicals for the 1D gel SDS PAGE. 20
Table 3.1. General predicted genome statistics of the *Codakia* symbiont genomes. 24
Table 3.2. Key genes comparison of *C. orbicularis* and *C. orbiculata* symbiont. 25
Table 3.3. Jspecies genome comparison. 26
Table 3.4. Overview about the identified symbiont proteins of freshly collected bivalves. 38
Table 3.5. Overview about the glycogen synthesis and degradation genes in *C. orbicularis* symbiont genome. 45
Table 3.6. Putative reverse TCA-cycle enzymes encoded in the *Codakia* symbiont genome. 47
Table 3.7. The RuBisCO gene cluster in *C. orbicularis* symbiont genome. 50
Table 3.8. NCBI protein blast of PP-PfkA and HppA. 51
Table 3.9. QmoABC proteins encoded in the *Codakia* symbiont genome. 53
Table 3.10. Putative dissimilatory nitrate reduction proteins encoded in the *C. orbicularis* symbiont genome. 55
Table 3.11. Assimilatory nitrate reduction and related proteins encoded in the *C. orbicularis* symbiont genome. 58
Table 3.12. Putative nitrogen metabolism regulatory proteins in the *Codakia* symbiont genome. 60
Table 3.13. Putative *rnf* gene cluster in *C. orbicularis* symbiont genome. 61
Table 3.14. Encoded TRAP type proteins in the *Codakia* symbiont. 64
Table 3.15. Type VI secretion system gene cluster in the *Codakia* symbiont genome. 65
Table 3.16. Sec/Tat- related secretion proteins encoded in the *Codakia* symbiont genome. 66
Table 3.17. Miscellaneous transporter encoded in the *Codakia* symbiont genome. 67
Table 3.18. NSAF quantification in the membrane enrichments of the "porine Gram-negative proteins". 68
Table 3.19. Overview about the sure identified proteins of the *C. orbicularis* symbiont after incubation. 69
Table 3.20. Identified glycolysis and related enzymes in *C. orbicularis* and *C. orbiculata* symbiont proteome. 84
Table 3.21. Identified TCA-cycle and related proteins in *C. orbicularis and C. orbiculata* symbiont proteome. 86
Table 3.22. Identified PPP enzymes in the symbiont proteome *C. orbicularis* and *C. orbiculata* bivalves. 87
Table 3.23. Identified CBB related proteins in *C. orbicularis and C. orbiculata* symbiont proteome. 88
Table 3.24. Identified proteins of the ATP synthase in *C. orbicularis and C. orbiculata* symbiont proteome. 89

Table 3.25. Identified proteins of the Sox system in *C. orbicularis and C. orbiculata* symbiont proteome. 90

Table 3.26. Identified proteins of the Dsr- sulfur oxidizing pathway in *C. orbicularis and C. orbiculata* symbiont proteome..... 91

Table 3.27. Identified sulfur related proteins in *C. orbicularis and C. orbiculata* symbiont proteome... 92

Table 3.28. Identified nitrogen fixation proteins in *C. orbicularis and C. orbiculata* symbiont proteome. 93

Table 3.29. Putative nitrogenase protection proteins of the nitrogenase in *C. orbicularis and C. orbiculata* symbiont proteome. 95

Table 3.30. Key proteins of the amino acid metabolism in the *C. orbicularis and C. orbiculata* symbiont proteome. 96

Table 3.31. Identified stress proteins in *C. orbicularis and C. orbiculata* symbiont proteome. 97

Table 4.1. Field geochemical data from the seagrass bed *Thalassium testudinum* in Guadeloupe... 103

Table 6.1. Putative sulfide dehydrogenase (FccAB) proteins encoded in the *C. orbicularis* symbiont genome. 157

Table 6.2. Overview of the Sox-proteins encoded in the *C. orbicularis* symbiont genome..... 158

Table 6.3. Sulfur metabolism proteins of the symbiont in *C. orbicularis*. 159

Table 6.4. *dsr* gene cluster encoded in the *C. orbicularis* symbiont genome. 160

Table 6.5. *nif* gene cluster in *C. orbicularis* symbiont genome..... 161

Table 6.6. Oxidative phosphorylation related proteins in *C. orbicularis and C. orbiculata* symbiont proteome. 164

Table 6.7. Gene cluster of the motility genes in the *C. orbicularis* symbiont genome. 167

Table 6.8. Encoded ABC transporter in the *Codakia* symbiont genome..... 170

List of figures

Figure 1.1. Growth of symbiosis field. 1
Figure 1.2. Ecological types of interaction between organisms. 3
Figure 1.3. (A) - *Codakia orbicularis* and (B) - C. *orbiculata.* 6
Figure 1.4. Phylogenetic tree of sulfur oxidizing Gammaproteobacteria. 7
Figure 1.5. Natural habitat of the *Codakia* bivalves. 8
Figure 1.6. Morphology of the *Codakia* bivalves and gill organization. 9
Figure 1.7. Electron microscopy of the gill of *C. orbicularis.* 10
Figure 1.8. Overview about the known physiology of the *Codakia* symbiont. 12
Figure 2.1. Endosymbiont purification with Percoll density centrifugation. 17
Figure 2.2. Schema of the gel free analysis with the SynaptG2. 23
Figure 2.3. Color coding used for the visualization of the $\log_2$ ratio. 23
Figure 3.1. Overview of the gene function of *C. orbicularis* symbiont. 27
Figure 3.2. Overview of the database creation. 28
Figure 3.3. Theoretical distribution of all encoded *Codakia* symbionts proteins. 29
Figure 3.4. Master 2D gel of the intracellular soluble proteome of the *C. orbicularis* symbiont. 31
Figure 3.5. Master 2D gel of the intracellular soluble proteome of the *C. orbiculata* symbiont. 32
Figure 3.6. False color 2D gel of the symbiont of freshly collected *C. orbicularis* and *C. orbiculata*. 33
Figure 3.7. Comparison of the 2D gels of the soluble proteins of *C. orbicularis* symbiont incubation samples. 35
Figure 3.8. False color 2D gel (control/+[O_2]) 36
Figure 3.9. False color 2D gel picture (+[O_2]/-[O_2]) 37
Figure 3.10. 1D-SDS-PAGE of the soluble and membrane symbiont fraction. 38
Figure 3.11. Venn diagram comparison of the identified proteins in freshly *C. orbicularis* and *C. orbiculata* symbionts. 39
Figure 3.12. Comparison of the TMH's of the membrane enriched and the soluble protein fraction from *C. orbicularis* symbiont 1D gel Orbitrap analysis. 40
Figure 3.13. Overview about the functions of the identified proteins with more than two TMHs in *C. orbicularis* symbiont. 40
Figure 3.14. Comparison of the most abundant metabolic key proteins of freshly *C. orbicularis* and *C. orbiculata* symbionts. 41
Figure 3.15. Identified carbohydrate metabolism related enzymes of the *C. orbicularis* and *C. orbiculata* symbiont 43
Figure 3.16. Putative glycolysis/gluconeogenesis pathway of the *Codakia* symbiont. 44
Figure 3.17. Comparison ($\log_2$ ratio) of the identified nitrogen fixation proteins in *C. orbicularis and C. orbiculata* symbiont. 44
Figure 3.18. Overview about the different putative glycogen synthesis and degradation pathways encoded in the *Codakia* symbiont genome. 46
Figure 3.19. Predicted TCA-cycle in *Codakia* symbiont. 47

Figure 3.20. Comparison (log_2 ratio) of the identified TCA-cycle proteins in *C. orbicularis and C. orbiculata* symbiont. 48

Figure 3.21. Schematic overview about the reconstructed PPP in the symbiont of *Codakia*. 49

Figure 3.22. RuBisCO gene cluster encoded in the *C. orbicularis* symbiont genome (ID 312-321). 50

Figure 3.23. The PP-PfkA and HppA are co-localized encoded in *Codakia* symbiont genome. 51

Figure 3.24. Identified proteins of the CBB in the *Codakia* symbiont in gel free approach. 51

Figure 3.25. Comparison (log_2 ratio) of the identified CBB proteins in *C. orbicularis and C. orbiculata* symbiont. 52

Figure 3.26. Identified proteins in the gel free approach, related to the sulfur metabolism in freshly *Codakia* symbionts. 54

Figure 3.27. Comparison (log_2 ratio) of the identified sulfur oxidation metabolism proteins in *C. orbicularis and C. orbiculata* symbiont. 54

Figure 3.28. Nitrogen fixation related gene cluster encoded in the *C. orbicularis* symbiont genome. 58

Figure 3.29. Overview about the best NCBI blast hit of the 44 nitrogen fixation related proteins in the *C. orbicularis* symbiont genome. 59

Figure 3.30. Identified nitrogen fixation genes in the *Codakia* symbiont proteome. 59

Figure 3.31. Comparison (log_2 ratio) of the identified nitrogen fixation proteins in *C. orbicularis and C. orbiculata* symbiont. 60

Figure 3.32. The phylogenetic affiliation, NCBI blast hits, of RubY encoded in *C. orbicularis*. 62

Figure 3.33. The NCBI blast hits of RubY encoded in *C. orbicularis* dependent on the oxygen requirements. 62

Figure 3.34. Key genes in amino acid metabolism encoded in the *Codakia* symbiont genome. 63

Figure 3.35. Alignment of the "porin Gram-negative type", PorT proteins. 68

Figure 3.36. 1D SDS PAGE of membrane protein fraction of the *C. orbicularis* symbiont incubation with saturated (+[O_2]) and low oxygen (-[O_2]) conditions. 69

Figure 3.37. Venn diagram comparison of the identified symbiont proteins of the *C. orbicularis* incubation. 70

Figure 3.38. Comparison of the most abundant metabolic key proteins of *C. orbicularis* incubation symbiont. 71

Figure 3.39. Identified carbon metabolism proteins in the *C. orbicularis* symbiont proteome. 72

Figure 3.40. Comparison (log_2 ratio) of the identified carbon metabolism proteins in *C. orbicularis* symbiont incubation experiment (+[O_2]/-[O_2]). 73

Figure 3.41. Comparison (log_2 ratio) of the identified carbon metabolism proteins in *C. orbicularis* symbiont control and incubation +[O_2] sample. 74

Figure 3.42. Identified CBB proteins in the *C. orbicularis* symbiont proteome. 75

Figure 3.43. Comparison (log_2 ratio) of the identified CBB proteins in *C. orbicularis* symbiont incubation experiment (+[O_2]/-[O_2]). 75

Figure 3.44. Comparison (log_2 ratio) of the identified CBB proteins in *C. orbicularis* symbiont control and incubation +[O_2] sample. 76

Figure 3.45. Overview about the identified sulfur metabolism related proteins in incubated *C. orbicularis* symbionts. 77

Figure 3.46. Comparison (log_2 ratio) of the identified sulfur metabolism proteins in *C. orbicularis* symbiont incubation experiment (+[O_2]/-[O_2]). 78

Figure 3.47. Comparison (log_2 ratio) of the identified sulfur metabolism proteins in *C. orbicularis* symbiont control and incubation +[O_2] sample. 78

Figure 3.48. Identified carbon nitrogen fixation proteins in the *C. orbicularis* symbiont proteome. 79

Figure 3.49. Comparison (log_2 ratio) of the identified nitrogen fixation proteins in *C. orbicularis* symbiont incubation experiment (+[O_2]/-[O_2]). 80

Figure 3.50. Comparison (log_2 ratio) of the identified nitrogen fixation proteins in *C. orbicularis* symbiont control and incubation +[O_2] sample. 80

Figure 3.51. Identified stress related proteins in the *C. orbicularis* symbiont proteome. 81

Figure 3.52. Incubation of isolated bacterial *C. orbicularis* symbiont outside of the host for 24 hours. 82

Figure 3.53. Venn diagram comparison of the soluble proteome of *C. orbicularis* symbiont fraction from freshly collected bivalves and "free living" symbiont incubation (+[O_2]). 82

Figure 3.54. Cytochemical detection of lysosomal enzymes arylsulfatase and acid phosphatase in thin gill sections of *C. orbiculata.* 99

Figure 3.55. Cytochemical detection of lysosomal enzymes arylsulfatase and acid phosphatase in thin gill sections of *C. orbicularis.* 101

Figure 4.1. Ammonium concentration in the seagrass sediment of *T. testudinum.* 103

Figure 4.2. Overview about the "sure" identified proteins in the *Codakia* symbiont proteome. 107

Figure 4.3. Light micrograph of freshly gills from *C. orbicularis* with stained glycogen particles. 110

Figure 4.4. Predicted unconventional CBB in the *Codakia* symbiont. 115

Figure 4.5. Overview about the predicted catalyzed reaction of the PP-PfkA of *R. magnifica*. 115

Figure 4.6. Schematic function of the HppA encoded in the symbiont genome. 115

Figure 4.7. Overview about the assumed sulfur metabolism in the *Codakia* symbiont. 119

Figure 4.8. Schematic map of the *sox* gene clusters in the *Codakia* symbiont genome compared to other known *sox* gene clusters in sulfur oxidizer. 121

Figure 4.9. Schematic map of the of the *dsr* gene cluster in the symbiont genome of *Codakia* and *dsr* gene cluster in related organism. 123

Figure 4.10. The *nif* gene cluster encoded in the *Codakia* genome and identified *nif* gene cluster in other organism. 125

Figure 4.12. Reaction equation of the nitrogen fixation with the nitrogenase. 125

Figure 4.11.The nitrogenase is a two-protein complex. 125

Figure 4.13. Regulation cascade of *nif* genes in *K. pneumoniae.* 129

Figure 4.14. Acetylne reduction assay of *Codakia* gills. 129

Figure 4.15. Overview about the reconstructed predicted physiology of the *Codakia* symbiont. 136

Figure 6.1. 16S rDNA comparison of the symbionts of *C. orbicularis* and *C. orbiculata*. 156

Figure 6.2. SEED Viewer closest neighbor of the *Codakia* symbiont. 173

Abbreviations

AA	amino acid
ABC	ATP-binding cassette
Aqua dest.	destilled water
ATP	adenosine triphosphate
BLAST	Basic Local Alignment Search Tool
CBB	Calvin Benson Bassham cycle
COA	*Codakia orbiculata*
COG	clusters of orthologous
contig	contiguous stretch of DNA
COO	Coomassie Brillant Blue
COS	*Codakia orbicularis*
CP	cytoplasmic
DHAP	dihydroxyacetonphosphate
DTT	dithiothreitol
EC	extracellular
IEF	isoelectric focussing
IM	inner membrane
L	litre
LC	liquid chromatography
MALDI-ToF	Matrix-Assisted Laser Desorption/Ionization Time of Flight
Mb	megabase
MS	mass spectroscopy
MW	molecular weight
NGS	Next Generation Sequencing
NSAF	Normalized Spectral Abundance Factor
OM	outer membrane
ORF	open reading frame
PAGE	polyacrylamide gel electrophoresis
PEP	phosphoenolpyruvat
PGA	phosphoglyceratealdehyd
pI	isoelectric point
P	phosphor
PP	periplasmic
PPi	pyrophosphate
PPP	pentose phosphate pathway
PTS	phosphotransferase system
RuBisCO	ribulose-1,5-bisphosphate carboxylase/oxygenase
SDS	sodium dodecyl sulfate
Sgp	sulfur globule protein
TCA-cycle	tricarboxylic acid cycle
TEM	transmission electron microscopy
TRAP	tripartie ATP-independent periplasmic transport system

1. Introduction

The first description of a microbial symbiosis was done by Antoine van Leeuwenhoek, he observed bacteria society in his own tooth plaque. Further studies demonstrated that the human microbiome is highly complex and has a huge impact on our physiology (Kinross et al., 2011). In our body houses 10x more bacteria than eukaryotic cells. The human microbiome with regard to function and complexity is still a big mystery, which is poorly understood until now (McFall-Ngai, 2008). One good example how bacteria influence our live is the digestion of food in the gut (Million et al., 2013). We are living in the -omics area, these new techniques give us new possibilities to investigate microbial symbiosis. The research in symbiosis field has growth rapidly in the last 20 years (Figure 1.1) The genomics revealed the genetic potential of the microbiome but gives no further information about function and expression, for this goal the proteomics approach can be useful.

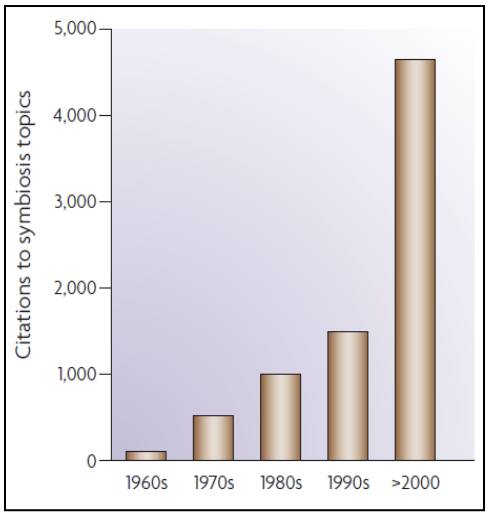

Figure 1.1. Growth of symbiosis field. This figure shows the rapid growth of symbiosis research over the last 40 years. (measured by citations mentioning symbiosis in the title, data derived from Pubmed) (M. McFall-Ngai 2008).

1.1 An overview of symbiosis

The term symbiosis comes from the ancient Greek, "syn" means “with” and “biosis” is the “life”. A definition from deBary 1879 says that a symbiosis describes "a close, long-term or persistent relationship of two or more different species that live together, independent of the fitness effect of each other". In my thesis I will use the term host for the bivalves and symbiont for the bacteria.

Symbiosis occurs since life arose on the earth and exists in all three domains of life (Dimijian, 2000). Symbiosis had and has a big impact on evolution, development and physiology of the species (Margulis et al., 1991; McFall-Ngai, 2001, 2002). One of the most known example is the endosymbiotic theory, which indicates that the origin of mitochondria and plastids and possibly other cell organelles are free-living bacteria.

It is suggested that the establishment of association between two or more organisms is triggered by cross-feeding. Syntrophism or cross-feeding, a special case of symbioses, is an interaction among organisms where one species live on the

products of another, typically through transfer of one or more metabolic intermediates (Schink et al., 2006).

According to the fitness effect on the partners named (i) mutualism association: all partner benefit from the relationship) (ii) parasitism: one partner benefit, the other is may be harmed, (iii) commensalism: one partner benefit, the other is neither harmed nor benefit (figure 1.2). This definition is highly controversial under scientists, some argued for a more narrow definition in which symbiosis describes only mutualism associations.

The distribution of symbionts is depending on where the host harbor their symbionts, ectosymbionts live attached outside of the host, while endosymbionts live inside their host either intra- or extracellular. However, intracellular symbiosis can be further divided in obligate (primary) and facultative (secondary) endosymbionts. Obligate symbionts stay in very close metabolic interaction with their hosts and are essential for the host survival. Both partners share a long evolutionary history together mostly with a genome reduction by the symbiont and the loose of the ability to return to a free living form (van Ham et al., 2003; Moya et al., 2008). Obligate symbionts are dependent by a strict vertical transmission wherever facultative symbionts are normally horizontal or environmental transmitted (Russell et al., 2003). The obligate symbionts harbor in specialized host tissues but also in a facultative symbiont relationship it occurs that the host create specialized tissues for the symbiont.

Most of such symbiotic associations are often established between organism phylogenetically widely separated. Therefore symbiosis between bacteria and eukaryotes allow eukaryotes to resolve metabolic limitations. Bacterial symbionts can provide many different activities to their eukaryotic host, like photosynthesis, chemosynthesis and nitrogen fixation.

One critical step in the association of two or more organisms is the transmission of the symbiont to a new host generation. Therefore three different transmission modes are distinguished, (i) vertically transmission, most without an aposymbiotic phase of the host and often obligate symbionts, (ii) horizontal-environmental, which includes an aposymbiotic phase and a symbiont recognition is necessary. The last transmission mode (iii) is the horizontal-lateral transmission, one host acquires symbionts from another host. However, this transmission modes are not clearly distinguished and many variation of the different modes are described (Bright et al., 2010).

The cultivation of obligate symbionts might by difficult because of the close metabolic interaction and the loose of the ability of a free living form (genome reduction). However, for some endosymbionts a cultivation outside their host could be demonstrated for example for the ship worm symbiont (Betcher, 2011).

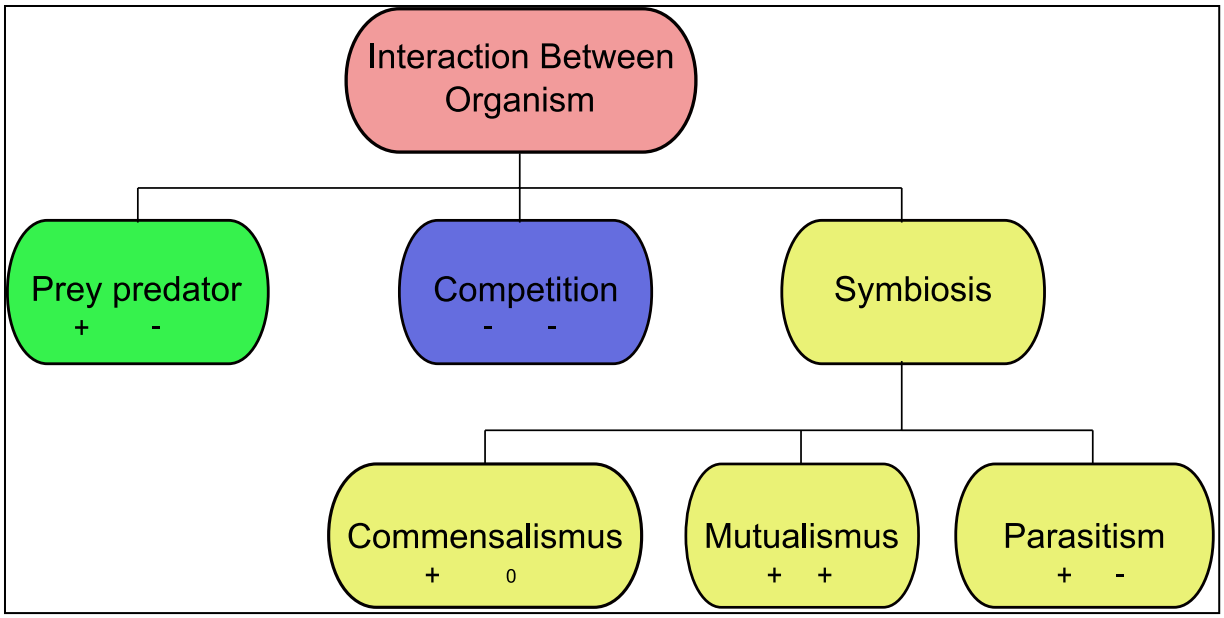

Figure 1.2. Ecological types of interaction between organisms. (+ positive effects of the organism, - negative effects of the organism, 0 no obvious interaction for both partners)

1.1.1 Marine chemosynthesis

Approximately 71% of the earth`s surface is covered by a marine environment and create a wide variety of different habitats. Some of these habitats are particularly characterized by strong physical (pressure, temperature, light) and chemical (e.g. oxygen, sulfide, methane) parameters.
The first description of chemosynthesis was done by Vinogradskii in 1890 and can only be carried out by bacteria and archaea. The energy yield of these reactions is dependent on the differences in the redox potential between electron donor and acceptor and their availability. Since oxygen has a high positive redox potential, it is the best electron acceptor (table 1.1). Reactions with high energy yield are for example sulfide and methane oxidation. The source of methane or sulfide compounds is variable in the environment. At hydrothermal vents, geochemical abitotic processes produce sulfide and methane. In seagrass beds, whale/wood falls and mangroves biotic processes like the anaerobic degradation of organic matter by sulfate reducer are the source. These habitats provide reduced components, for example sulfide, methane and hydrogen to the chemosynthetic bacteria (in this thesis the term sulfide show the total dissolved sulfides: H_2S, HS^- and S_2^-). Autotrophic chemosynthetic bacteria can use sulfide and methane as energy source to fix carbon dioxide and create organic compounds. The chemosynthetic autotroph bacteria produce biomass and constitute an important role in the ecology as primary producer. Therefore in such habitats a variety of different obligate nutritional symbiotic relationships could develop.

Table 1.1. Overview about symbiosis relevant redox reactions and their energetic yield.
(adapted to Wentrup 2012)

Metabolism	e-donor	e-acceptor	$_{\Delta}G_o$[kj/mx][2]	i.a. Symbiotic reference
Aerobic				
methane oxidation	CH_4	O_2	-813	(Childress et al., 1986)
sulfide oxidation	H_2S	O_2	-796	(Berg et al., 1984)
hydrogen oxidation	H_2	O_2	-237	(Petersen et al., 2011)
Anaerobic				
sulfide oxidation/Nitrate reduction	HS^-	NO_3^-	-3722	yes
H_2 oxidation/sulfate reduction	H_2	SO_4^{-2}	-152	yes

$_{\Delta}G_o$[kj/mx][2] - energy available in kJ per reaction

1.2 Chemoautotrophic symbiosis

The association between chemoautotrophic bacteria and marine invertebrates was first discovered in the deep sea at hydrothermal vents in *Riftia pachyptila* (Corliss et al., 1979). *R. pachyptila* is a giant tube worm with up to 2 m in length, without digestive system like gut or mouth and obtain their nutrients from the chemoautotrophic sulfur oxidizing symbionts (Cavanaugh et al., 1981; Felbeck, 1981). *R. pachyptila* is one of the fastest growing marine invertebrates (Lutz et al., 1994) that implies the very effective association between worm and symbiont. To date more than 200 invertebrates species are known to be associated with bacterial symbionts.

Similar symbiotic relationships were described in shallow-water environments, involving bivalves (Felbeck, 1983; Reid, 1990), nematodes (Ott et al., 1982) and annelids (Cavanaugh et al., 1981; Felbeck, 1981). Most of the chemosynthetic symbionts belong to Gammaproteobacteria and are divided in two main functional groups (i) the sulfur oxidizing symbionts and (ii) the methane oxidizing (methanotrophic) symbionts (Cavanaugh et al., 2006). Sulfur oxidizing symbionts were also described in Epsilon- and Alphaproteobacteria (Dubilier et al., 2008; Gruber-Vodicka et al., 2011). All these symbiotic relationships are characterized by carbon transfer to the host, the autotrophic symbionts use the chemosynthetic generated energy for the carbon fixation. Two cases of main carbon transfer are described as (i) "farming" the host digest the symbiont and (ii) "milking" the symbiont provide organic compounds and feed the host. A change between these two strategies is not barred (Streams et al., 1997). Therefore most of the host of chemosynthetic symbionts show a reduced digestive system. The benefit for the symbiont is not so distinct as for the host. One point is the increased availability of substrates, like oxygen and energy sources. Therefore, the tube worm *R. pachyptila* developed a vascular blood system with hemoglobin. The hemoglobin binds reversible oxygen and sulfide (HS^-) for transport to the trophosome and prevents spontaneous oxidation of sulfides (Kraus et al., 1995). The occurrence of hemoglobin appears to be variable in symbiont bearing bivalves. In *Lucina pectinata* hemoglobin is described in the bacteriocytes (Frenkiel et al., 1996). For the bivalve *Codakia* no

hemoglobin is described, probably through the water pumping of surface and pore-water the bivalve provides enough substrates for the symbiont. However the simultaneous providing of sulfides and oxygen is a unique challenge because reduced sulfur compounds show an auto-oxidation in aerobe conditions.

The deep sea mussels *Bathymodiolus* and *Calyptogena*, found at hydrothermal vents, show a reduced digestive system and were one of the first described bivalves with chemoautotrophic symbiont (Cavanaugh, 1983; Fiala-Médioni et al., 1986; Fiala-Médioni et al., 1986). After this discovery chemosynthetic symbiont bearing bivalves were found in several reducing habitats like mud, mangroves, seagrass beds and sunken wood (Dubilier et al., 2008). In chemotrophic symbiont bearing bivalves both paradigms for carbon transfer were found, symbiont digestion "farming" is described for a large tropical bivalve *Lucina pectinata* (Liberge et al., 2001) and carbon translocation "milking" is shown in the gutless bivalve *Solemya reidi* (Fisher and Childress, 1986).

Currently five distinct bivalve families (*Lucinidae*, *Mytilidae*, *Solemyidae*, *Thyasiridae* and *Vesicomyidae*) harbor chemosynthetic symbionts (Roeselers and Newton 2012). Among these bivalve families the symbiotic relationship is diverse. However, in obligate symbiont bearing bivalves the digestive tract and the filtering capacity is reduced (Reid, 1990) and most of them developed a specialized symbiont harboring tissue, like bacteriocytes.

For the bivalve family *Thyasiridae* around 100 species are described and may inhabit coastal and shelf sediments (Southward, 1986; Duperron et al., 2013). Some bivalves harbor chemoautotrophic endosymbionts, one species the symbiont is located extracellular other harbor no symbionts (Batstone et al., 2014; Dufour, 2005). The association is very flexible, the thyasirids can survive time periods without symbionts only by particulate feeding. In nature both occurs, symbiotic and aposymbiotic species (Dufour et al., 2006). It is shown that the thyasirids can digest the symbionts and use them as nutrients ("farming"), no evidence for a direct carbon transfer from symbiont to host ("milking") was found (Dufour et al., 2006).

The bivalve family *Mytilidae* is ubiquitous at the deep sea, including cold seeps, hydrothermal vents or organic falls, in which methane, reduced sulfur or hydrogen is available (Childress et al., 1986; Dubilier et al., 2008). Therefore, the feeding regime of the bivalves is flexible from particle feeding (Page et al., 1991) to receive organic compounds from the symbiont (Dubilier, 1998). It is described that one mussel harbor up to six different symbiont species, corresponding to sulfur-oxidizing or methanotrophic Gammaproteobacteria (Duperron et al., 2013). In members of the *Bathymodiolus* mussels dual symbiosis between sulfur- and methane oxidizing bacteria were observed (Distel et al., 1995; Duperron et al., 2005).

The bivalve family *Solemyidae* occur ubiquitous in the ocean from shallow water to deep sea and live in an obligate symbiosis with sulfur-oxidizing Gammaproteobacteria as endosymbionts (Eisen et al., 1992; Duperron et al., 2013). A vertical transmission of the symbiont is documented (Cary, 1994; Krueger et al., 1996). One very well studied bivalve is *Solemya velum* (Stewart et al., 2006).

The *Lucinidae* family inhabit a broad geographical range in shallow water to at least 2500 m sulfide rich environments, like seagrass and mangroves (Schweimanns et al.,

1985; Taylor et al., 2014). To date it is known that the entire bivalve family *Lucinidae* harbors thiotrophic Gammaproteobacteria specialized host tissue, suggesting an obligatory association (Frenkiel et al., 1995; Taylor et al., 1997). Each of the bivalve species is occupied by a single symbiont species, which is environmental transmitted from a free living stock of bacteria (Duperron et al., 2013; Gros et al., 1996). Based on the 16S rDNA the lucinid bivalves always harbor only one symbiont species, (Durand et al., 1996; Woyke et al., 2006).
The *Vesicomyidae* family comprise at least 100 species and is distributed worldwide at depths from 100 to 9500 m (Goffredi et al., 2003; Krylova et al., 2006, 2010). Currently approximately 29 *Vesicomyidae* species were found in the Atlantic and Mediterranean Sea, only 11 bivalves harbor symbionts (Duperron et al., 2013). Previous studies suggest a vertical transmission of the symbionts (Endow et al., 1990). A new study mentioned the complex story of transmission and claimed it as transovarial transmission, because symbionts were not found in male gonads and gametes (Szafranski et al., 2014). One very well investigated symbiotic interaction is *Candidatus "Ruthia magnifica"* (Newton et al., 2007).

1.3 *Codakia orbicularis* and *C. orbiculata* and its chemoautotrophic symbiont

The lucinid bivalves *C. orbicularis* (Linnaeus, 1758) and *C. orbiculata* (Montagu, 1808) are widespread found in the Caribbean sea and west India. For *C. orbicularis* the maximum reported size is about 75 mm, *C. orbiculata* is much smaller with about 10-20 mm (figure 1.3). The morphology of both bivalves is very similar.

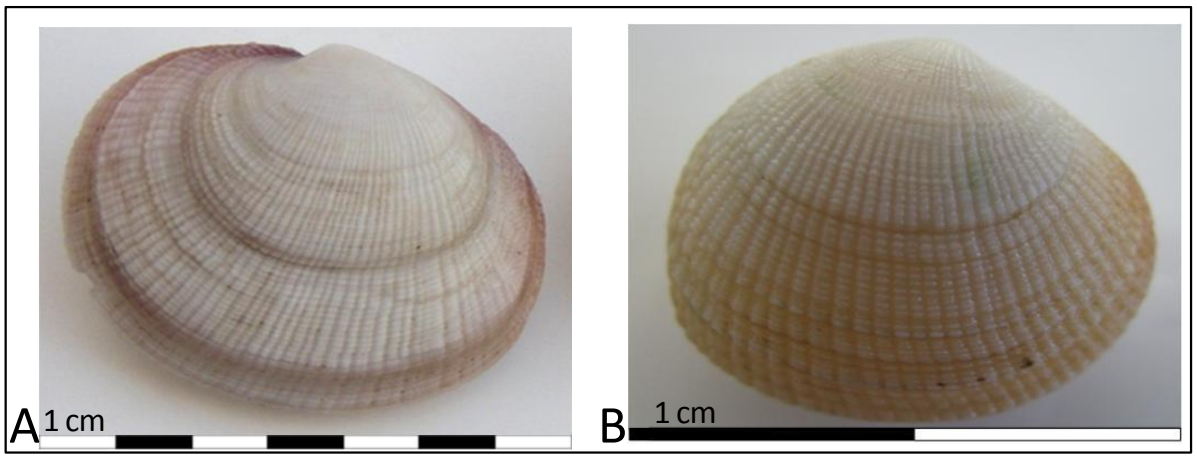

Figure 1.3. (A) - *Codakia orbicularis* and (B) - C. *orbiculata.*
(scale bar 1 cm) (picture: Elisabeth, 2011)

The bivalves inhabit the shallow water seagrass beds of *T. testudinum*. The bivalves burrow in the sediment and live in around 3-7 cm in close association with the seagrass roots (Reynolds et al., 2007). Lucinid bivalves live at the interface of oxic and anoxic sediment zones. They build burrows in sulfide rich anaerobic sediments and are connected to the surface water, which is saturated with oxygen. The bivalves move and burrow with their foot long and branched burrows to connect with sulfur

pockets, sulfide mining (places with high sulfide concentration inside the sediment) (Dufour et al., 2003) (figure 1.5). The bivalves grow about one cm per year (personal communication O. Gros). The big bivalve, *C. orbicularis* has a sweet and good tasting meat, it is a traditional food in the Caribbean (Berg et al., 1984). According to 16S rDNA sequence analysis, both *C. orbicularis* and *C. orbiculata* shelter the same and only one bacterial symbiont species inside the gills (Durand P. et al., 1996). The endosymbiont of *R. pachyptila* is closely related to the symbiont of the *Codakia* bivalves (Figure 1.4)(Durand P. et al., 1996; Imhoff et al., 2003; Brissac et al., 2011).

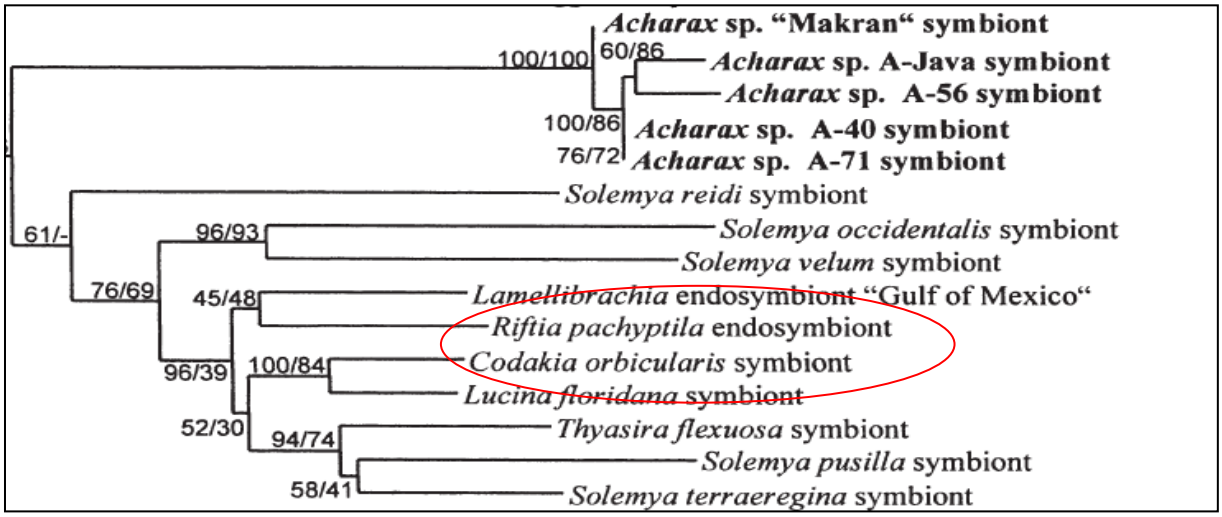

Figure 1.4. Phylogenetic tree of sulfur oxidizing Gammaproteobacteria.
Free living and symbiotic bacteria within the Gammaproteobacteria based on the 16S rDNA sequence similarity. (Imhoff et al., 2003)

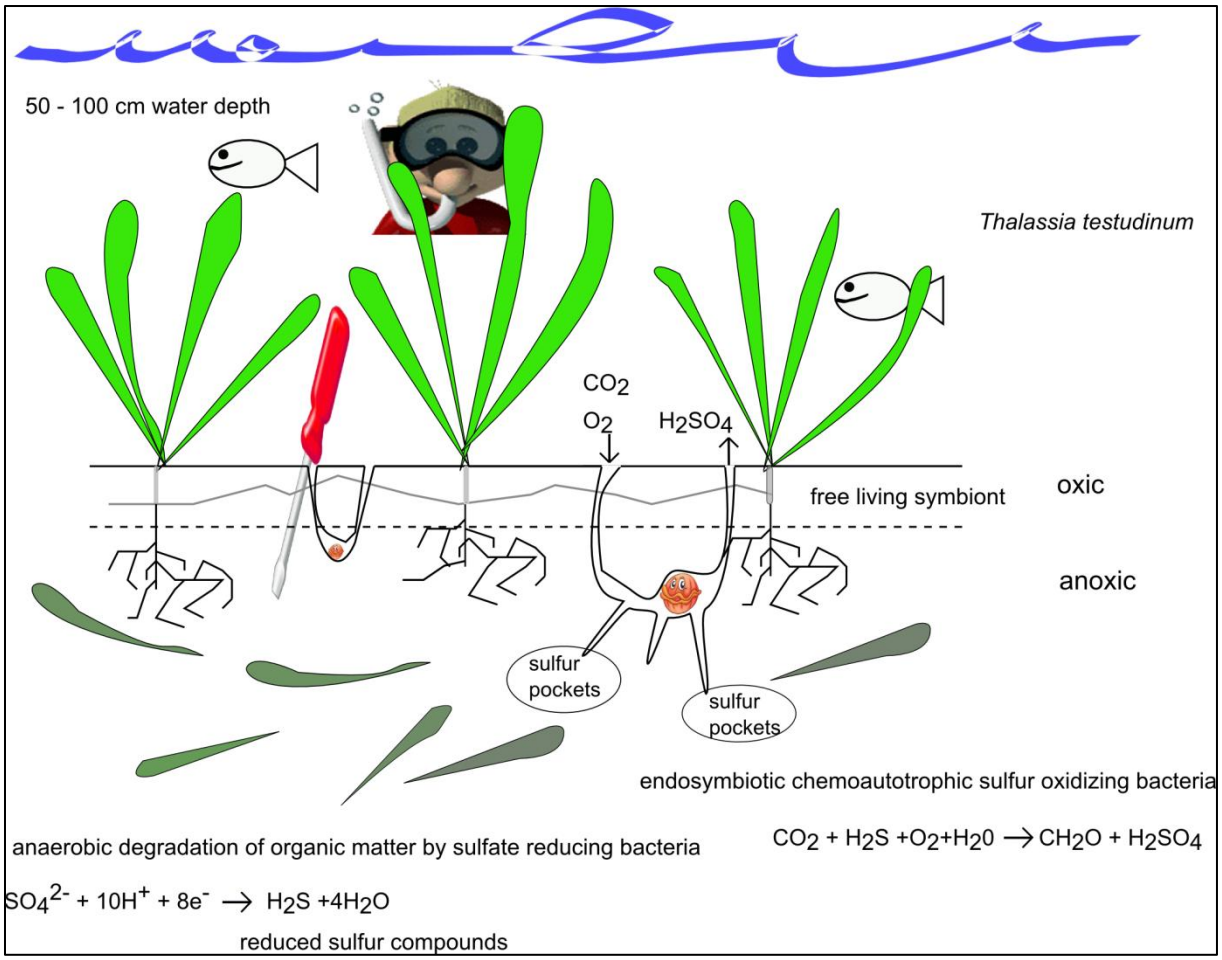

Figure 1.5. Natural habitat of the *Codakia* bivalves.
The seagrass bed is very productive of organic matter, the anaerobic degradation of organic matter by sulfate reducer produce sulfides, this display a biotic source of the sulfides. *Codakia* live burrowed in the seagrass sediment to have access to sulfide and oxygen.

In figure 1.6- A an open *C. orbiculata* is shown. The gills of freshly collected bivalves are in lightly beige color. In starved bivalves the gills are dark brown and thin. The color arises from the content of bacteria and from the amount of elemental sulfur in the periplasmic space of the bacteria (Johnson et al., 2001; Lechaire et al., 2008; Caro et al., 2009). It is visible that the gill tissues capture a big part in the bivalves.
The digestion system and filtering capacity of the bivalve are reduced compared to other bivalves without endosymbionts (Reid, 1990). But the particle feeding from the pore water is still part of the feeding regime of the bivalve (Duplessis et al., 2004a). It is predicted that the nutrient concentration, like oxygen, sulfides and organic particle varies in space and/or in time, the flexible feeding regime can may help to survive these conditions (Dufour et al., 2006; Rossi et al., 2013).
The gills of both bivalves are divided in three zones i) the ciliated zone ii) the intermediary zone, iii) the lateral zone, which takes up most of the volume of the gills (Frenkiel et al., 1995)(figure 1.6- B). The first two zones are free of symbiotic bacteria. The lateral zone is mainly occupied by bacteriocytes (which houses symbionts) (figure 1.7), the upper two thirds and by intercalary cells. The lower third is occupied by granule cells, considered to be important for storage and metabolic conversion of sulfur compounds (Frenkiel et al., 1995). The organization of the gills is

highly variable. Under freshly collected conditions the gills are around 32% occupied by symbionts. It is estimated that the bivalve loose under starvation conditions around one-third of the symbiont population per month. In both bivalves a progressive decrease of the bacteriocytes associated with a reducing number of symbiont populations were observed during starvation, while no release of bacteria were detectable with CARD-Fish (Brissac et al., 2009; Caro et al., 2009; Elisabeth et al., 2012; Gros et al., 2012). That means the symbiotic relationship is a dead end for the symbionts (Brissac et al., 2009). Lysosomes are eukaryotic membrane bound cell organelles with hydrolytic enzymes to chemically degradation of polymers and whole bacteria. In freshly collected bivalves less lysosomes were observed in the gills (Frenkiel et al., 1995). After three months of starvation (no particle food, no reduced sulfur compounds) the amount and size of lysosomes increased (Caro et al., 2009). Still unknown is if it is an active lysosomal host driven lysis, a bacterial autolysis or suicide, because of the sulfide starvation, and the bivalves just use the death bacteria with lysosomal degradation. In both cases the death bacteria are a good nutrient source.

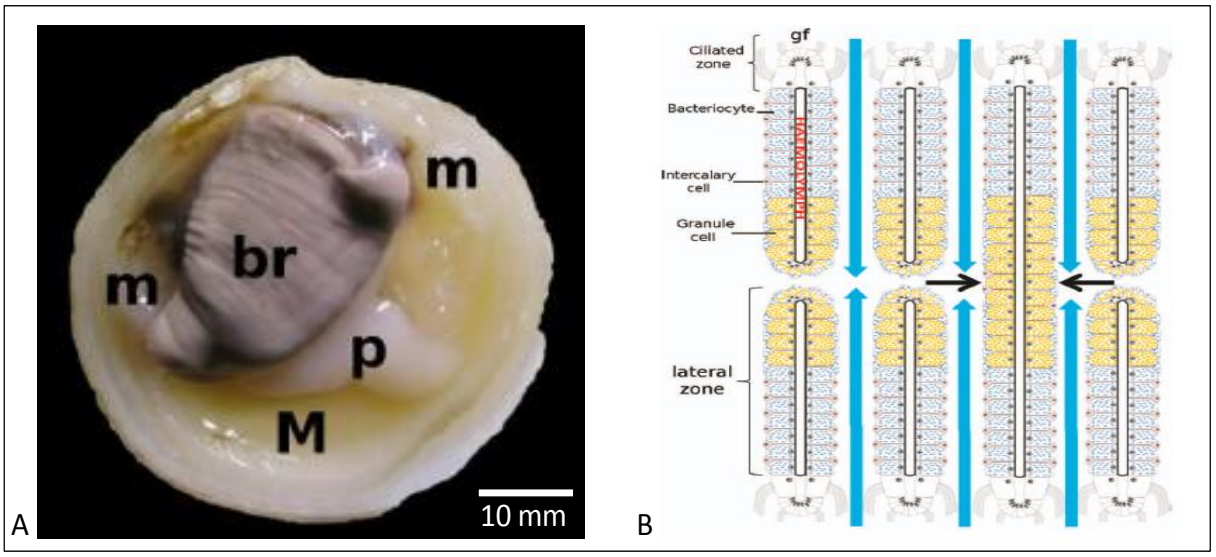

Figure 1.6. Morphology of the *Codakia* bivalves and gill organization.
(A) An open freshly collected *C. orbiculata* bivalve (picture: (Elisabeth, 2011)and (B) schematic overview about the gills organization in freshly collected *C. orbiculata* gills (Gros et al. 2012). The gills tissue harbor the symbiont in specialized cells, called bacteriocytes. The gill is composed of three zones, the ciliated, the lateral and the intermediary zone (not shown). Only the lateral zone harbors symbionts. (br - gills, M - mantle, p - food, m - muscle)

Moreover, a free living stock of symbionts was detected in the sediment of the seagrass bed, an evidence that the symbiosis is facultative for the symbiont (Gros et al., 1996). In the nature an aposymbiotic bivalve was never observed (personal observation). Hence it is assume that for the growth of the bivalve the symbiont is obligate. No endosymbionts were detectable in either larvae or juveniles of *C. orbicularis*, suggesting that the bivalve acquires symbiosis-competent bacteria once their gill tissue is sufficiently developed (Gros et al., 1998). Cultivation of aposymbiotic juveniles of *C. orbicularis* in natural seagrass bed sand demonstrate the transmission from a free living bacterial stock in the sediment (Gros et al., 1996). An environmental/horizontal transmission seems to be common in lucinid bivalves

(Gros, 1999). *C. orbiculata* can acquire bacteria during their whole life by phagocytosis at the apical pole of the bacteriocytes (Gros et al., 2012), until now it could be not shown for *C. orbicularis* but is likely. However, an environmental transmission need a complex molecularly machinery to recognize the symbiont and the selective phagocytosis. Lectins are carbohydrate-binding proteins and often involved in immune defense and predicted to be responsible for the host-symbiont recognition (Sharon et al., 1989; Bright et al., 2010). In the gill of *C. orbicularis* a codakine protein was found with similarities to various animal c-type lectins and the author hypothesized that the codakine protein is involved in the symbiont recognition (Gourdine et al., 2007), further studies indicated a specificity for mannose and fucose monosaccharides and for N-linked biantennary complex type glycans (Gourdine et al. 2008).

Interestingly, it was shown that there are up to seven sub-populations of this symbiont phenotype within a single bivalve. These bacterial subpopulations differ in size 0,5 - 5 µm, number of chromosome copies (1-7) and content of sulfur granules (Caro et al., 2007, 2009). For the reasons behind that polymorphism are only speculative, poly genome copies might indicate fast growing bacteria or highly metabolic active bacteria , because more genes can be translated. But inside the gills dividing bacteria were rarely observed, the authors suggest that the cell division is inhibited within bacteriocytes. (Caro et al., 2007). It is assumed that these different sub-populations also have different metabolic physiologies. It is suggested that the host can control the symbiont populations, the entry and distribution by inhibiting symbiont cell division and harvesting bacterial biomass (Caro et al., 2007).

However, the molecular mechanisms that regulate the symbiont population likely involve a complex symbiont-host signaling cascade and is still unknown (Stewart et al., 2006).

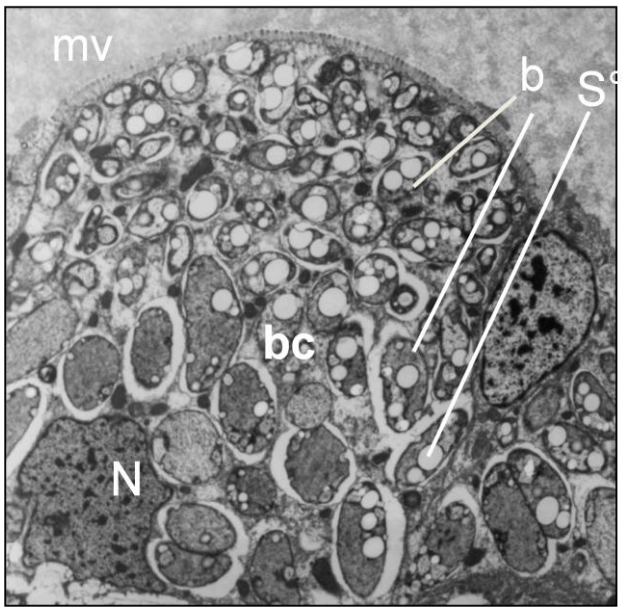

Figure 1.7. Electron microscopy of the gill of *C. orbicularis.*
The picture shows bacteriocytes with membrane enclosed symbionts. The symbionts store elemental sulfur globules in the periplasmic space (visible as white dots) (Lechaire et al., 2008) (bc - bacteriocytes, b - symbiont with elemental sulfur globules, mv - microvilli, N - cell nucleus, S° - elemental sulfur globules) (picture by O. Gros).

Little is known about the physiology of the symbiont and the metabolic interaction with the host, an overview of the known physiology is shown in figure 1.8. The

symbiont uses oxygen as terminal electron acceptor while the evidence of respiratory nitrate reduction is missing (Duplessis et al., 2004a). Inside the gills enzyme activity was found for RuBisCO, APS reductase and ATP sulfurylase (SopT) (Berg et al., 1984; Brissac et al., 2011). The symbiont oxidizes reduced sulfur compounds to generate energy to fix carbon dioxide to create organic matter via the Calvin Benson Bassham Cycle (CBB) (Van Dover, 2000). The seagrass bed is a very productive system of organic compounds (Jørgensen, 1982). The anaerobe degradation of organic matter by inorganic sulfate reducer (biotic process) produces reduced sulfur compounds like hydrogen sulfide or thiosulfate. Reduced sulfur compounds are not stable in the presence of high oxygen concentrations, a process like called auto-oxidation. Therefore the bivalves live at the interface between anoxic and oxic zone in the sediment to provide oxygen as electron acceptor and reduced sulfur compounds as electron donor (Duplessis et al., 2004a). The simultaneous provision of this is an unique challenge because reduced sulfur compounds are oxidize in the presence of oxygen through auto-oxidation (Stewart et al., 2005). It is apparent that a chemoautotrophic lifestyle is very oxygen consuming hence the bivalve filters permanently oxygen saturated surface water to provide the high oxygen concentration to the symbiont. Sulfur exhibits oxidations states -2 to +6. Different enzymatic systems were described in prokaryotes for sulfur oxidation in the past. Thiosulfate is one of the most abundant forms of reduced sulfur compounds in the nature and rather stable. Phyologenetic diverse groups can use thiosulfate, but sometimes close related bacteria can use thiosulfate while the other cannot (Ogawa et al., 2008). Thiosulfate is less toxic than sulfides (Childress et al., 2011). The symbiont store elemental sulfur in the periplasmic space as an energy storage for starvation conditions. If oxygen is not available as an electron acceptor the symbiont can use the elemental sulfur as electron sink and produce sulfides (Duplessis et al., 2004a). In the past nitrate reduction enzyme activity was detected in the bivalve gills (Berg et al., 1984). In later studies no evidence of a dissimilatory nitrate reduction was found predicted for the utilization of nitrate as alternative electron acceptor (Duplessis et al., 2004a). Nevertheless nitrate can be used as nitrogen source.

The symbioses between *Codakia* and the Gammaproteobacteria symbiont is, at least for the symbiont, not obligate. The symbiont is until now not cultivable but it is possible to keep the symbiont alive for around three days in sterile sea water. During this time period, no cell division was observed and after three days, the symbionts die (personal communication of O. Gros). That means such an experiment is not a classical cultivation but rather an incubation of the symbionts for a short time period.

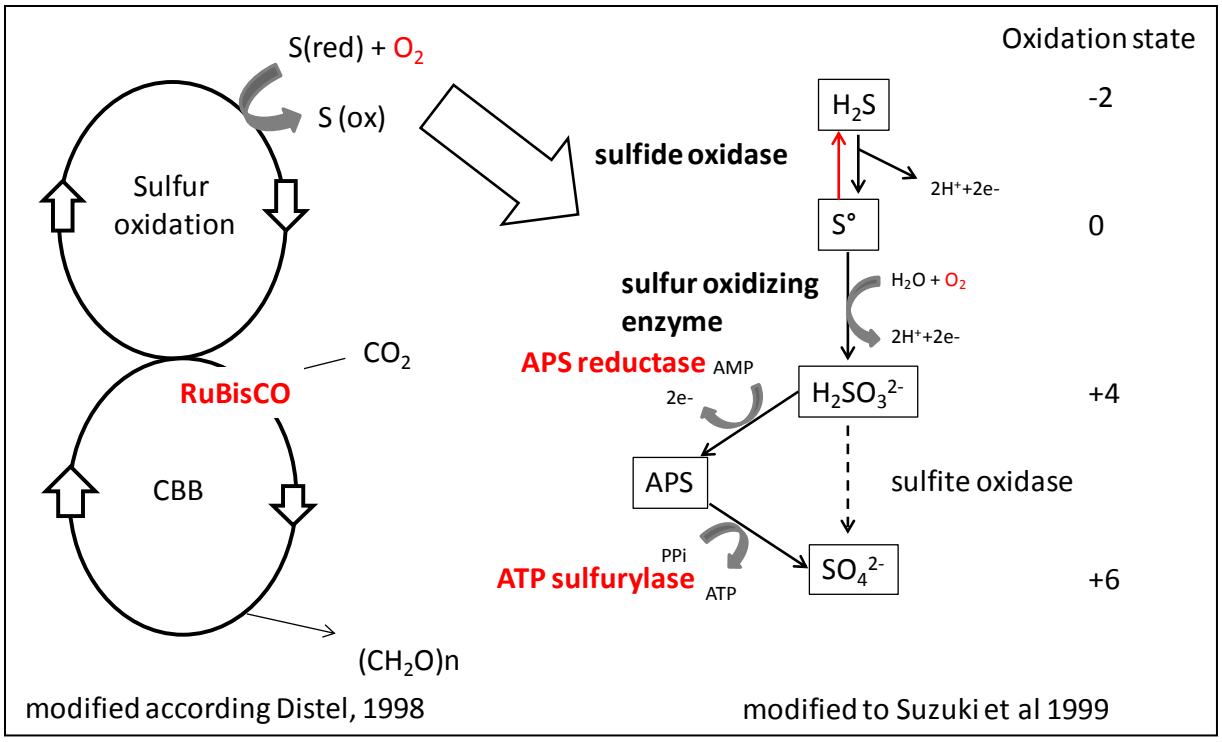

Figure 1.8. Overview about the known physiology of the *Codakia* symbiont.
The symbiont oxidize reduced sulfur compounds to generate the energy for the CO_2 fixation via the CBB to produce organic compounds und provides these compounds to the bivalve. The activity of the red marked enzymes RuBisCO and ATP Sulfurylase (SopT) was detected inside the gills, further the genes for APS reductase (*aprAB*) were found, *aprA* was sequenced from Brissac et al., (2011). The red arrow from elemental sulfur to hydrogen sulfide display an evidence of hydrogen sulfide production under anaerobic conditions (Arndt et al., 2001; Distel, 1998; Suzuki, 1999).

1.4 Ecosystem *Thalassia testudinum* - habitat of *Codakia*

Seagrass is an angiosperm flower plant with leaves, roots and flowers which is permanently under water. Because seagrass needs a lot of light it is only found in shallow water (one to five meter). The photosynthesis rate of the seagrass is very high, thus the surface water is 200% saturated with oxygen (Duplessis et al., 2004a). The light generally has been considered to be the limited growth factor. Mostly the seagrass ecosystem is dominated by one species. Seagrass ecosystems occupy 10% of the world's shallow coastal waters and are an important nursing ground for marine organisms. Because the seagrass bed it is very complex and has a lot of places to hide. The roots of the seagrass stabilize the sediment and help to filter the water. The fauna in the sediment is divided in groups dependent on the living place. The lucinid bivalves dominate the infauna of the seagrass ecosystem, they occupy 97% of the tropical seagrass sites (van der Heide et al., 2012). It is hypothesized that the seagrass provide optimal habitat conditions for symbiont bearing bivalves. This ecosystem is a very productive system and is enriched in organic matter (Duarte et al., 1999). The anaerobic degradation of death organic matter in seagrass bed sediments by sulfate reducers produces reduced sulfur compounds such as hydrogen sulfide (Jørgensen et al., 1974). Reduced sulfur compounds (herein called sulfides) are toxic to most plants and animals as it poisons the cytochrome c oxidase

(Dorman et al., 2002) that means reduced productivity for the seagrass. The bivalve and the sulfur oxidizing endosymbionts oxidize sulfide and thus detoxify the sediment. Van Heide et al. (2012) found a positive correlation between seagrass growth and lucinid bivalves and conclude that the three stage symbiosis between the bivalve the sulfur oxidizing bacteria and the seagrass forms the basis of seagrass ecosystems.
The seagrass bed of *T. testudinum* is typically a nitrate poor environment but the organic and sulfide content is high. Most of the inorganic nitrogen exists as ammonia (Touchette et al., 2000). In a seagrass bed of *Thalassia hemprichii* in Indonesia the ammonium concentration range between 20 µM at the top of the sediment and around 150 µM in six to eight cm sediment depth (Stapel et al., 1996). The ammonium concentration seems to be dependent on the sediment depth. In the oxygenic zone ammonium can oxidize to nitrate, this process consume ammonium. In the anoxic zone an anaerobic dissimilatory nitrate reduction can produce ammonium (An et al., 2002).
The seagrass ecosystem is a very important habitat and achieve different ecological key functions. It is a nursery ground for a different marine organisms, protect coastal sediments, is a feeding place for birds, fix carbon dioxide to carbohydrates thus play a key role in the nutrient cycling. The ecosystem is also very sensitive because it is mostly dominated by only one seagrass species, destructive fishing practices like dragnet, close interaction with bivalves and bacteria. The bivalves filter much water thus they are sensitive for changes of the water quality. Other indirect factors can have a negative effect on the ecosystem like overfishing and the global climate change. In the past decades a loss of seagrass worldwide have been documented. In the future this can be a big threaten for coastal ecosystems (Waycott et al., 2009).

1.5 Objectives - aim of this thesis

The shallow water bivalves *C. orbicularis* and *C. orbiculata* live in the sediment in association with the seagrass *T. testudinum*. The bivalves harbor a chemoautotrophic sulfur oxidizing endosymbiont. However, little is known about the physiology of the symbiont and the symbiotic interaction between host and symbiont.
Hence, the goals of this project were to investigate

1) the physiology of the symbiont extracted from *C. orbiculata* and *C. orbicularis* and potential metabolic differences
2) the interaction between host and symbiont ("communication" and the metabolic exchange of organic compounds)
3) if the bacterial population decreases in case of starvation while no release of bacteria is detectable, is it caused by bacteria auto-lysis or host driven active lysosomal digestion.

2. Material and Methods

2.1 Material

2.1.1 Software and databases

Software and other *in silico* tools used in this study are listed in table 2.1.

Table 2.1. Software and *in silico* tools.

Software/Tool	Company/Reference/Web page
Delta2D Software 4.0 / 4.2	Decodon GmbH (Greifswald) https://www.decodon.com/delta2d.html
Protein localization tools	
Cello	http://cello.life.nctu.edu.tw/
SignalP	http://www.cbs.dtu.dk/services/SignalP/
PsortB	http://db.psort.org/
Phobius TM	http://phobius.sbc.su.se/
LipoP	http://www.cbs.dtu.dk/services/LipoP/
TatP	http://www.cbs.dtu.dk/services/TatP/
TMHMM	http://www.cbs.dtu.dk/services/TMHMM/
Image Quant	Molecular Dynamics
KEGG Pathways	www.kegg.jp./kegg/pathway.html
Scaffold III/IV	www.proteomesoftware.com
NCBI-Pubmed	www.pubmed.com
GPS-Explorer	Software Version 3.5, Applied Biosystems
Jcoast 1.6 / 1.7	MPI, Ribcon GmbH, www.jcoast.net
Venn diagram	http://bioinfogp.cnb.csic.es/tools/venny/index.html
NCBI database	ftp://ftp.ncbi.nih.gov/genomes/
Jspecies	http://www.imedea.uib.es/jspecies/

2.1.2 Chemicals

All chemicals used in this study were purchased from Carl Roth, Sigma Aldrich, Applichem, Serva Electrophoresis, GE Healthcare Life Science, Pro-Lab, VWR, Promega.

2.2 Methods

2.2.1 Sampling of bivalves

Adult individuals of *Codakia orbicularis* (Linné, 1792) and *C. orbiculata* (Montagu, 1802) were manually collected in *Thalassia testudinum* seagrass beds on “îlet cochon” (16°12’53’’N; 61°32’05’’W) in Guadeloupe (French West-Indies, Caribbean). The bivalves live in the sediment 1 to 10 cm depth. The average bivalve shell length of the collected *C. orbicularis* were around 5 cm and of *C. orbiculata* around 1 cm. Specimens of *C. orbicularis* were also bought on the street market.

2.2.2 Bivalve incubation - Induction of symbiont's loss

Around 20 individuals of *C. orbicularis* and up to 50 bivalves of *C. orbiculata* were kept in 50 L plastic tanks filled with 0.22 µm filtered seawater at 26°C up to six months. The bivalves were visibly checked of liveliness twice every day and death bivalves were removed. For the cytochemistry experiments bivalves were cultivated for six months (T_1, T_2, T_3, T_4, T_5). For proteomic investigations the bivalves were incubated up to four weeks under the same conditions. The water was oxygenated with an aquarium pump. No food (organic particles) and no reduced sulfur compounds were added into the water throughout the starvation experiments. The water in the tanks was changed every seven days to remove toxic waste products and avoid algae growth in the sea water tanks.
In order to reach low oxygen concentrations in the starvation experiments, the sea water tanks were bubbled with argon gas. The oxygen concentration was measured every day using a microsensor multimeter (Unisense). The incubation tank was bubbled again with argon gas if the oxygen concentration increases more than 0.5 mg/l. In the tank with low oxygen concentration 5 mM potassium nitrate was added as a potential alternative electron acceptor. During this incubation at different time points bivalves were killed and dissected to extract the bacteria. These sample were stored at -20°C for further investigations.

2.2.3 Cytochemistry - arylsulfatase and acid phosphatase detection

The gills of the bivalves were dissected and fixed (2% glutaraldehyde, 1% paraformaldehyde in 0.1 M cacodylate adjusted to 900 mOsm,) for 2 h at 4°C. The fixed gills were washed overnight at 4°C in 0.2 M sodium cacodylate buffer with an adjusted osmolarity at 1800 mOsm with NaCl and Ca_2Cl. The gills were sliced in sodium cacodylate buffer into 100 µm sections using a Vibratome OCT Slicer at 4°C. Only one enzyme activity can be detected per gill section, therefore for each enzyme reaction three replicate sections were prepared. Arylsulfatase activity was detected according to a method from Lewis et al. (1992). Briefly, the 100 µm sections were transferred directly to the enzyme reaction solution (60 mM sodium acetate, 15 mM acetic acid, 8 mg/mL sulfate p-nitrocatechol, 60 mM barium chloride) and incubated for 45 min at 37°C. After the incubation, the samples were washed in cacodylate buffer with several changes of overnight.
Acid phosphatase was detected by a method from Barka and Anderson (Barka et al., 1962). The sections were pre-incubated for 10 to 20 min at 37°C in 50 mM acetate buffer with 15 mM acetic acid. The enzymatic reaction took place in the pre-incubation buffer containing an additional 0.1 M Sodium ß-glycerophosphate substrate, saturated with lead-(II)-nitrate. The sections were incubated for 45 min at 37°C. After incubation the sections were washed in acetate buffer over a period of 60 min with several changes of the washing buffer.
For both acid phosphatase acid and arylsulfatase assays, a precipitate containing the heavy metal was formed as a result of enzyme activity during the incubation stage.

For both enzymes a negative control was prepared therefore the 100 µm-gill sections were incubated in the particular enzyme reaction solution without the enzyme substrate but with the heavy metal.
Following incubations, 100 µm sections were processed for embedding. Samples were dehydrated through a graded aceton series and embedded in Epon-araldite resin according to Glauert (1991). Ultra-thin (100 nm) sections were made using an ultracut E microtome and two grids containing 4-5 sections were observed with a Tecnai G2 TEM at 200 kv. The barium or lead precipitates were detected using a 30 mm^2 EDX detector. The EDXS analysis corresponds to an analytical method used for elemental characterization. Micrographs presented in this study are representative of the three replicate bivalves examined at each sampling time during starvation.

2.2.4 Purification of the bacterial fraction - density centrifugation

The symbionts were purified using a Percoll density centrifugation (Distel et al., 1988) (figure 2.1). Only gills were used which looks thick and lightly beige. The thick and lightly beige gills were dissected and homogenized in cold sterile sea water using a handheld Dounce homogenizer. These homogenates were centrifuged (30 g, 1 min, 4°C). The supernatants were centrifuged again (400 g, 2 min, 4°C). The bacteria containing pellet, was resuspended in sterile sea water and gently filled on a cold 3 mL Percoll gradient (60% Percoll (Sigma Aldrich) and 40% IBS buffer (0.49 M NaCl, 30 mM $MgSO_4$, 11 mM $CaCl_2$, 3 mM KCl and 50 mM Imidazol (pH 7.5)). This gradient with the bacterial solution loaded on the top was centrifuged (4000 g, 10 min, 4°C). The bacteria, which contains elemental sulfur as granules in the periplasmic space (Lechaire et al., 2008), accumulate as a white pellet at the bottom of the gradient, while the lighter host tissue stays at the top of the gradient. These pellets were washed twice in sterile sea water. The bacterial pellets were stored at -20°C. This procedure was used for the genomic and proteomic samples. For *C. orbicularis* symbiont, a single bivalve proteomic analysis was possible, the gills of *C. orbiculata* are too small for a single bivalve analysis so four to six individuals were merged. This has to keep in mind by the quantification of the proteins of the *C. orbiculata* symbiont.

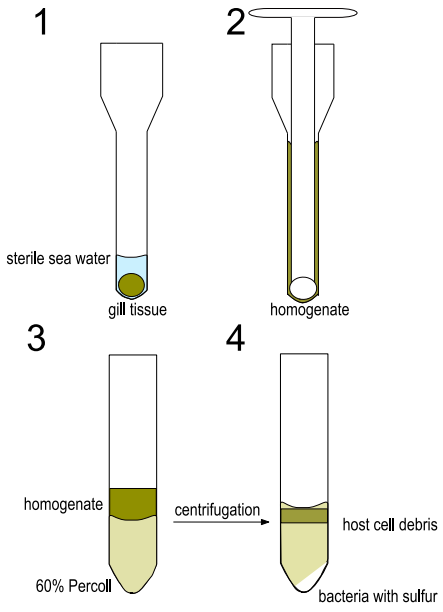

Figure 2.1. Endosymbiont purification with Percoll density centrifugation.
The first step is the dissection of the gills. The gill tissue was homogenized with sterile sea water in a Dounce homogenzier (1-2). The big host pieces were removed during a first gently centrifugation step. After a second faster centrifugation step the bacteria are in the pellet and could be resuspended in sterile sea water. The received solution was carefully filled on a 60% Percoll gradient (3). After a fast centrifugation step the bacteria form a white pellet at the bottom because of the elemental sulfur. The host tissue without sulfur is lighter and in the upper part of the gradient (4).

2.2.5 Sequencing of the *Codakia* symbionts

For *C. orbicularis,* the two gills of a single bivalve were used for the enrichment of the bacterial symbionts for sequencing. For *C. orbiculata* one bivalve was not enough because the gills are tiny. In this case it was necessary to merge 4 to 8 individuals. Only bivalves with a lightly beige gill tissue were used for the extraction of the bacterial symbionts. The symbionts were purified as described above and were send on dry ice to the Göttingen sequencing labor.
The DNA extraction of the purified symbionts was done in the Göttingen Genomics Laboratory in using a MasterPure DNA purification Kit (Epicentre, Madison, WI, USA), following the manufacturer's manual. The 454 libraries were produced with a GS FLX General Shotgun DNA Library Preparation Method Manual (Roche, Mannheim, Germany). The reads were assembled *de novo* using Ray assembly (Boisvert et al., 2010). The genomes of the symbionts of *C. orbiculata* and *C. orbicularis* were sequenced by the Göttingen Genomics Laboratory using the 454-GS-FLX Titanium and Illumina next generation sequencing techniques. The contigs were provided as .fasta files, this files were used for further analyze.

2.2.6 Gene annotation

The data were up-loaded by IMG (Integrated microbial genomes and metagenomes) and at once evaluated by IMG (Markowitz et al., 2014). Additional the raw data were analyzed in the workgroup of T. Schweder with the help of S. Heiden. For the gene annotation only contigs over 500 bp were used. The sequence analysis and annotation were performed with the Jcoast software package (http://www.jcoast.net/). Jcoast is a comparative analysis and search tool (Richter et al., 2008), that enables an automatic ORF finding. After the automatic annotation, a human manual gene annotation revision was done.

2.2.7 Database creation for proteomics

To remove homolog genes in the both *Codakia* symbiont genomes, a bidirectional blast from NCBI was used with general criteria: protein identity over 95% and in both blast direction the same protein hit. The resulting database was used for the proteome analyses of this study.

2.2.8 Metabolic Pathways

The annotated genes were classified according to their metabolic functions in the KEGG Pathway maps with adaptations to marine Gram-negative bacteria. The metabolic classes are provided in the supplementary material.

2.2.9 Proteomics

2.2.9.1 Extraction of soluble proteins

The white pellet of purified symbionts was washed in lysis buffer (10 mM tris, 1 mM EDTA, Roche complete protease inhibitor) and resuspended in 300 µL lysis buffer. The cell disruption was done with ultra sonication, three times for 20 sec with permanent cooling. The cell debris were separated by a centrifugation step (12.000 g, 10 min, 4°C). The proteins in the supernatant were precipitated with acetone to remove salt from the sample. Therefore two volumes of -20°C cold acetone were added to the protein solution and incubated overnight at -20°C. The precipitated proteins were spin down in a centrifugation step (9000 g, 30 min, 4°C). The pellet was washed twice in 99% ethanol and one time in 70% ethanol and afterward solubilised in 8 M urea and 2 M thiourea. Therefore the samples were gently shaken at RT for at least one hour. The protein concentration was measured with Roti® NanoQuant, an adapted procedure to the Bradford method (Bradford, 1976).

2.2.9.2 Enrichment of the membrane proteins

Membrane proteins contain usually hydrophobic domains. This means that they are not soluble in a normal water based buffer solution. Because of that, a special separation protocol was used for the accumulation of membrane proteins (Eymann et al., 2004; Wolff et al., 2008). The cell disruption was done as described above with ultra sonication. To remove the cell debris, the solution was centrifuged (8000 g, 10 min, 4°C). The supernatant was filled in a ultracentrifugation tube (Herolab, Wiesloch Germany). The ultracentrifugation was done with a fixed rotor (Beckman L7-65, rotor Ti 70) (100.000 g, 60 min, 4°C). The pellets were homogenized and incubated under permanently shaking in a high salt buffer (20 mM tris HCl pH 7.5, 1 M NaCl) for 30 min at 4°C and again ultra-centrifuged (100.000 g, 60 min, 4°C). The pellets were homogenized in alkaline buffer (0.1 M Na_2CO_3, pH 11, 0.1 M NaCl) and afterwards again incubated for 30 min at 4°C in alkaline buffer and centrifuged

as describe above. These steps were repeated with TEAB (triethylammonium bicarbonate) buffer (50 mM TEAB, pH 7.8). The pellets were homogenized in 100-200 µL TEAB buffer. The protein concentrations were measured as described above with Roti®NanoQuant. Therefore a calibration curve was made with BSA and TEAB as buffer.

2.2.9.3 1D gel SDS PAGE

The proteins were separated according to their size in a 12% acrylamide/ bisacrylamide gel (Laemmli, 1970). The 1D gels have two different parts, a stacking gel with low, 5% of acrylamide/bisacrylamide concentrations above and a separating gel with higher concentration, 12% of acrylamide/bisacrylamide concentration (table 2.2) The gels were run on a BioRad Mini-PROTEAN® system. To destroy secondary and tertiary structures of the proteins, the samples were heated for five min to 90°C in sample buffer. The running buffer was prepared freshly. The gel electrophoreses was done at 120 to 150 V and maximal 200 mA for around 1-1.5 hours in BioRad Mini-PROTEAN® Tetra Cell (table 2.3)To protect the separated proteins in the gels for mass spectroscopy, the gels were fixed for one hour in a fixing solution (50% ethanol (v/v), 12% acetic acid (v/v)) and afterwards stained with gently shaking at RT in a Coomassie Brilliant Blue staining solution (table 2.3) over night. To remove the background coloration the gels were decolorized with *A. dest.* and gently shook with several changes of the water until the background of staining on the gel disappeared. After this step, the gels were scanned with a light scanner (Light scanner X-finity Ultra, Quatographic Technologie GmbH) and stored at 4°C.

Table 2.2. Composition of 1D SDS gel.

Chemical	Separating gel 12%	Stacking gel 4%
Aqua dest.	3.35 mL	1.4 mL
1.5 M tris-HCl (pH 8.8)	2.6 mL	----------
1 M tris-HCl (pH 6.8)	----------	2.6 mL
10% SDS	75 µL	25 µL
Rotiphorese® Gel30 *	2.5 mL	0.42 mL
10% APS	75 µL	25 µL
TEMED	3 µL	2.5 µL

(*acrylamide-mix 37.5:1)

Table 2.3. Chemicals for the 1D gel SDS PAGE.

	Chemical	Concentration
4 x SDS loading-buffer	tris	200 mM
	DTT	400 mM
	SDS	8%
	bromphenolblue	0.4%
	glycerol	40%
Running-buffer	tris	25 mM
	glycin	190 mM
	SDS	3.5 mM
Coomassie Brillant blue solution 1l	*aqua dest.*	900 mL
	o-phosphoric acid	118 mL
	ammonium sulfate	100 g
	Coomassie G-250	1.2 g
	methanol	200 mL

2.2.9.4 In-gel protein digestion and ESI MS/MS analyses

The 1D gels were cut into ten equal bands and were completely decolorized with 200 mM ammonium carbonate ($(NH_4)CO_3$) and 30% acetonitrile in Eppendorf tubes at 37°C for one hour and gently shaking with two to three times changing of the buffer. The gel pieces were dried in a Speed Vac to remove the acetonitrile. The gel pieces were covered with a 1:20 trypsin (Promega) solution (diluted in a trypsin activation buffer provided by Promega). The dry gel pieces absorb the trypsin solution. The gel pieces with the trypsin solution were incubated over night at 37°C without shaking. To eluate the peptides from the gel, the gel pieces were covered with *A. dest*. and submitted to sonication in an ultra sonication bath for 15 min. The tubes were placed in a swimmer so that they stay at the surface of the bath. The supernatant with the peptides inside was transferred to LC-MS glass vials. In order to clean the samples a Zip Tipping (MERCK Millipore^{TM}, Billerica) step was done according to the manufacturer's instruction manual. Only samples, which were treated with the same procedure, were compared after the MS analysis.

2.2.9.5 Protein identification and data analysis with Scaffold

The program Scaffold (4.2 Proteome Software Inc., Portland OR, USA) was used to validate MS/MS based peptide identifications. Only peptides that exceed the threshold criteria were used for sure identifications, that means the deltaCN score was greater than 0.1 and XCorr scores greater than 2.2, 3.3 and 3.8 for doubly, triply and quadruply charged peptides respectively. Only proteins were accepted with at least two identified unique peptides. The relative quantification of the proteins in the 1D gel analysis was done by using the normalized spectral abundance factors (NSAF) (Zhang et al., 2010).

2.2.9.6 2D SDS PAGE

The proteins were separated in two dimensions. The first one is the isoelectric point and the second one is the molecular weight. After the precipitation the protein pellet was resuspended in a buffer (8 M urea, 2 M thiourea) and afterwards the protein concentrations was determined with the Bradford assay. For separation by 2D PAGE, 100 µg of protein were loaded onto immobilized pH gradient (IPG) stripes in the pH range 3-10 or 4-7. The stripes were loaded with the particular amount of protein extract supplemented with a detergence solution (8 M urea, 2 M thiourea, 1% Chaps, 20 mM DTT and 0.52% Pharmalyte pH 3-10) and incubated at room temperature for at least 16 hours. The isoelectric focusing (IEF) was done with a Ettan IPGphor III (GE Healthcare) (1^{st}. phase: 500 V and 500 Vh, 2^{nd} Phase: 1000 V, 800 Vh, 3^{rd} phase: 8000 V, 13500 Vh, 4^{th} phase: 8000 V 21200 Vh) for around 8 h at 20°C. The IPG stripes were equilibrated for 20 min in equilibration solution A (50 mM tris-HCl (pH 6.8)), 6 M urea, 30% glycerol, 4% SDS, 3.5 mg/mL DTT) and for 20 min in equilibration solution B (50 mM tris-HCl (pH 6.8), 6 M urea, 30% glycerol, 4% SDS, 45 mg/mL iodacetamide and a small amount of bromphenolblue). For the second dimension the separation after the molecular weight the IPG stripes were put on the SDS gels (0.38 M tris-HCl, pH 8.8, 0.01% SDS, 0.34% bisacrylamide, 12.2% acrylamide, 0.025% ammonium persulfate and 0.05% TEMED). The gels with the stripes were run in a Protein Plus Dodeca Cell (BioRad) at 15°C for approximately 16 h. The gels with the separated proteins were fixed in 50% ethanol and 7% acetic acid for at least one hour at RT. The fixed gels were washed in *A. dest*. for 10 min with two changes of water. The proteins were visible with a fluorescence dye, Sypro Ruby and Flamingo, after an incubation in the dye solution at RT over night. The stained gels were scanned with a fluorescence scanner (Typhoon 9400, GE Healthcare) in an intensity range from 84000 to 104000 for the quantification of the protein spots with Delta2D. The Delta2D software allowed the combination of three gel replicates, the spot detection on these gels and the determination of the spot intensities.

2.2.9.7 2D gel protein spot digestion and identification with MALDI-ToF

The protein spots were picked automatically with the Ettan Spot Handling workstation (Amersham Biosience Uppsala Sweden). The protein spots were transferred in a 96 well plate. The 96 well plate was automatically digested with the Ettan spot digestion station (Ettan™ Spot Handling Workstation, GE Healthcare). The samples were measured with a MALDI-ToF 4800 (Matrix-Assisted Laser Desorption/Ionization Time of Flight Mass Spectrometry) (Applied Biosystems, Darmstadt, Germany) from Dr. D. Albrecht. For analysis only proteins were used with a probability-based mowse score of $p < 0.05$, at least 50% protein score and exhibiting a minimum peptide count of two unique peptides of the protein.

2.2.9.8 Gel free proteomics - Synapt G2

The bacterial pellet was washed in 50 mM TEAB buffer (pH 7.4) and resuspended in 300 µl TEAB buffer for the cell disruption with ultra sonication (3x20 sec). Afterwards the protein concentration was determined with the Roti® NanoQuant assay. For the gel free digestion 250 µg of proteins diluted in 80 µl TEAB buffer were added in a Eppendorf tube and additional 20 µL RapiGest^{TM} (Waters) SF Surfactant solution were added to the protein buffer solution. The RapiGest^{TM} SF Surfactant is used to enhance enzymatic digestion of proteins. For reducing of the protein solution, 1 µL of 500 mM TCEP (tris(2-carboxyethyl)phosphine) was added and incubated for 45 min at 60°C with gently shaking. To prevent the disulfide bonds from refolding 2 µL of 500 mM iodacetamide were added to modify the reactive cysteine-SH groups and incubated for 15 min at RT in the dark. Before the digestion the trypsin was activated in a trypsin activation buffer for 15 min at 37°C. For digestion, 6.25 µL of 10 times activated trypsin were added to the protein solution and incubated for 5 h at 37°C with 900 min^{-1} shaking. To remove the RapiGest^{TM} 1 µL triflouracetic acid was added and incubated for 45 min at 37°C to decrease the pH. Under low pH conditions the RapiGest^{TM} is precipitated. The precipitated RapiGest^{TM} was spin down at least three times at a table top centrifuge (max rpm, 10 min, 4°C). The samples were desalted with a Stage Tip Purification from Proxeon according to manufacturer's instructions. The following washing and equilibration steps were done in a standard table top centrifuge at 7400 g at RT. To improve the desalting of the Stage Tips of Proxeon (Thermo scientific), C^{18} material diluted in methanol were obtained around 5 mm on the top of the original column in the Proxeon Stage Tips. The column was rinsed three times with 100 µL buffer A (0.1% acetic *A. dest*) and afterwards washed two times with 100 µL buffer B (0.1% acetic acid in acetonitrile). It is important that all these steps were done without drying the column. The columns were loaded with the peptide solution and washed twice with buffer A. The peptides were eluted with 30 µL buffer B by using the elute syringe. To remove the acetonitrile 20 µL A. dest. water were added and incubated in the Speed Vac until the volume decrease to 10 µL. The exact peptide solution volume was determined refuel to the same volume. Tryptic digested hemoglobin, used as an internal standard in a determined concentration was added before measuring. The vials were stored at -80°C until measuring with the Synapt G2 mass spectroscopy. The sample measurements were done by the workgroup of Prof. D. Becher, mainly by S. Meyer. Three technical replicates from three biological replicates respectively were measured. The three technical replicates were merged and proteins which were identified twice were supposed to be "sure" identified, same principle for the biological replicates (figure 2.2).

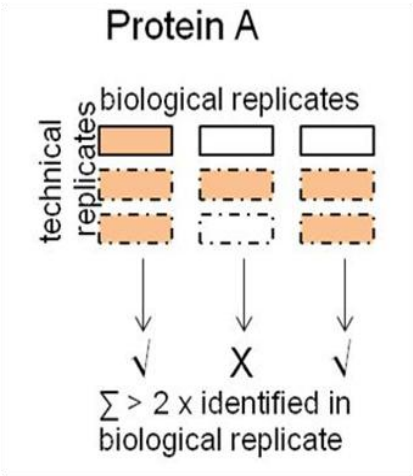

Figure 2.2. Schema of the gel free analysis with the SynaptG2.
Three biological replicates were measured and from each of them three technical replicates were prepared. To assume a protein as sure identified, each technical replicate has to be identified twice as well as each biological replicate (Protein A: random protein, beige color - Protein A identified, X - not sure identified, √ - sure identified).

2.2.9.9 Gene expression comparison

To compare the protein quantification under different conditions, we used a $\log_2$ ratio to measure the expression of the proteins (Bergemann et al., 2011). That means a protein identified with the equal amount would have a $\log_2$ ratio of zero, a twofold change of the protein amount would have a $\log_2$ ratio of one. We used a $\log_2$ ratio of 0.5 as a threshold for significant different protein expression. In tables a color coding was used to visualize different gene expression pattern (figure 2.3).

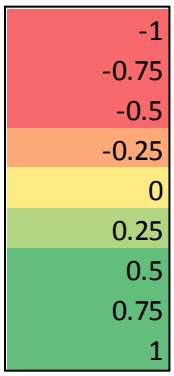

Figure 2.3. Color coding used for the visualization of the $\log_2$ ratio.

3. Results

3.1 Genomics

The sequencing of both symbiont fractions was done by the Göttingen Genomics Laboratory group headed by Prof. R. Daniel in Göttingen. They used different sequencing techniques, Illumina and 454 sequencing to get the best result. The raw sequencing data were aligned in Greifswald with the help of S. Heiden, afterwards the predicted genes were automatically and manual annotated, by Jcoast and human power.

3.1.1 General genome characteristics

The shotgun genome projects of both symbiont fractions from *C. orbicularis* and *C. orbiculata* were up-loaded to IMG (Integrated microbial genomes and metagenomes; (http://img.jgi.doe.gov)) by A. Thürmer from Göttingen. This group provided us a file with the assembled large contigs. The assembly of the DNA reads was done by Ray assembly (Boisvert et al., 2010). The DNA reads were assembled to 179 contigs in *C. orbicularis* and 903 contigs in *C. orbiculata.* Thus, the draft genome of both symbionts is fragmented in 179 and 904 pieces. The calculated general genome statistics from IMG are shown in table 3.1. The predicted genome size of the symbionts of *C. orbicularis* and *C. orbiculata* are 4.8 Mb and 5.3 Mb respectively.

Table 3.1. General predicted genome statistics of the *Codakia* symbiont genomes.
(calculated from IMG http://img.jgi.doe.gov/)

	***Codakia orbiculata* symbiont**	***Codakia orbicularis* symbiont**
contigs (>500 bp)	903	179
predicted size (Mb)	5.3	4.8
GC content (%)	51.84	52.12
protein coding genes	5378 (100%)	4534 (100%)
protein coding genes with predicted function	3856 (71%)	3450 (76%)
RNA genes	77 (1.43%)	81 (1.79%)

3.1.2 Genome comparison

The phylogeny is the study of the evolutionary relationships among groups of organisms usually based on highly conserved DNA sequences. In bacteria in general the 16S rDNA is used to reconstruct the phylogenetic relationships.
Previously the 16S rDNA sequence of the symbionts of *C. orbiculata* and *C. orbicularis* was sequenced by P. Durand and O. Gros (1996). Based on the 16S rDNA comparison it was suggested that both symbionts belong to the same species. In the course of the genome sequencing, we re-sequenced the 16S rDNA (figure

6.1). The comparison showed that there is a high similarity (99% identity) between the respective gene sequences of both symbionts. Due to decreased sequencing quality at the beginning and the end of the reaction there are not determined bases (N) visible. The visible check of the alignment showed that around 900 nucleotides fits perfect to each other, only two not determined nucleotides are the differences. To verify the results of the 16S rDNA analyses, we furthermore compared selected key genes of both symbionts (figure 3.2).

Table 3.2. Key genes comparison of *C. orbicularis* and *C. orbiculata* symbiont.
(NCBI nucleotide blast http://blast.ncbi.nlm.nih.gov)

Category	Gene	Function	Identity
	16S	small subunit ribosomal RNA	99%
sulfur metabolism			
	aprA	adenylylsulfate reductase, alpha subunit	100%
	aprB	adenylysulfate reductase, beta subunit	100%
	dsrA	dissimilatory sulfite reductase, alpha subunit	100%
	dsrB	dissimilatory sulfite reductase, beta subunit	100%
	sopT	ATP sulfurylase	100%
Calvin Benson Bassham cycle (CBB)			
	cbbM	ribulose-1,5-bisphosphate carboxylase/oxygenase, large subunit	100%
	gapA	glyceraldehyde-3-phosphate dehydrogenase	100%
	pgk	phosphoglycerate kinase	99%
	prkB	phosphoribulokinase	100%
nitrogen fixation			
	nifA	nitrogenase molybdenum iron protein subunit alpha	100%
	nifB	nitrogenase molybdenum iron protein subunit beta	99%
	nifH	nitrogenase reductase	100%
nitrogen metabolism			
	narH	respiratory nitrate reductase, alpha chain	100%
	narJ	respiratory nitrate reductase, beta chain	99%
	narI	respiratory nitrate reductase, delta chain	99%
	nosZ	nitrous-oxide reductase	100%
oxidative stress response			
	sodA	superoxide dismutase [Fe]	99%

To further analyze the relatedness of the *Codakia* symbionts Jspecies, a tool which measure the probability whether two genomes belong to the same species, was used. Jspecies compares genomes with different alignments methods. The software calculates the average nucleotide average (ANI) based on BLAST (ANIb) and MUMmer (ANIm). The average nucleotide identity (ANI) is described in Goris et al. (2007). The tetra nucleotide frequencies calculation is described in Teeling et al. (2004). The MUMmer software calculate the ANIm values, as described in Kurtz et al. (2004). In table 3.3 it is shown, that all blast parameter achieve the threshold that both symbionts belong to the same species. The sequencing quality differs, which

leads to sequencing mistakes. This might lead to the small distinctions in the key genes of the symbiont genomes.

Table 3.3. Jspecies genome comparison.
The green marked numbers achieve the threshold that both organism belong to the same species (identity in %) (http://www.imedea.uib.es/jspecies/). (COA - *C. orbiculata*, COS - *C. orbicularis*)(ANI - average nucleotide identity)

		COA	COS
Tetra (tetranucleotide signature):	COA	--------	**99.915**
	COS	**99.915**	--------
ANIb (based on blast) average nucleotide identity (ANI):	COA	--------	**98.53**
	COS	**98.86**	--------
ANIm (based on MUMmer):	COA	--------	**98.75**
	COS	**98.74**	--------

3.1.3 Genome annotation

The contigs of the draft genomes were scanned by Jcoast for open reading frames (ORF). Jcoast predicted for the *C. orbiculata* symbiont 5129 and for *C. orbicularis* 4158 protein coding genes. The gene annotation tool Jcoast tries to give these potential genes a predicted function. This comparison is based on an aminoacid (AA) blast, because of the degeneracy code of the DNA, with proteins of known function. But all these predicted function results have to be handled with care, because it is just a probability for a predicted function and sometimes several different hits share the same similarity. The Jcoast tool enables an automatic annotation by using several databases (NCBI, SwissProt, Pfam, Kegg, Interpro). This first annotation is susceptible for mistakes, a manual correction was done to create the database for the proteomic investigations. Because the data (16S rDNA, key genes, Jspecies) indicate that the symbiont of the bivalves belong to the same species, only the genes of the *C. orbicularis* symbiont were further analyzed. The genome of the *C. orbicularis* symbiont was choosen because of the higher sequencing quality.

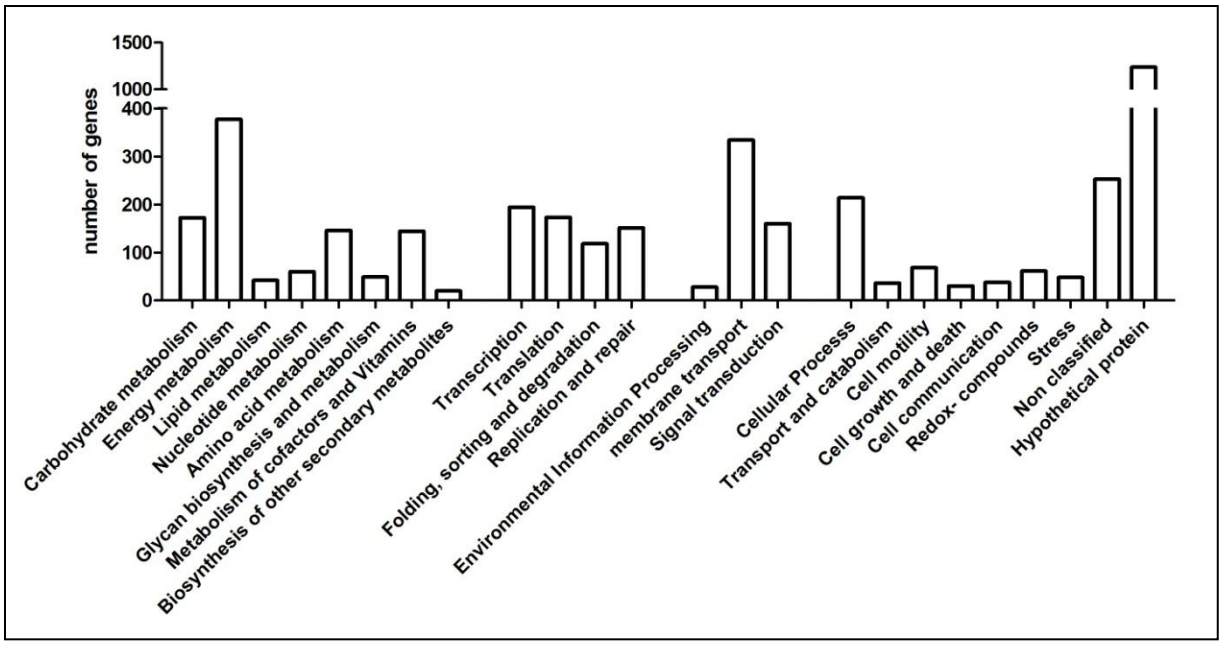

Figure 3.1. Overview of the gene function of *C. orbicularis* symbiont.
The data show the protein classes after manual gene annotation correction (categories adapted from Kegg metabolic pathways).

In the figure 3.1 an overview is shown about the functions of genes in the *C. orbicularis* symbiont genome. One main part of genes belongs to the energy metabolism which includes sulfur oxidation metabolism related genes, nitrogen metabolism related genes and genes related to the oxidative phosphorylation. Also a large number of genes for the membrane transport are encoded in the genome, this includes several types of multi-component transporter, like ABC, TRAP and PTS. Noticeable is that many genes cannot be affiliated to a function, but they contain conserved gene regions with unknown functions. Interestingly the amount of genes related to carbohydrate metabolism and transport seems quite high. This can indicate for an autotrophic lifestyle of the bacteria.

3.1.4 Database creation

The protein database is the backbone for proteome analyses. For our studies we used the draft genomes of both lucinid symbionts as database. The genomes are in principle very similar but could contain distinctions in the gaps of the genome or sequencing mistakes. As mentioned above, both genomes were automatically annotated by Jcoast and manual corrected. To get a non-redundant database that comprises the complete genome we used a bidirectional blast from NCBI to remove homolog genes with the general criteria of a protein identity over 95% and in both blast directions the same gene hit (figure 3.2). Due to the fact that the quality of the *C. orbicularis* genome with only 179 contigs was better than the *C. orbiculata* genome with more than 903 smaller contigs we used the *C. orbicularis* genome as a "master-genome". That means all proteins of *C. orbicularis* are in the database plus the proteins of *C. orbiculata* which did not achieve the general protein criteria. In summary 5709 proteins were summarized in the fasta sequence database, which

was used for all proteomic investigations of this study (electronic supplemetary material (fasta file: DB_COS_COA; Exel file: Codakia genome_classification_ Proteomics).

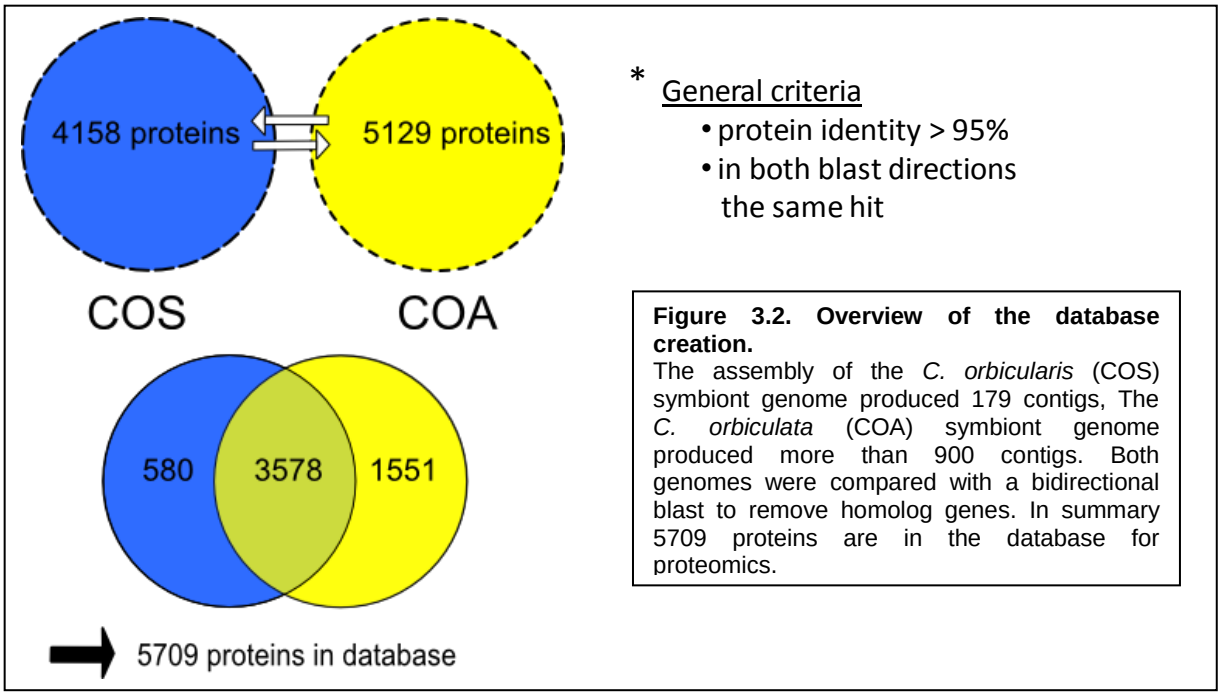

Figure 3.2. Overview of the database creation.
The assembly of the *C. orbicularis* (COS) symbiont genome produced 179 contigs, The *C. orbiculata* (COA) symbiont genome produced more than 900 contigs. Both genomes were compared with a bidirectional blast to remove homolog genes. In summary 5709 proteins are in the database for proteomics.

3.2 Proteomics

The proteome of the cell determines the physiology under certain environmental conditions to a certain time point. We reconstructed the metabolic pathways by the identification of the enzymes of the respective pathway. By the determination of the abundance of the proteins we were able to estimate the importance of distinct pathways under certain conditions for the organism.
For *C. orbicularis* we used for all proteomic approaches one bivalve as one biological replicate. For *C. orbiculata* it was necessary to merge the gills of four to eight specimens. This is relevant for the quantification and comparison of the results of both bivalves. For a broad proteomics investigation different approaches were used. First, the soluble intracellular protein fractions were isolated and separated by two dimensional polyacrylamide gel electrophoresis (2D SDS PAGE). Detectable protein spots were measured by MALDI-ToF analyses. Second, 1D SDS PAGE was used for the pre-separation of the soluble intracellular or enriched membrane protein fractions. The peptide mixtures of the in-gel digested protein fractions were measured with an Orbitrap mass spectrometer. Third, a gel free approach by using the SYNAPT G2 technique, a high definition mass spectrometer, was used for absolute protein quantification.

3.2.1 2D - PAGE

The 2D PAGE was independently described by Klose and O'Farrell (Klose, 1975; O'Farrell, 1975). The proteins were separated in the first dimension according to their isoelectric point and in the second dimension according to the molecular weight. This approach helps to reduce the complexity of protein mixtures. It is also a good method to visualize and estimate the abundance of distinct proteins.

3.2.1.1 Theoretical 2D gel protein distribution

The created database with 5709 proteins was used to separate the proteins *in silico* in two dimensions, the isoelectric point and the molecular weight. For this approach the internet tool JVirGel was used to separate the proteins. (http://www.jvirgel.de/index.html). The theoretical 2D gel picture is shown in figure 3.3.

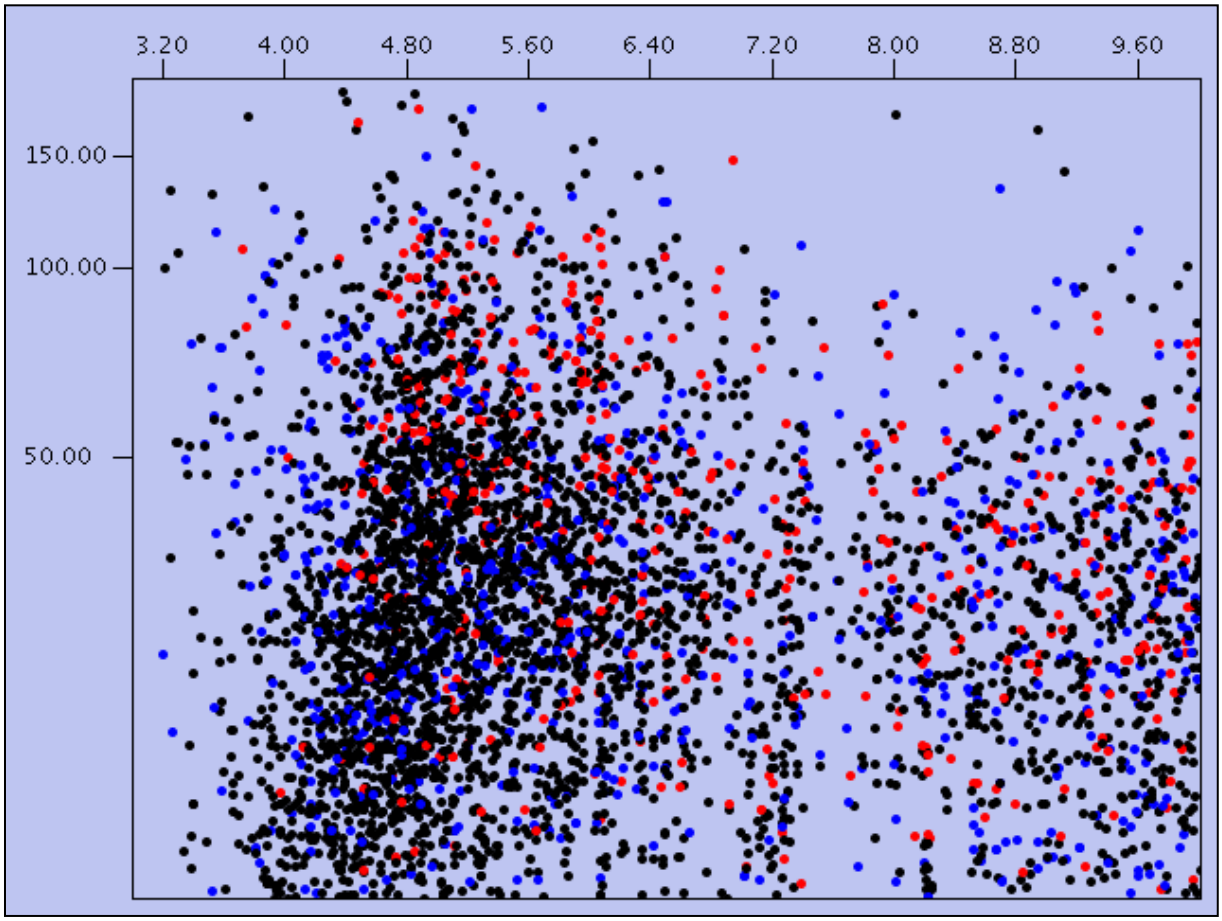

Figure 3.3. Theoretical distribution of all encoded *Codakia* symbionts proteins.
All proteins with a molecular mass between 200 to 15 kDa and an isoelectric point between pH 3-10 were virtually unraveled. (red spots - indicate potential membrane proteins, blue spots - indicate that the protein containing a signal peptide, black spots - normal intracellular proteins)

The biggest amount of proteins is between pH 4-7 and in a molecular weight range of 25 to 100 kDa. The theoretical 2D distribution has shown that many proteins are in

the main pI region between pH 4-7, therefore we used a pH range between 4-7 for our 2D-gel based protein separations.

3.2.1.2 2D gels of freshly *C. orbicularis* and *C. orbiculata* symbionts

The 2D gels were stained either with the fluorescence dye SYPRO® Ruby or Flamingo. The change of the dye was done due to economic reasons, flamingo is less expensive. However, both dyes are of similar sensitivity, but a direct quantitative comparison between the different stained gels is not possible. The spot size is dependent on the protein amount of the particular protein. The quantification of the 2D gels are based on a visible measurement of the black protein spot area stained by the respective fluorescence dye. All labeled proteins spots were identified three times in biological replicates with an MALDI ToF mass spectrometer. The mass spectrometer measurement was done by D. Albrecht. The protein search after the MS measurement was done using the first sequencing results based database of the *C. orbiculata* symbiont. The quality of the used first preliminary database was not satisfying and thus the proteomics results were improvable. Further analysis are necessary to improve these 2D gel data. Therefore, only selected identified key proteins were used for the investigation of the physiology of both *Codakia* symbionts.

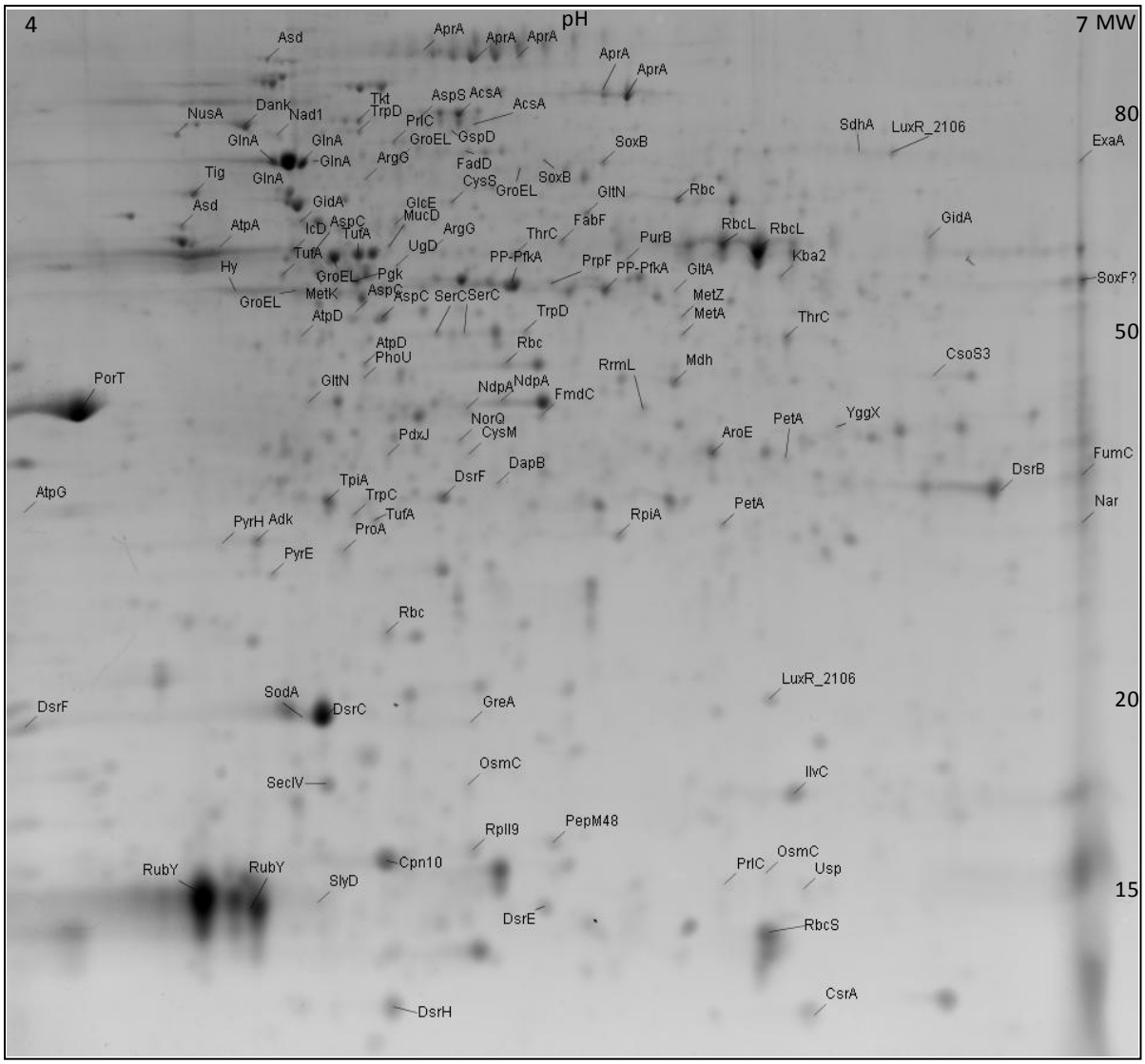

Figure 3.4. Master 2D gel of the intracellular soluble proteome of the *C. orbicularis* symbiont.
The proteins were separated in the pH range 4-7 and stained with SYPRORuby. "Sure" identified proteins identified with MALDI-ToF are indicated with protein label.

Currently, 87 proteins could be identified in the 2D gel of the soluble *C. orbicularis* symbiont protein fraction (figure 3.4). 19 identified proteins are part of the carbohydrate metabolism, the both RuBisCO (RbcL1/2, RbcS1) forms were abundant identified. 14 proteins belong to the energy metabolism, mainly sulfur oxidation proteins. The APS reductase (AprAB), key enzyme of the sulfur oxidation is abundant identified. A "porine of gram negative bacteria" (PorT) is high abundant identified in the 2D gels. A small protein was abundant identified as rubrerythrin (RubY). An overview about the identified proteins is provided in the supplementary material (file: 2 D gel results).

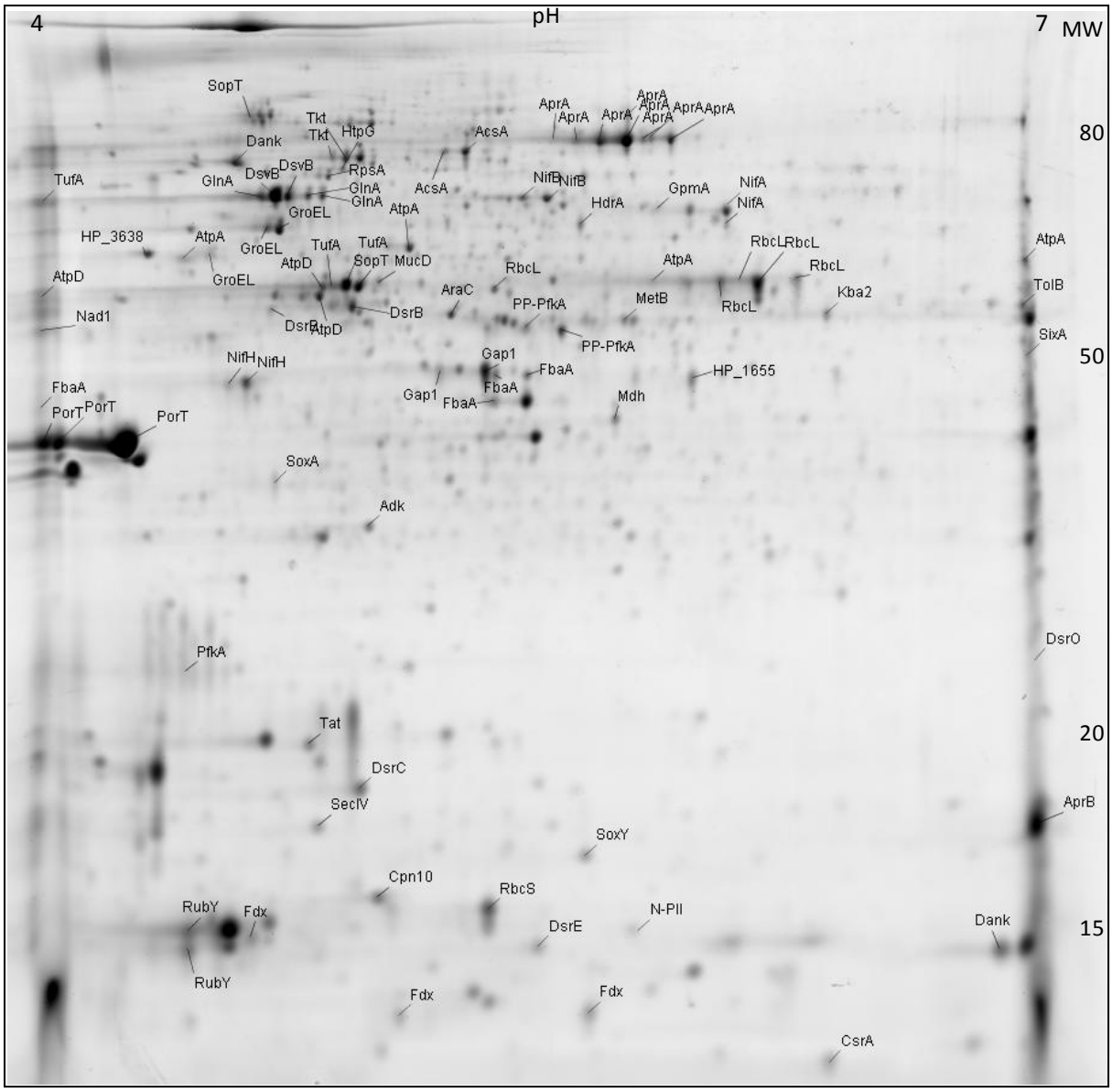

Figure 3.5. Master 2D gel of the intracellular soluble proteome of the *C. orbiculata* symbiont.
The proteins were separated in the pH range 4-7 and stained with Flamingo™. All indicated proteins are "sure" identified with MALDI-ToF.

In the symbiont of *C. orbiculata* 50 proteins could be identified in the 2D gel of the soluble protein fraction (figure 3.5). 11 proteins related to carbohydrate metabolism were identified, including the PP dep. 6-phosphofructokinase proteins (PP-PfkA), transketolase and ATP-dep. phosphofructokinase (PfkA). RuBisCO (RbcL1/S1) form I was identified in a huge amount. The key enzymes, nitrogenase and some regulatory proteins for the gaseous nitrogen fixation were identified (NifK/D/H, N-PII) in the 2D gels. The glutamine synthetase (GlnA), a key enzyme of the amino acid metabolism, is abundantly expressed in both bivalve symbionts.
The major protein expression is in both *Codakia* symbionts the same. The biggest spots are similar in both symbionts RubY, PorT, RuBisCO, AprAB and GlnA. This indicate the same metabolic physiology of the symbionts in both bivalves.
A table of all identified proteins is provided in the supplementary material (file: 2 D gel results).

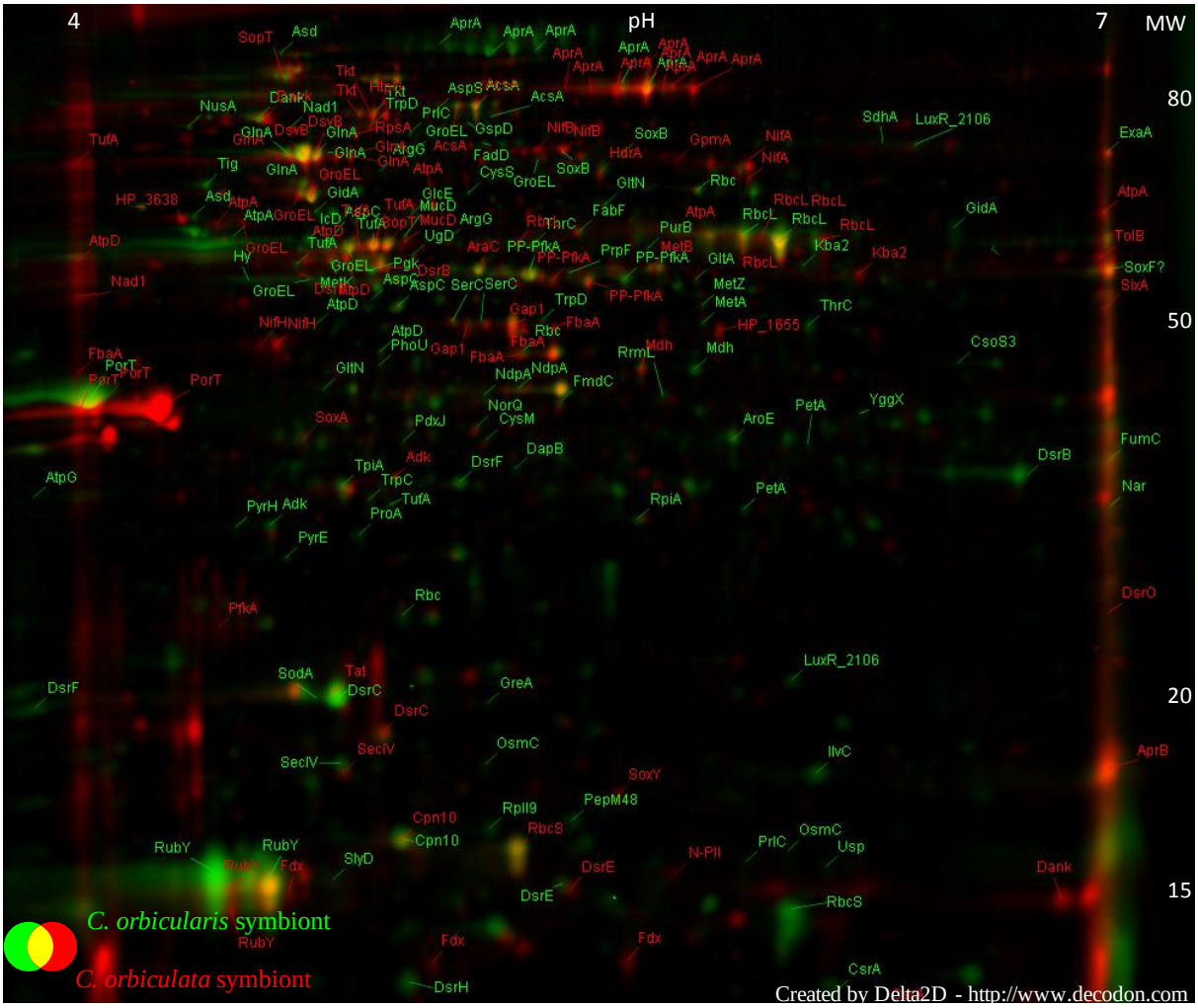

Figure 3.6. False color 2D gel of the symbiont of freshly collected *C. orbicularis* and *C. orbiculata*.
The proteins were separated in the pH range 4-7. All identified proteins are indicated, green marked proteins were identified in *C. orbicularis* symbiont and red marked proteins were identified in *C. orbicularis* symbiont.

In figure 3.6 a false color comparison is shown between symbionts of both bivalves. Two gels of different conditions were warped on each other with the help of the software Delta2D (http://www.decodon.com) and spots of the same protein were located at the same position on the gel image. With the false color the expression pattern of the two different conditions can be compared. Yellow spots indicate proteins with similar spot intensity while red or green spots indicate different spot intensity or spots which were only detected in one sample. The biological aberrance was quite high between both bivalve symbionts. For both symbionts more than six 2D gels were produced and the protein spots were identified. On the false color picture it is visible that the abundant spots are in both bivalve symbionts the same, but also a lot of green spots are visible and some red spots. These spots are exclusively identified in the symbiont of one bivalve. In this respect it is important to emphasize, that for the small bivalve around five gills were merged. Furthermore the bivalve is a mobile system and moves inside the sediment, the geochemical parameter can fluctuate very much.

Noteworthy is the remarkably high abundant rubrerythrin protein (RubY ID 1096) spot which was identified at a low pH and low molecular weight. This protein was identified

in two different spots which are located closely together. Further analysis are necessary to explain the two protein spots of one protein and to confirm the identification results.

3.2.1.3 *C. orbicularis* incubation and 2D gels of the symbiont

After sampling of the bivalves in the seagrass bed, bivalves were immediately transferred into the sea water tank. In the first week, the mortality of the bivalves was around 10%, after this time the mortality decreased to approximately 5% per week. My personal observation was that under low oxygen conditions the bivalve shells were more and often open than under normal conditions. That implies the bivalve filter more water under low oxygen conditions.
The bacterial extraction is based on the high density of the bacteria because of the elemental sulfur inside the periplasmic space. The content of the elemental sulfur in the bacteria decreases in case of starvation, as well as the symbiont population themselves (Lechaire et al., 2008). The bacterial extraction with freshly collected bivalves gills worked mostly well, sometimes it occurred that during the homogenizing step a hard white foam was formed which contained most of the bacteria. It was impossible to destroy this foam and extract the bacteria via the density centrifugation. We tried an extraction with several filter steps of different sizes but without any success in case of the hard white foam. The filter extraction method worked not as clean as the density extraction but is totally upgradeable in the future. One advantage of this method is that the sub-population of the symbionts without sulfur particle could be enriched too.
The 2D gels are a good method to control the quality of the symbiont extraction because of the visualization of the proteome. After two weeks the proteome in the 2D gels changed completely (figure 3.7). The amount of detectable protein spots decreased while the amount of two big spots with low molecular weight increased. The gels indicate that the enrichment of the bacteria via the density centrifugation worked well in the first two weeks after collecting. After this time the enrichment quality decreased. For the proteomics a clean enrichment of the symbionts is essential, because of that one week of starvation time was choose to investigate the physiology under different conditions.

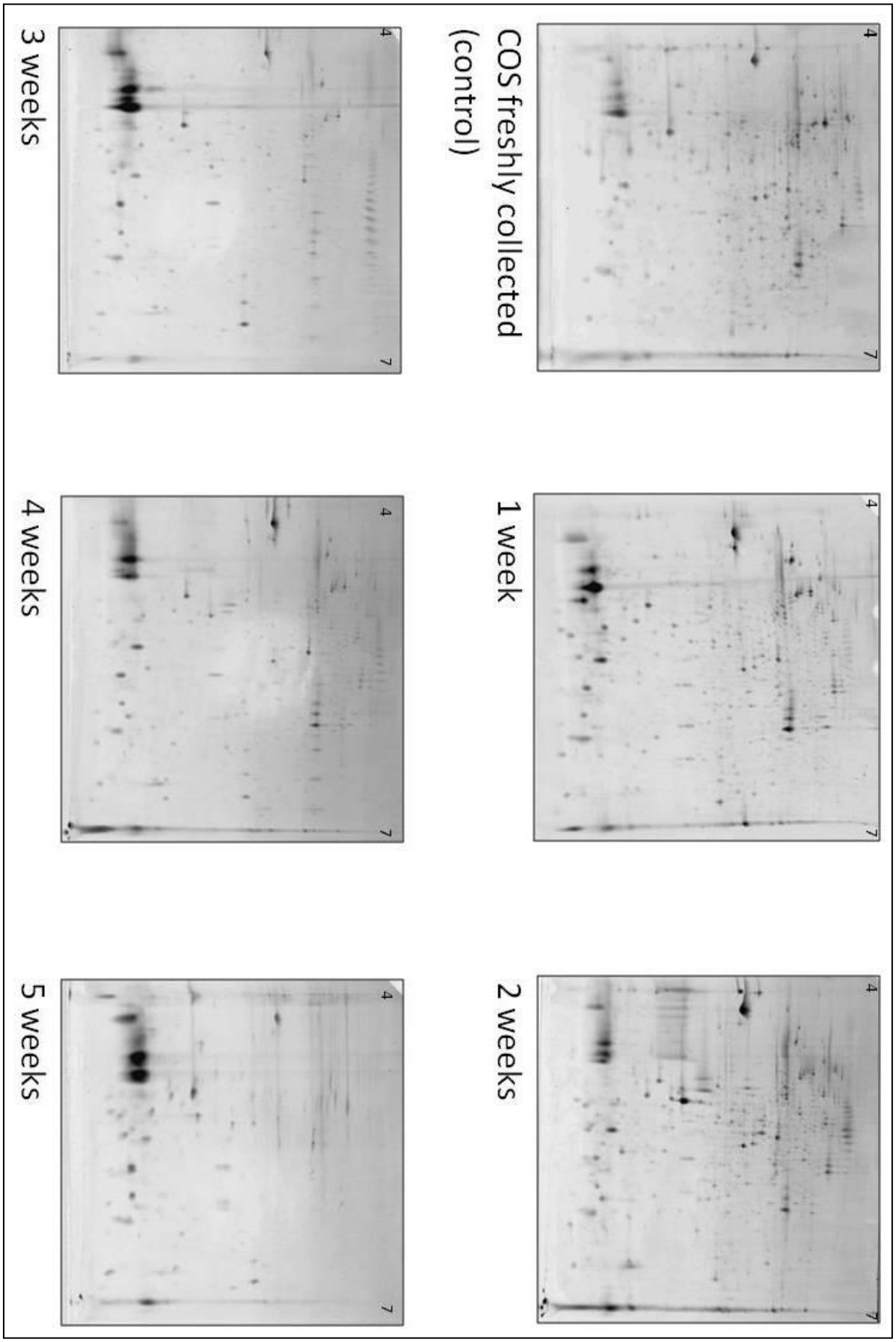

Figure 3.7. Comparison of the 2D gels of the soluble proteins of *C. orbicularis* symbiont incubation samples.
The bivalves were incubated in sterile sea water tanks with saturated oxygen conditions. The soluble proteins of the symbiont were separated in the pH range 4-7 and stained with the fluorescence dye Flamingo™. (COS - *C. orbicularis*)

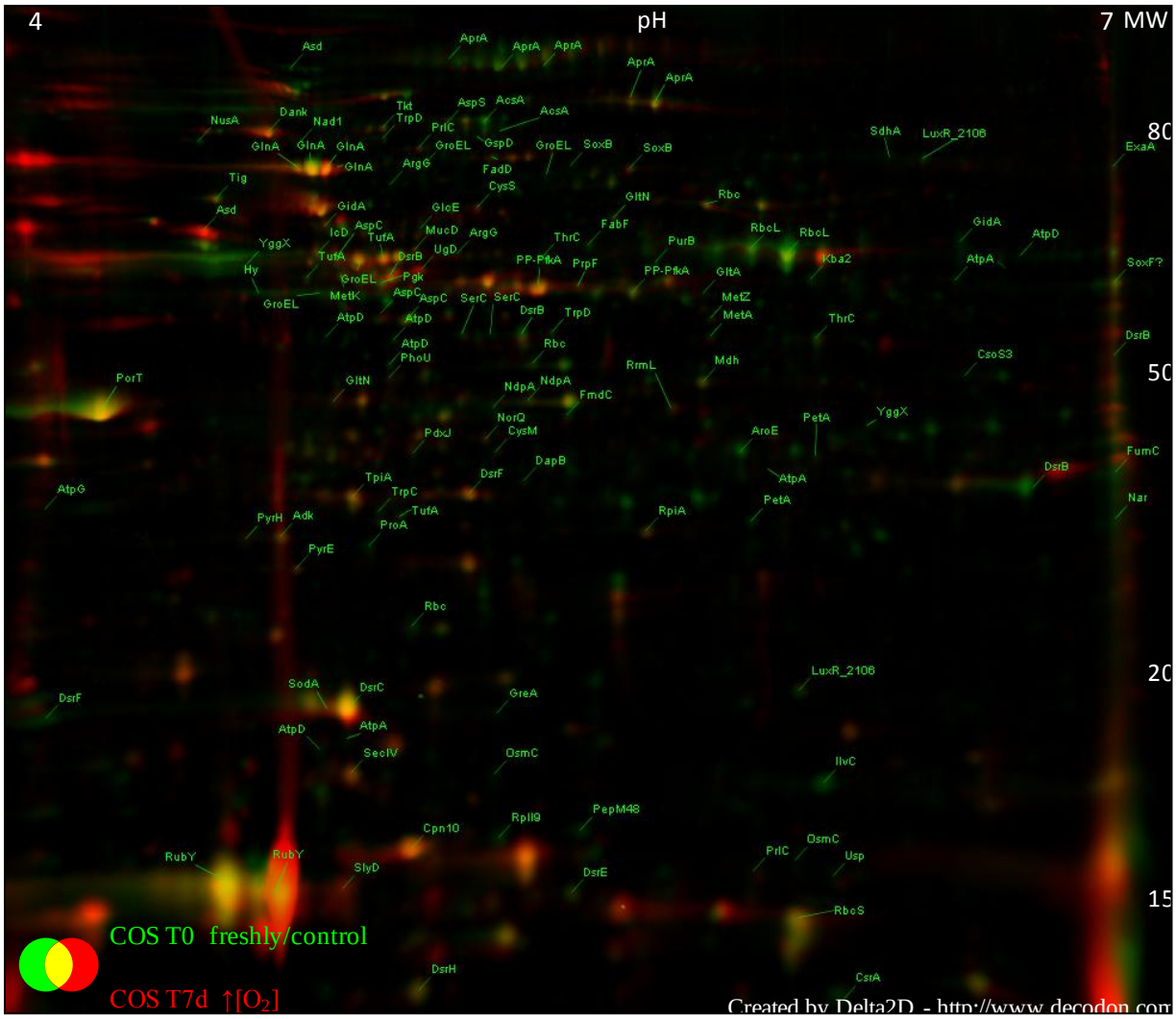

Figure 3.8. False color 2D gel (control/+[O_2]) of the soluble proteins of *C. orbicularis* (COS) symbiont freshly collected (green) and one week of incubation of the bivalve under starvation conditions with oxygen (red). The proteins were separated in the pH range 4-7. All identified proteins are indicated, green marked proteins were identified in freshly *C. orbicularis* symbiont. (d - days)

In figure 3.8 a false color 2D gel picture shows the comparison of the proteome of the symbiont in freshly collected bivalves and in bivalves after one week of incubation with oxygen. Most of the protein spots are yellow that means they have the same expression level under both conditions. Small green spots are visible, these spots are only identified in the control sample. The RuBisCO protein is down-regulated after one week incubation. One of the two rubrerythrin spots is abundant up-regulated in the one week starvation sample with oxygen. The APS reductase AprA protein spots are green, that means under reduced sulfur starvation conditions it is down-regulated.

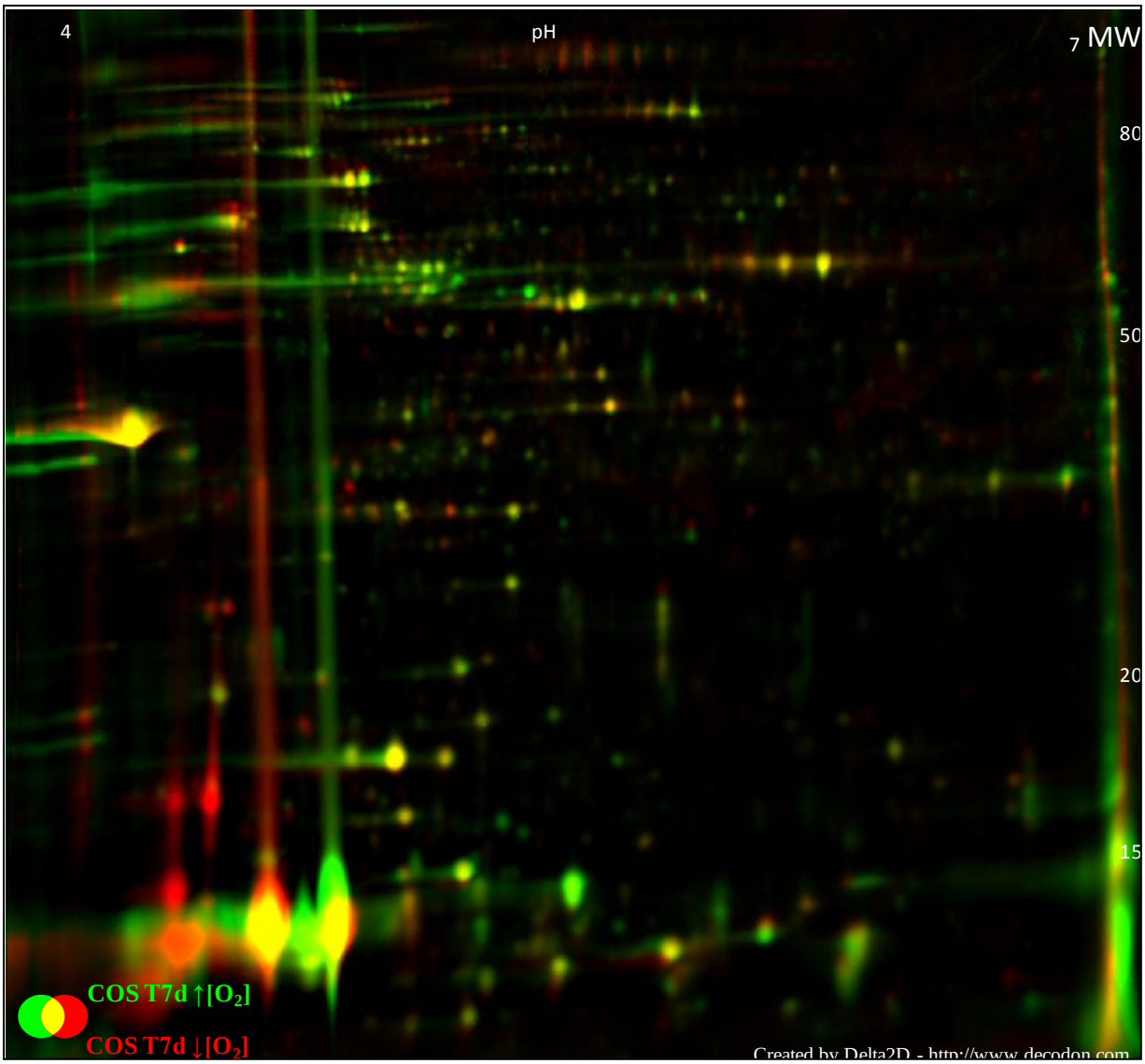

Figure 3.9. False color 2D gel picture (+[O_2]/-[O_2]) of the *C. orbicularis* (COS) symbiont seven days of incubation with normal and saturated oxygen conditions and sulfide starvation.
The proteins were separated in the pH range 4-7. (green - one week of starvation with saturated oxygen concentration, red - one week of starvation with low oxygen conditions)(d-days)

In the figure 3.9 a false color picture of one week starvation with saturated and low oxygen concentration is shown. Most of the proteins are in a green color, which indicate that these proteins are up-regulated under saturated oxygen conditions.
The two spots of rubrerythrin show a differential expression level, the right one is up-regulated under saturated oxygen concentrations and the left one is up-regulated under low oxygen conditions. Further analysis are necessary to explain this expression schema. Also some big red spots are visible these spots are up-regulated under low oxygen conditions.

3.2.2 Proteome of freshly *C. orbicularis* and *C. orbiculata* symbionts

We analyzed the soluble proteins and an enrichment of membrane proteins in 1D SDS PAGE. The samples were analyzed with a LC/MS-Orbitrap, a very sensitive mass spectrometer. The identified proteins were relatively quantified with a

normalized spectral abundance factor (NSAF) based on the measured spectral counts (Zhang et al., 2010). These 1D gel proteomics approaches are dependent on the possibility that the proteins are soluble and can be separated in a SDS gel. In the gel free approach no pre-separation of the proteins is necessary. Another advantage of the gel free approach with the Synapt G2 mass spectrometer is the label free absolute quantification of the identified proteins with the help of the internal standard.

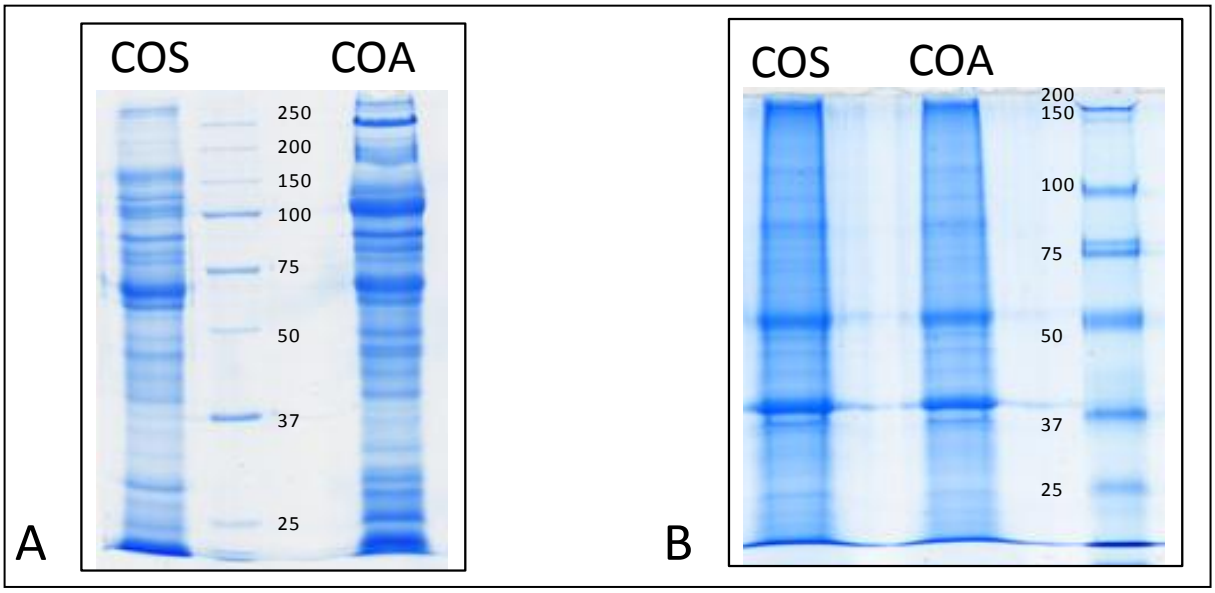

Figure 3.10. 1D-SDS-PAGE of the soluble and membrane symbiont fraction.
The gels were stained with Coomassie Brillant Blue. A) soluble proteins *of C. orbicularis* (COS) and *C. orbiculata* (COA) symbiont B) Enrichment of the membrane proteins of both *Codakia* symbionts. (marker in kDa)

The freshly collected bivalves were transported to the labor and were as fast as possible killed and the gill dissected to extract the symbionts. This samples works as a control sample. In the 1D gel of the soluble protein fraction of both bivalves are different proteins bands visible, at 100 kDa in the *C. orbiculata* symbiont a big protein band is visible in contrast to the *C. orbicularis* symbiont only a very weak protein band is visible (figure 3.10-A). In the 1D gel of the membrane protein enrichment no differences in protein bands are visible (figure 3.10-B).
An overview about the proteomics results in control samples is shown in table 3.4. The spreading between the two bivalves in the different approaches is quite high in the membrane enrichment and the soluble gel free analysis. The reasons for that are diverse. Salt is a disruptive factor of the mass spectrometer measurement and contain in the marine environmental samples. For the gel free analysis we could adapt the protocol to remove more salt from the sample this increased the protein identifications.

Table 3.4. Overview about the identified symbiont proteins of freshly collected bivalves.
The 1D gel approach was used for the soluble and membrane symbiont protein fraction and measured with a Orbitrap mass spectrometer. Furthermore the soluble proteins were analyzed with a SynptG2 in a gel free approach. The shown proteins are "sure" identified. (COS - *C. orbicularis symbiont*, COA - *C. orbiculata symbiont*)

	1D soluble	1D membrane	Gel free soluble
COS	846	1020	451
COA	764	629	190

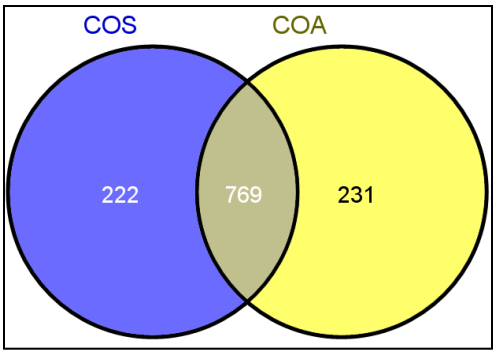

Figure 3.11. Venn diagram comparison of the identified proteins in freshly *C. orbicularis* and *C. orbiculata* symbionts. For the comparison identified proteins from all fractions and approaches were jointed, duplicates were removed. (COS - *C. orbicularis*, COA - *C. orbiculata*)

The comparison of the two bivalve symbionts with the Venn diagram (figure 3.11) shows that the main part of identified proteins was found in both bivalves symbionts. Unfortunately, between the different proteomic approaches is a fluctuating of the identified proteins.

The membrane proteins were enriched following the method to Eymann et al., (2004). The gel picture comparison (figure 3.10-A/B) of soluble and membrane enriched proteins looks different. This is a first indication that different proteins could be enriched. In summary in *C. orbiculata* 1020 and in *C. orbicularis* 629 proteins were identified in the membrane enrichment. Membrane proteins are mainly characterized by transmembrane helices (TMH) to fix the proteins in the membrane. In *C. orbiculata* 149 putative membrane proteins with more than two TMHs were identified and in C. *orbicularis* 81 putative membrane proteins were found. In figure 3.12-A/B an overview about the proteins which contain TMHs in the soluble and membrane enrichment fraction are shown. In both cases most of the proteins are without any TMH. The percentage of the TMHs containing proteins is higher in the membrane protein enrichment. In the genome 78 % of the genes own no TMH, the other genes have one or more TMHs. In figure 3.13 a functional classification of the proteins with more than two TMHs is shown. Proteins related to the environmental information processing and energy metabolism are the biggest part of proteins with more than two TMHs. Also a lot of the identified hypothetical proteins contain TMHs.

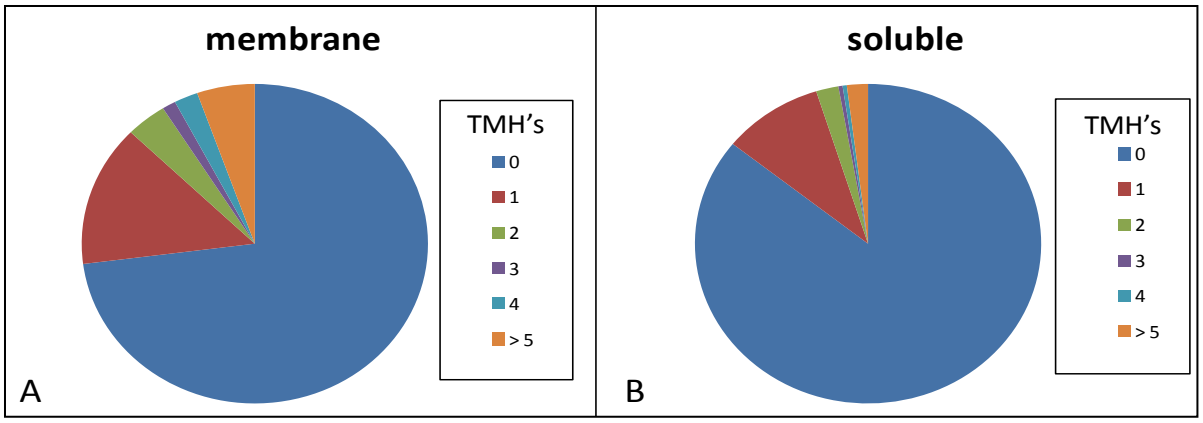

Figure 3.12. Comparison of the TMH's of the membrane enriched and the soluble protein fraction from *C. orbicularis* symbiont 1D gel Orbitrap analysis.
A) - Number of TMH's of all identified proteins in the membrane protein fraction. B) - Number of TMH's of all identified proteins in the soluble protein fraction. Both fractions were analyzed in a 1D gel and measured with a Orbitrap mass spectrometer.

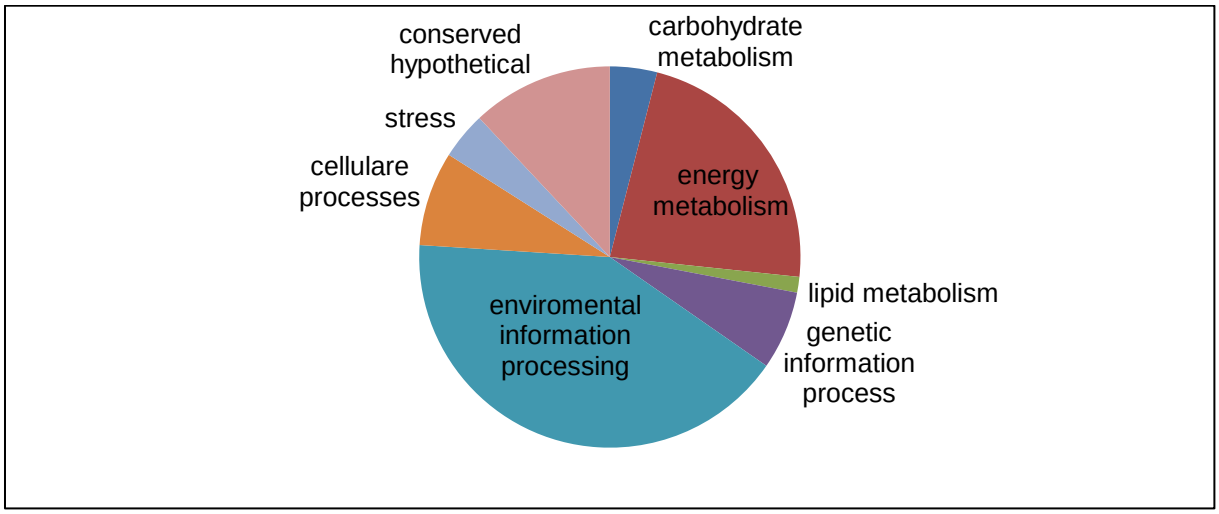

Figure 3.13. Overview about the functions of the identified proteins with more than two TMHs in *C. orbicularis* symbiont. (classified to the Kegg metabolic categories (TMH - transmembranehelices)

The comparison of the symbiont collected from *C. orbicularis* and *C. orbiculata* on the proteomic level was one aim of this study. The most abundant key proteins in the gel free analysis of *C. orbicularis* compared to *C. orbiculata* are shown in figure 3.14.

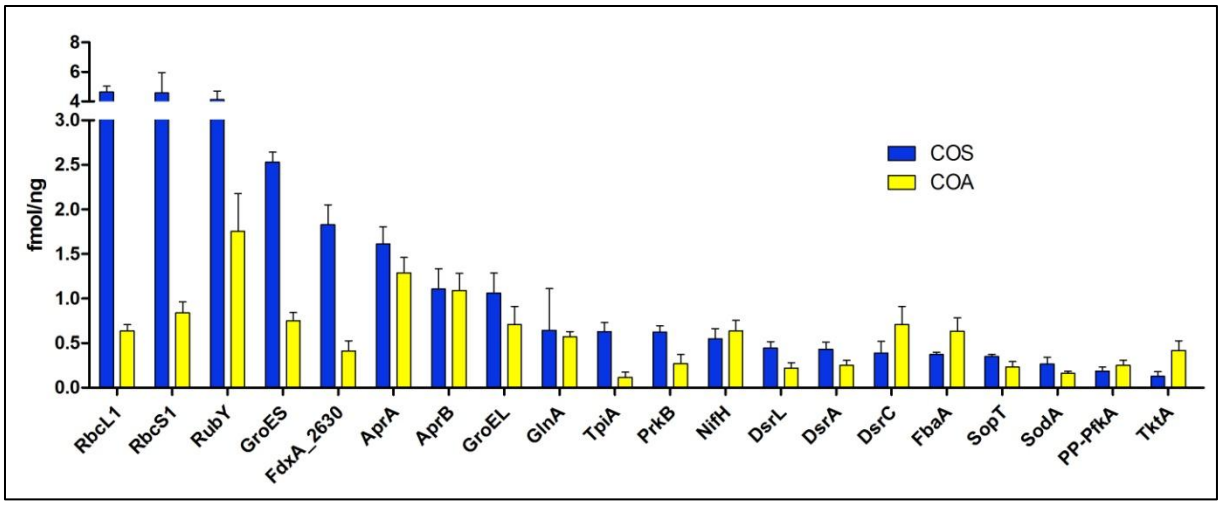

Figure 3.14. Comparison of the most abundant metabolic key proteins of freshly *C. orbicularis* and *C. orbiculata* symbionts. The data represent the mean ± standard deviation of three independent experiments each performed in three technical replicates measured with the Synapt G2 mass spectrometer. (RbcL1 - ribulose bisophosphate carboxylase large subunit, RbcS1- ribulose-1,5-bisphosphate carboxylase/oxygenase small subunit, RubY- rubrerythrin, GroEL/ES- chaperone, FdxA_2630- ferrodoxin, AprAB- adenylylsulfate reductase subunit alpha/beta, GlnA- glutamine synthetase, type I, TpiA- triosephosphate isomerase, PrkB- phosphoribulokinase, NifH- nitrogenase, Dsr- dissimilatory sulfite reductase and associated proteins, FbaA- fructose-1,6-bisphosphate aldolase, SopT- ATP sulfurylase, SodA-superoxide dismutase (Fe), PP-PfkA- PP dep. 6-phosphofructokinase, TktA- transketolase) (COS - *C. orbicularis*, COA - *C. orbiculata*)

The two subunits of RuBisCO form I show the biggest difference between the bivalves. The RuBisCO subunits are in the *C. orbicularis* symbiont the most abundant proteins. Also other proteins, phosphoribulokinase (PrkB) and transketolase (TktA), related to the carbon dioxide fixation via the CBB were identified in a high amount. This indicates the importance of the CBB for carbon fixation of the bivalve symbionts. In the *C. orbiculata* symbiont the amount of RuBisCO is much lower, the most abundant protein is rubrerythrin.

Proteins related to the sulfur oxidation metabolism like adenylylsulfate reductase subunit alpha/beta (AprAB) and dissimilatory sulfite reductase and associated proteins (DsrL, DsrA) were identified in a high amount. Further proteins for nitrogen fixation, nitrogenase (NifH) and stress adaption for oxygen, superoxide dismutase (SodB) were identified in both bivalve symbionts in comparable amounts.

The RuBisCO proteins are much more abundant in *C. orbicularis* symbiont, the PPi-dependent phosphofructokinase (PP-PfkA) is in both bivalves expressed in the same level. The sulfur metabolism has in both bivalves symbionts in the same expression level.

Notable is the high amount of the chaperon GroEL/ES, even under freshly conditions in *C. orbicularis* symbiont. GroEL can be an indication for the stress level of the organism, especially heat stress. The freshly collected bivalves were dissected quickly after collecting. The bivalves were stored for the transport from the seagrass bed to the labor in fresh sea water.

3.2.3 Protein functions in freshly collected *Codakia* symbiont

The identified proteins were manual classified in the Kegg metabolic pathways and used to reconstruct the metabolic pathways of the symbiont.

3.2.3.1 Carbon metabolism

3.2.3.1.1 Carbohydrate metabolism

(1) Glycolysis - Gluconeogenesis - Glycogenesis

In the genome all genes of the glycolysis and gluconeogenesis are encoded. In the proteome all enzymes necessary for the conversion of glucose-6-phosphate to pyruvate were identified in the symbiont proteome (table 3.20). In figure 3.16 the glycolysis pathway is shown as suggested by the genome and proteome data. The glucosidase and the ketohexokinase are encoded in the genome but were not expressed. The figure 3.15 show the glycolytic enzymes identified in the gel free approach.
This makes sense because the symbiont grows autotrophic and an activation of different sugar molecules for heterotrophic degradation is not necessary.
Most of the enzymes for gluconeogenesis are enzymes of the glycolysis working in the reverse direction. The phosphoenolpyruvate synthase (PpsA ID 1820) and the fructose-1,6-bisphosphatase (FbpA ID 1631) were very less identified in the proteome. In summary in the genome all glycolytic and gluconeogenesis genes were encoded but only less expressed. Even in a predicted autotrophic lifestyle of the symbiont, the intermediates of these pathways to synthesize amino acid and other intermediate metabolism compounds that could explain the low expression.
The genome of the *Codakia* symbiont encodes two fructose-1,6-bisphophsphate aldolases, one is predicted specific for the CBB (FbaAII ID 3037) and the second one the glycolysis (FbaAI ID 478). The FbaAI was only once identified in the proteome. The FbaAII is encoded in a gene cluster with glycolysis genes and abundant identified in a high amount. Additional a ketose-bisphosphate aldolase, class II (KbaII ID 895) is encoded in the genome and abundant identified but in a low amount.

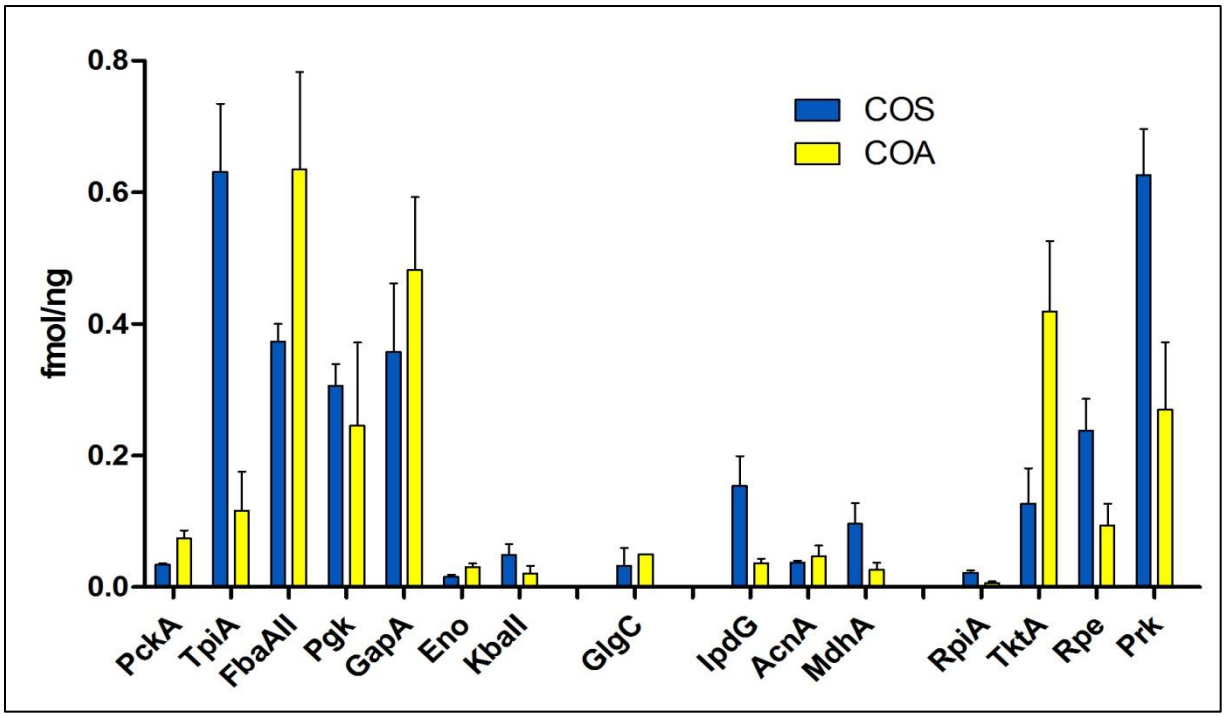

Figure 3.15. Identified carbohydrate metabolism related enzymes of the *C. orbicularis* and *C. orbiculata* symbiont The proteins were classified to the metabolic pathway glycolysis, PPP, TCA-cycle and glycogen metabolism. The data represent the mean ± standard deviation of three independent experiments each performed in three technical replicates measured with the Synapt G2 mass spectrometer. (PckA - phosphoenolpyruvate carboxykinase, TpiA - triosephosphate isomerase, FbaAII - fructose-1,6-bisphosphate aldolase, Pgk - phosphoglycerate kinase, GapA - glyceraldehyde-3-phosphate dehydrogenase, type I, Eno - enolase, KbaII - ketose-bisphosphate aldolase, class-II, GlgC - Glc-1-P adenylyltransferase, IpdG - dihydrolipoamide dehydrogenase, AcnA - aconitate hydratase 2, MdhA - malate dehydrogenase, RpiA - ribose-5-P isomerase, TktA - transketolase, Rpe - ribulose-P 3-epimerase, Prk - phosphoribulokinase)(COS- *C. orbicularis*, COA - *C. orbiculata*)

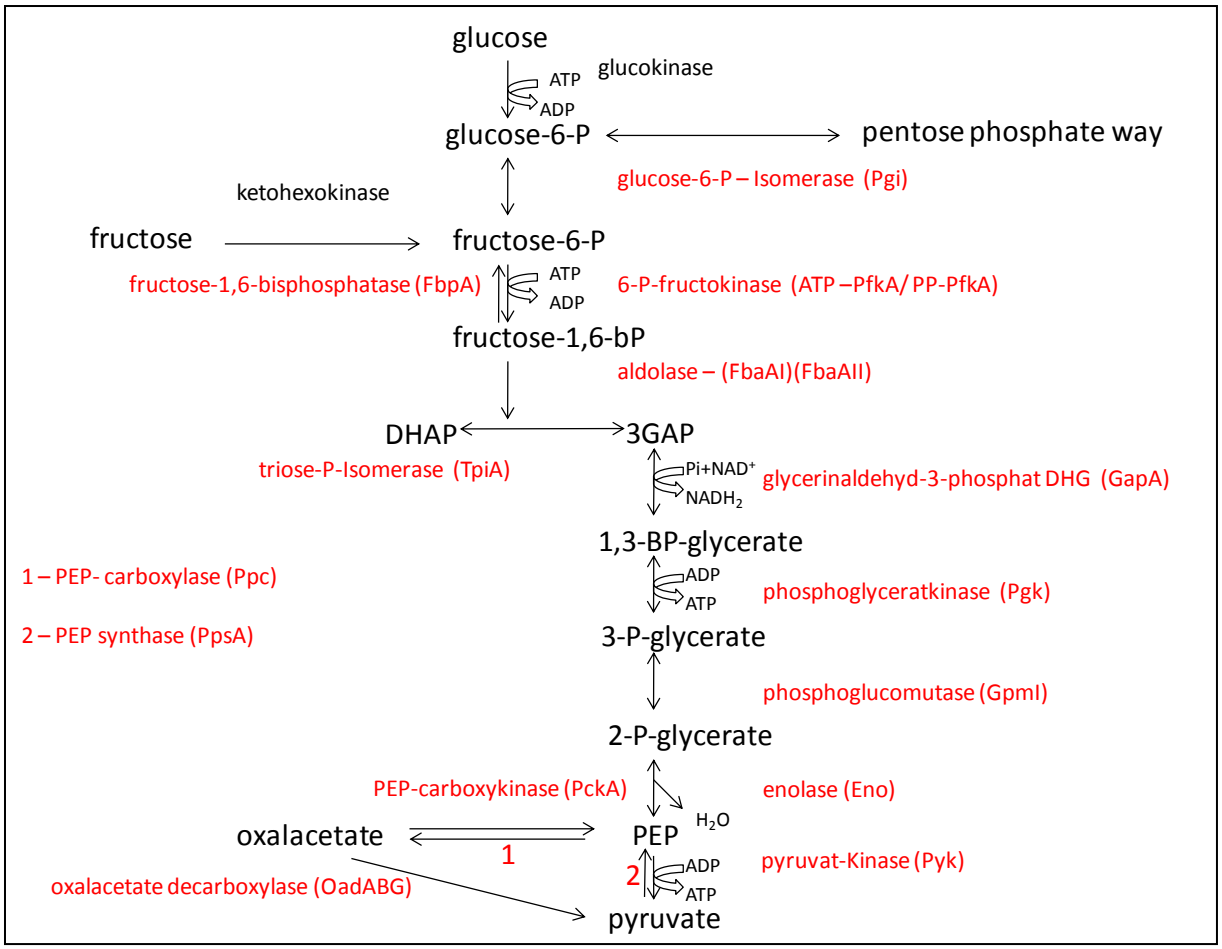

Figure 3.16. Putative glycolysis/gluconeogenesis pathway of the *Codakia* symbiont.
(red marked enzymes identified in the symbiont proteome, black marked enzymes only encoded in the genome). (P - phosphor, bP - bisphosphate)

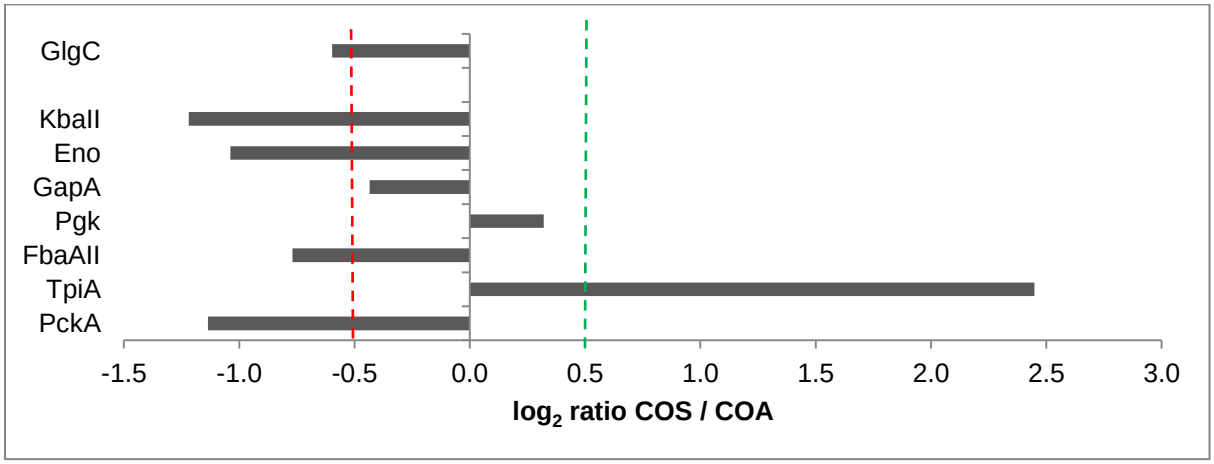

Figure 3.17. Comparison (log_2 ratio) of the identified nitrogen fixation proteins in *C. orbicularis and C. orbiculata* symbiont. Proteins with a positive log_2 ratio are up-regulated in *C. orbicularis* symbiont, proteins with a negative log_2 ratio are up-regulated in *C. orbiculata* symbiont. Ratios more than -0,5/0,5 indicate a significant change of the protein expression.

The comparison of the glycolytic related enzymes in both *Codakia* bivalves shows an up-regulation in the *C. orbiculata* symbiont, only the triosephosphate isomerase (TpiA) is significant up-regulated in the *C. orbicularis* symbiont (figure 3.17).

Glycogen is a polysaccharide of glucose, it serves as energy and carbon storage. In the genome several genes related to the glycogenesis and glycogenolysis are encoded. Four genes are organized in a gene cluster (marked in table 3.5 with double lines). The putative synthesis and degradation pathways of glycogen are shown in figure 3.18.

The degradation of the glycogen starts at the non reducing end one by one, catalyzed by the glycogen phosphorylase to glucose-1-phosphate. In the genome three genes were found annotated as glycogen phosphorylase (ID 1129/2103_2104/ 3764), in the proteome it is only one time identified independent of the conditions of the proteomics investigation.

The glucose-1-phosphate must be converted to glucose-6-phosphate to enter the glycolysis, for this conversion the phosphoglucomutase (Pgm ID 197) is responsible. Unfortunately only the glucose-1-phosphate adenylyltransferase (GlgC ID 38) was identified in the gel free approach, thus absolute quantified. The GlgC is up-regulated in the *C. orbiculata* symbiont.

Table 3.5. Overview about the glycogen synthesis and degradation genes in *C. orbicularis* symbiont genome. Genes organized in a cluster are separated by double lines. ("Proteomics" - number of identifications in the symbiont proteome, Cello - subCELlular LOcalization prediction, CP - cytoplasm).

ID	Product	Label	IEP	MW (kDa)	Cello	TMH	Prote-omics
37	glycoside hydrolase family protein	Gh_37	4.8	64.8	CP	0	3
38	glucose-1-phosphate adenylyltransferase	GlgC	5.8	47.2	CP	0	10
39	1,4-alpha-glucan-branching enzyme	GlgB	5.1	79.3	CP	0	8
41	glycogen/starch synthase, putative ADP-glucose type	GlgA	4.8	54.0	CP	0	4
49	UTP-glucose-1-phosphate uridylyltransferase	GalU	4.5	31.1	CP	0	4
197	phosphoglucomutase	Pgm	5	59.5	CP	0	7
796	glycogen debranching protein	GlgX	6	74.6	CP	0	0
1099	glucokinase	Glk	5.3	36,6	CP	0	0
1129	glycogen/starch/alpha-glucan phosphorylase	GlgP	5.3	92.9	CP	0	1
2103/ 2104	glycogen/starch/alpha-glucan phosphorylase	GlgP	6.7	92.0	CP	0	0
3764	glycogen/starch/alpha-glucan phosphorylase	GlgP	6.6	94.9	CP	0	0

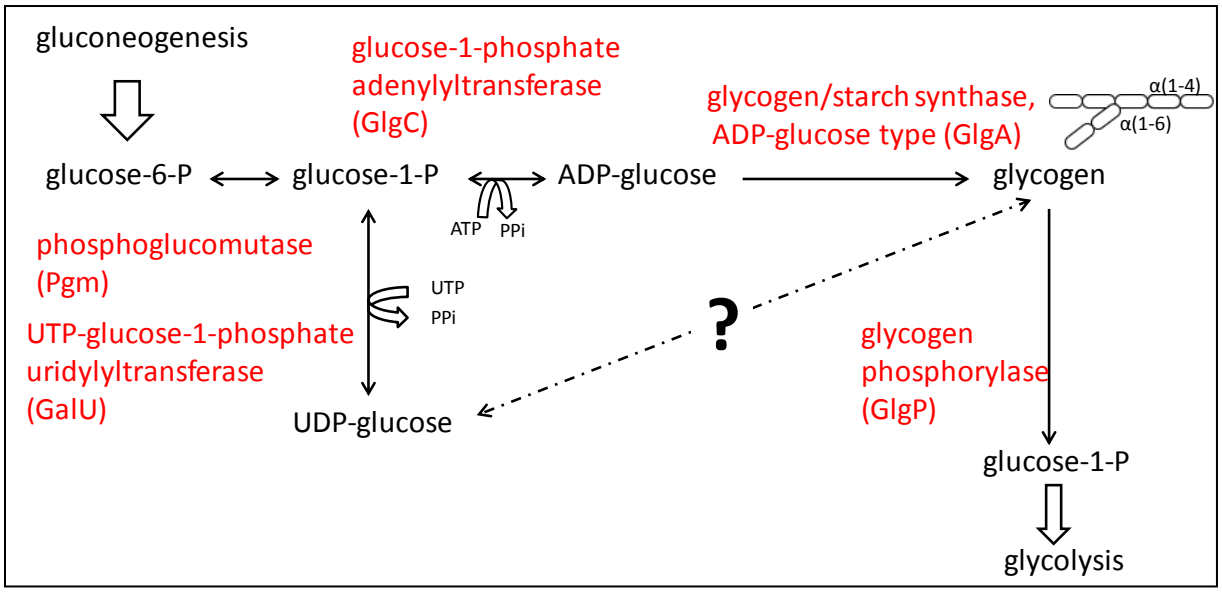

Figure 3.18. Overview about the different putative glycogen synthesis and degradation pathways encoded in the *Codakia* symbiont genome. (red marked proteins are encoded in the genome and identified in the symbiont proteome)(P- phosphor)

(2) Tricarboxylic acid cycle (TCA-cycle) / reverse tricarboxylic acid cycle (rTCA-cycle)

All genes of the TCA-cycle were annotated in the genome and identified in the proteome (table 3.21, figure 3.15./3.19). In the genome genes were found for the glyoxyate cycle, the isocitrate lyase (AceA ID 1805) and the malate synthase (AceB ID 1806), co-localized in the genome. In the proteome only the AceA was identified. The malate synthase A is difficult to identify with the mass spectroscopy (personal communication T. Schweder). It seems that the glyoxylate pathway plays under the investigated conditions not a key role in the symbiont metabolism.

All TCA-cycle genes were identified in a low amount, the reason for that is described in the glycolysis chapter (above). Noticeable is that in freshly collected *C. orbiculata* symbiont much more TCA-cycle enzymes were identified but only in a very low amount, the reason for that are not clear (figure 3.20). However in symbionts of freshly collected *C. orbicularis* only two TCA-cycle enzymes were identified. Nevertheless the proteomic data indicate that the TCA-cycle pathway is expressed in the endosymbiont lifestyle of the symbiont in freshly collected bivalves.

Most of the genes of the reverse TCA-cycle are the enzymes from the energy generated TCA-cycle working in reverse direction. In table 3.6 the putative encoded proteins unique of the rTCA-cycle in the *Codakia* symbiont are shown. Only one protein is expressed, the citrate lyase (AlcB ID 4029). The citrate lyase is composed of three subunits: alpha, beta and gamma. The alpha and beta subunit are required for a catalytic active enzyme, the gamma subunit serves as an acyl carrier protein. In the genome only a citrate lyase subunit beta is encoded, but the annotation of this protein is not very strong. The encoded citrate lyase shows no similarity to the ATP-citrate lyase of the endosymbiont of *Tevnia jerichonana*, in the genome of the *Codakia* symbiont no homolog protein was found for this. The other putative reverse TCA-cycle enzymes could not be identified in any symbiont proteome under the investigated conditions.

The genomic data are not completely verified, due to the fact that missing subunits of the AlcB might be in the sequencing gaps. However, the proteomic data suggest that the symbiont do not fix CO_2 via the rTCA-cycle under the investigated conditions.

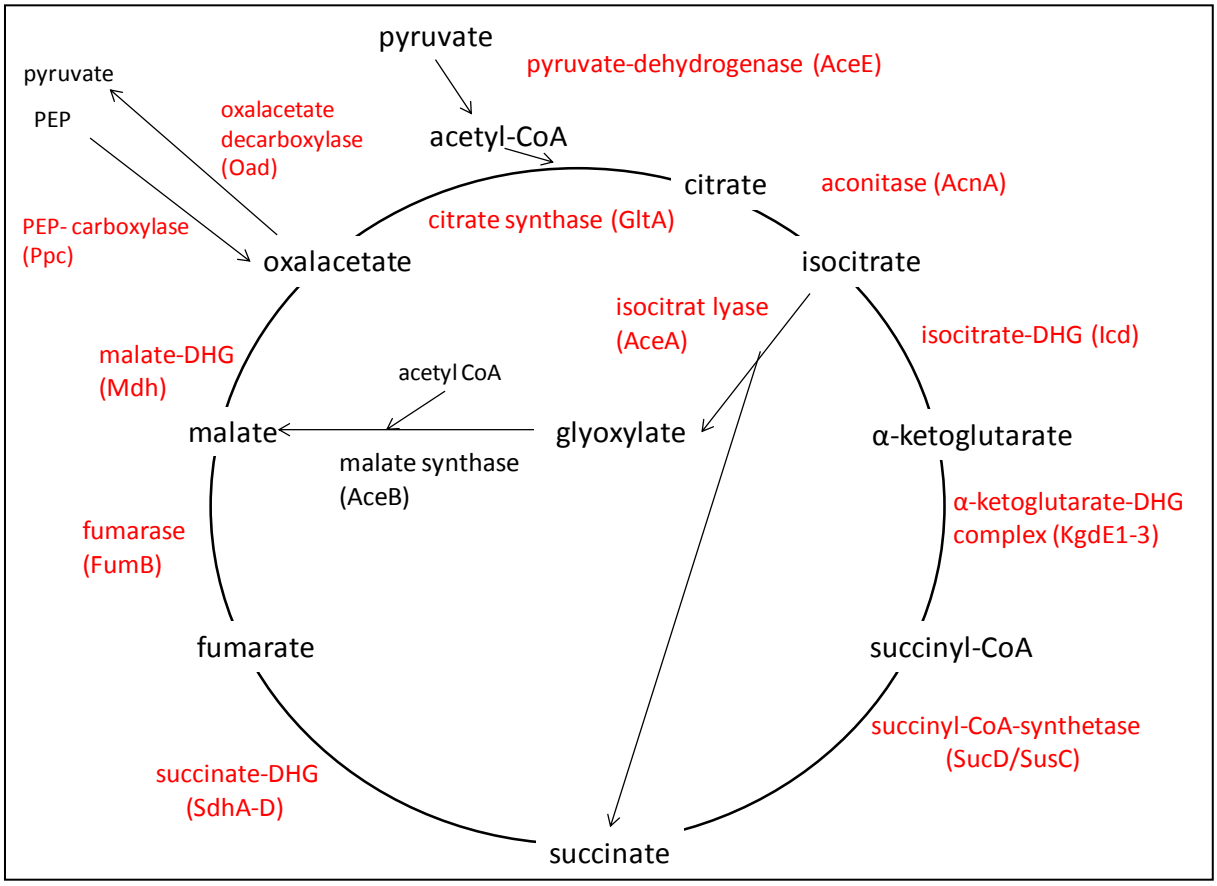

Figure 3.19. Predicted TCA-cycle in *Codakia* symbiont.
(red marked enzymes were identified in the proteome, black marked enzymes are encoded in the genome)

Table 3.6. Putative reverse TCA-cycle enzymes encoded in the *Codakia* symbiont genome.
(eValue and identity from NCBI blastP) ("Proteomics" - number of identifications in the proteome of the symbiont)

ID	Product	Label	IEP	MW (kDa)	eValue	Identity in %	Prote-omics
131	2-oxoglutarate ferredoxin oxidoreductase, beta subunit	KorB	6	26.0	$3.0e^{-81}$	76	0
132	2-oxoglutarate ferredoxin oxidoreductase, alpha subunit	KorA	5.2	61.9	0.0	68	0
968	pyruvat synthase alpha	PorA	4.8	45.6	$2.0e^{-180}$	73	0
969	pyruvat synthase beta	PorB	8.5	35.8	$4.0e^{-128}$	72	0
2680	2-oxoglutarate ferredoxin oxidoreductase subunit beta	KorB	6.4	35.3	$1.0e^{-133}$	69	0
2681	2-oxoglutarate oxidoreductase, alpha subunit	KorA	5.1	68.6	0.0	70	0
4029	putative citrate lyase subunit beta	AclB	5	35.4	$3.0e^{-134}$	71	10

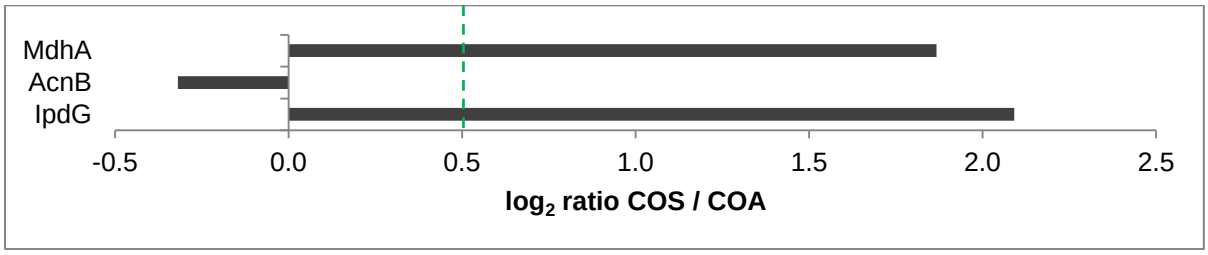

Figure 3.20. Comparison (log_2 ratio) of the identified TCA-cycle proteins in *C. orbicularis and C. orbiculata* symbiont. Proteins with a positive log_2 ratio are up-regulated in *C. orbicularis* (COS) symbiont, proteins with a negative log_2 ratio are up-regulated in *C. orbiculata* (COA) symbiont. Log_2 ratios more than -0,5/0,5 indicate a significant change of the protein expression (green/red line).

Two proteins of the TCA-cycle are up-regulated in the *C. orbicularis* symbiont. But only three proteins were identified in the gel free approach in both bivalve symbionts. In freshly collected *C. orbiculata* symbiont more TCA-cycle enzyme were identified but in a very low amount.

(3) Pentose Phosphate pathway (PPP)

The oxidative phase genes of the pentose phosphate way are encoded in the genome and organized in a gene cluster (table 3.22, figure 3.21). The 6-phosphogluconate dehydrogenase, decarboxylating (Gnd ID 9) were never identified in the proteome. The 6-phosphogluconolactonase (PgI ID 10) could only be identified, the glucose-6-phosphate-1-dehydrogenase (Zwf ID 11) were identified three times in the proteome. The oxidative part of the PPP convert sugar molecules and create NADPH. It seems that the symbiont can avoid these reactions and compensate the lack of pre-coursers and NADPH production with other unknown reactions.
The non-oxidative part is incomplete encoded in the symbiont genome, no transaldolase homolog gene could be detected. How the symbiont avoid the catalyzed steps of the transaldolase is until now not clear. The other enzymes of the non oxidative way are abundant identified in the proteome, that indicate the importance of the pathway for the symbiont. The main function of the non oxidative pathway is the conversion of phosphate sugars as pre-courser for other pathways. The comparison between the two bivalve symbionts shows that all the enzymes of the non oxidative way are more abundant in the *Codakia* symbiont, except the transketolase. The incubation of the *C. orbicularis* with high and low oxygen exhibit that these proteins are more expressed under starvation conditions with high oxygen concentration. The comparison between freshly collected and one week of starvation shows that these enzymes are more abundant in the freshly collected sample.

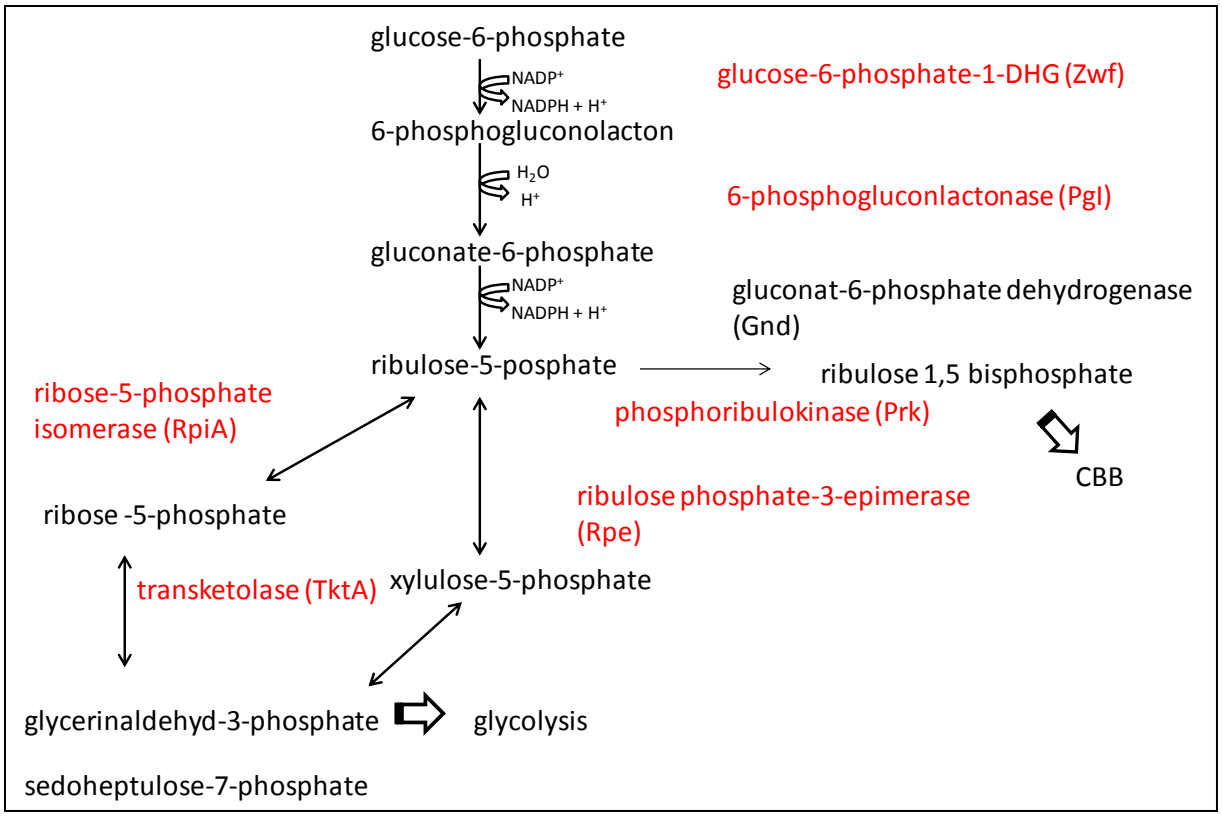

Figure 3.21. Schematic overview about the reconstructed PPP in the symbiont of *Codakia*.
No homolog of transaldolase was found in *Codakia* symbiont genome. (red marked enzymes were identified in the proteome, black marked enzymes are only encoded in the *Codakia* genome)

3.2.3.1.2 Calvin Benson Bassham Cycle (CBB)

RuBisCO is the key enzyme of the CBB. In the genome a cluster with seven CBB related genes is encoded (figure 3.22, table 3.7). The proteomic results of the CBB and related proteins are shown in table 3.23 and figure 3.24 In the genome two different types of RuBisCO were annotated. One gene is related to the form I RuBisCO, homolog proteins were found in the endosymbiont of *Bathymodiolus* sp.. The second RuBisCO gene is related to form II RuBisCO with high similarities to the endosymbiont of *Riftia pachyptila*. Both RuBisCO forms were identified in nearly every measured sample. Interestingly the form I RuBisCO is much more expressed as the form II RuBisCO, which is around 39 fold more expressed in the *C. orbiculata* and around 59 fold in the *C. orbicularis* symbiont. In case of starvation the expression of the RuBisCO form I decreased whereas the expression of form II increased. In the genome two copies of the RuBisCO activation proteins CbbQ (ID: 315/320) and CbbO (ID 316/321) were found but only one set of them is expressed. The CbbQ genes show a quite high similarity (eValue $1.0e^{-148}$, identity 71%), the CbbO genes show only moderate similarity (eValue $1.0e^{-157}$, identity 36%) to each other. Only the activator protein set close to form I RuBisCO was abundant identified in the proteome, the second set was never identified.

The regeneration of ribulose-1,5-bisphosphatase in the CBB is not complete encoded in the *Codakia* symbiont genome, no homolog of the sedoheptulose-1,7-bisphosphatase was found, although both forms of RuBisCO were found.
In the genome two 6-phospofructokinase (ID 984/1107) genes were annotated with low similarities (26% identity and eValue $2.0e^{-20}$) to each other. The ID 984 PfkA shows high similarities to the PPi-dependent 6-phosphofructokinase of *Calyptogena magnifica* (table 3.8). The putative PPi-dependent phosphofructokinase (PP-PfkA) is co-localized in the genome with pyrophosphate energized proton pump (HppA ID 985) (figure 3.23).
In the genome a fructose-1,6-bisphosphate aldolase, class II calvin subtype (FbaAII, ID 3037) is encoded and expressed in the proteome. But this enzyme is encoded in a cluster encoding for glycolysis genes.

Table 3.7. The RuBisCO gene cluster in *C. orbicularis* symbiont genome.
The "NCBI hit" show the best NCBI protein blast hit. (Evalue and identity from NCBI protein blast)

ID	Product	Label	IEP	MW (kDa)	Stra-nd	NCBI hit	eValue	Iden-tity %
312	RuBisCO operon transcriptional regulator CbbR	CbbR	6.6	34.3	-	*Candidatus Thiodictyon syntrophicum str. Cad 16*	$7.0e^{-115}$	68
313	ribulose-1,5-bisophosphate carboxylase, large subunit	RbcL1	5.7	52.4	+	*Thiothrix nivea* DSM 5205	0.0	93
314	ribulose-1,5-bisphosphate carboxylase/oxygenase, small subunit	RbcS1	5	13.4	+	*Solemya* gill symbiont	$2.0e^{-57}$	88
315	RuBisCO activation protein	CbbQ	4.9	30.1	+	*Solemya* gill symbiont	$3.0e^{-142}$	95
316	RuBisCO activation protein	CbbO	4.7	91.4	+	*Solemya* gill symbiont	0.0	86
319	ribulose-1,5-bisphosphate carboxylase, large subunit	RbcL2	5.4	50.8	+	*Thiothrix nivea* DSM 5205	0.0	88
320	RuBisCO activation protein	CbbQ	5.3	29.9	+	endosymbiont of *Tevnia jerichonana*	$2.0e^{-137}$	86
321	RuBisCO activation protein	CbbO	5.1	85.3	+	endosymbiont of *Tevnia jerichonana*	0.0	79

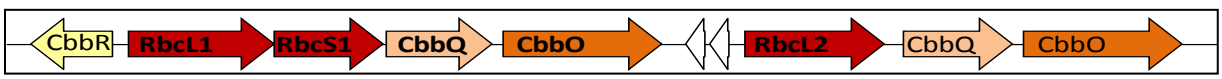

Figure 3.22. RuBisCO gene cluster encoded in the *C. orbicularis* symbiont genome (ID 312-321).
(bold - identified in the proteome)

Table 3.8. NCBI protein blast of PP-PfkA and HppA.
(PP-PfkA - pyrophosphate, HppA - PPi-dependent phosphofructokinase pyrophosphate energized proton pump

Organism	eValue	Identity	eValue	Identity
	PP-PfkA		HppA	
Calyptogena magnifica endosymbiont	0.0	79 %	0.0	71 %
Riftia pachyptila endosymbiont	0.0	90 %	0.0	83 %
Methylococcus capsulatus	0.0	74 %	0.0	73 %

Figure 3.23. The PP-PfkA and HppA are co-localized encoded in *Codakia* symbiont genome.
(PP-PfkA - PPi-depdendent phosphofructokinase ID 984, HppA - pyrophosphate energized proton pump protein ID 985)

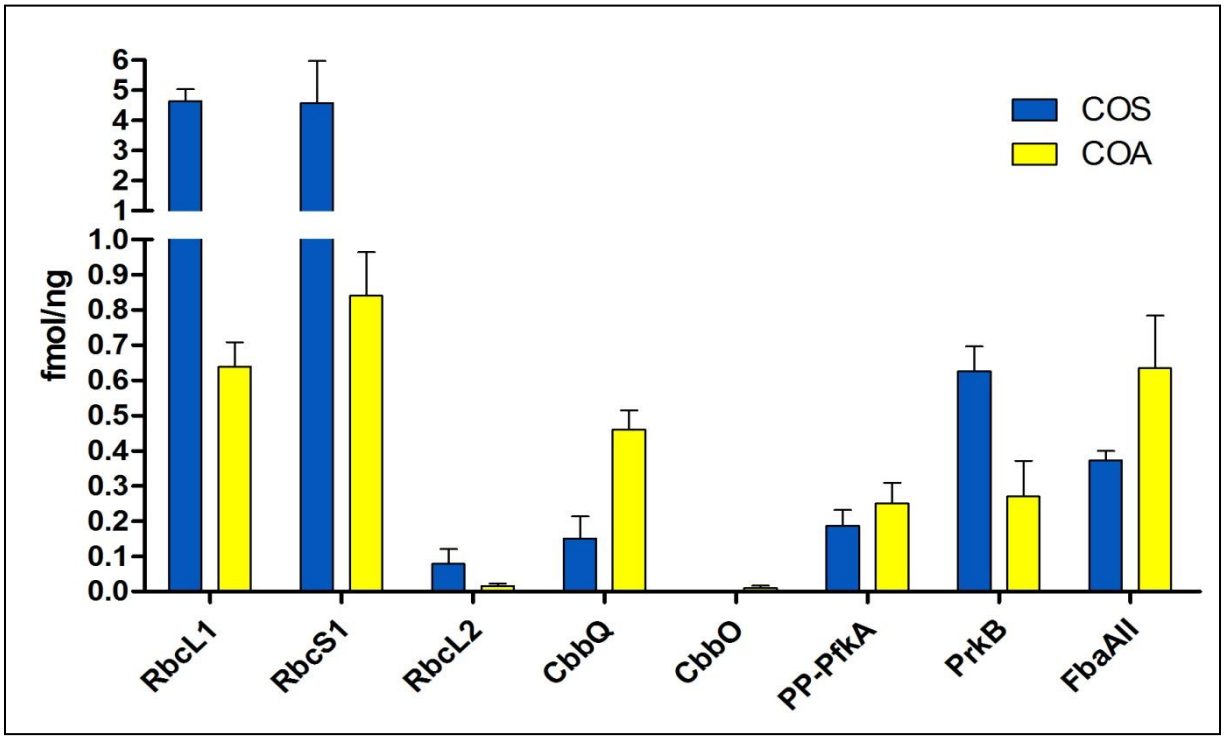

Figure 3.24. Identified proteins of the CBB in the *Codakia* symbiont in gel free approach.
The data represent the mean ± standard deviation of three independent experiments each performed in three technical replicates measured with the Synapt G2 mass spectrometer. (blue - symbiont of freshly collected *C. orbicularis* (COS), yellow - symbiont of freshly collected *C. orbiculata* (COA)) (Rbcl1 - ribulose bisophosphate carboxylase, RbcS1 - ribulose-1,5-bisphosphate carboxylase/oxygenase small subunit, RbcL2 - ribulose-1,5-bisphosphate carboxylase large subunit, CbbQ/CbbO - RuBisCO activation protein, PP-PfkA - 6-phosphofructokinase, PrkB - phosphoribulokinase, FbaAII - fructose-1,6-bisphosphate aldolase, class II)

In both bivalves the key enzyme of the CBB, RuBisCO, is abundant expressed. The two different encoded forms of RuBisCO are diverse expressed. The homolog to form I is much more expressed, nevertheless both forms are expressed at the same time.

Only one set of RuBisCO activator, CbbO and CbbQ (ID 315/316) proteins were identified.

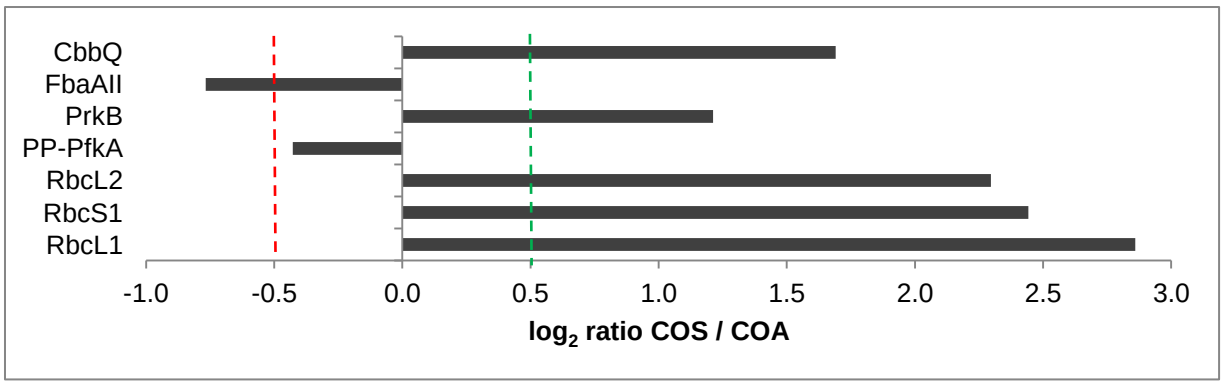

Figure 3.25. Comparison (log_2 ratio) of the identified CBB proteins in *C. orbicularis and C. orbiculata* symbiont. Proteins with positive log_2 ratio are up-regulated in *C. orbicularis* (COS) symbiont, proteins with negative log_2 ratio are up-regulated in *C. orbiculata* (COA) symbiont. Ratios more than -0,5/0,5 indicate a significant change of the protein expression (green/red line).

The key genes, both RuBisCO forms are significant up-regulated in *C. orbicularis* symbiont. The PP-PfkA show the same expression level in both bivalves. The FbaAII is significant up-regulated in the *C. orbiculata* symbiont (figure 3.25).

3.2.3.2 Energy metabolism

3.2.3.2.1 Oxidative phosphorylation

The oxidative phosphorylation requires oxygen as terminal electron acceptor. The related proteins are shown in table S6.6. The ATP-synthase genes are organized in a gene cluster (ID 1067-1072) in the genome and were abundant and in a high amount identified in the proteome (table 3.24). In *C. orbicularis* symbiont the amount of identified ATP synthase proteins was higher.

3.2.3.2.2 Sulfur metabolism

The sulfur oxidation metabolism is a key metabolism in the symbiont, an overview about the identified proteins is shown in figure 3.26 In the genome different enzymatic mechanisms are encoded to oxidize reduced sulfur compounds to sulfate. The first step is the reaction of sulfide to polysulfide or elemental sulfur, three system are encoded for these processes. The (1) Sox-system (Sulfur oxidation), (2) the sulfide dehydrogenase FccAB and (3) the sulfide quinone reductase Sqr (ID 2280) are encoded in the symbiont genome.
The genes are widespread in the genome, but *soxYZ* and *soxXA* are always co-localized in the genome. The genes for *soxB* and *soxW* are encoded on single sites in the genome. In the genome two copies of *soxYZ* are encoded, both proteins show low similarities to each other (SoxY 563 ↔ SoxY 4159 eValue $3.0e^{-15}$, 32% identity / SoxZ 564 ↔ SoxZ 4160 eValue $2.0e^{-10}$, 33% identity). The SoxYZ cluster 4159/4160

was not expressed under the investigated conditions. Most of the predicted *sox* genes were identified in the proteome (table 3.25, S6.2).
In the genome several candidates were found for the sulfide dehydrogenase (*fccAB*) (table S6.1). Only two putative FccAB complex candidates were identified under the given conditions. The sulfide quinone reductase Sqr is encoded in the genome and was identified in the proteome but in a very low amount (table 3.27, table S6.3).
The symbionts store elemental sulfur in the periplasmic space. In *A. vinosum* three different sulfur globule proteins are described (SgpABC) (Brune, 1995). In the *C. orbicularis* symbiont genome a gene was found encoding for a putative sulfur globule protein (*sgpA*), but only with medium similarity (ID 2472, NCBI blast $4.0e^{-14}$, 44% identity). For *sgpB* and *sgpC* no homolog genes were found. This putative SgpA was not identified in the proteome.
In the genome a *dsr*-cluster (Dsr - dissimilatory sulfite reduction) of 14 genes is encoded (table 6.4). The *dsr* cluster encodes a pathway of the oxidation from sulfur to sulfite. The proteomic results are shown in table 3.26. A cytoplasmic adenylylsulfate reductase (AprAB, ID 3699/3700) is encoded (table 6.3) and abundant expressed in the proteome (table 3.27). The AprAB is one of the most abundant proteins in the symbiont proteome. The last step in the sulfur oxidation is catalyzed by the sulfate adenylyltransferase (Sat) / ATP sulfurylase (SopT), which is encoded and abundant expressed. The electrons from the AprAB reaction are assumed to be transfered to the quinone-interacting membrane bound oxidoreductase QmoABC complex, putative genes are shown in table 3.9. To remove the sulfate from the cytoplasm a sulfate transporter SulP (ID 2768/1792) is encoded, one of these putative SulP is expressed (ID 1792) (table 6.3).

Table 3.9. QmoABC proteins encoded in the *Codakia* symbiont genome.
(Cello - subCELlular LOcalization prediction, CP - cytoplasm, "Proteomics" - number of identifications in the symbiont proteome, Qmo - quinone-interacting membrane bound oxidoreductase)

ID	Product	Label	IEP	MW (kDa)	TMH	Cello	Prote-omics
3709	heterodisulfide reductase subunit B	HdrB	4.5	32.4	0	CP	10
3710	heterodisulfide reductase, putative QmoC	QmoC	9.7	23.6	0	CP	9
3711	heterodisulfide reductase, putative QmoB	QmoB	4.4	81.2	0	CP	12
3712	heterodisulfide reductase, putative QmoA	QmoA	5.1	46.4	0	CP	12

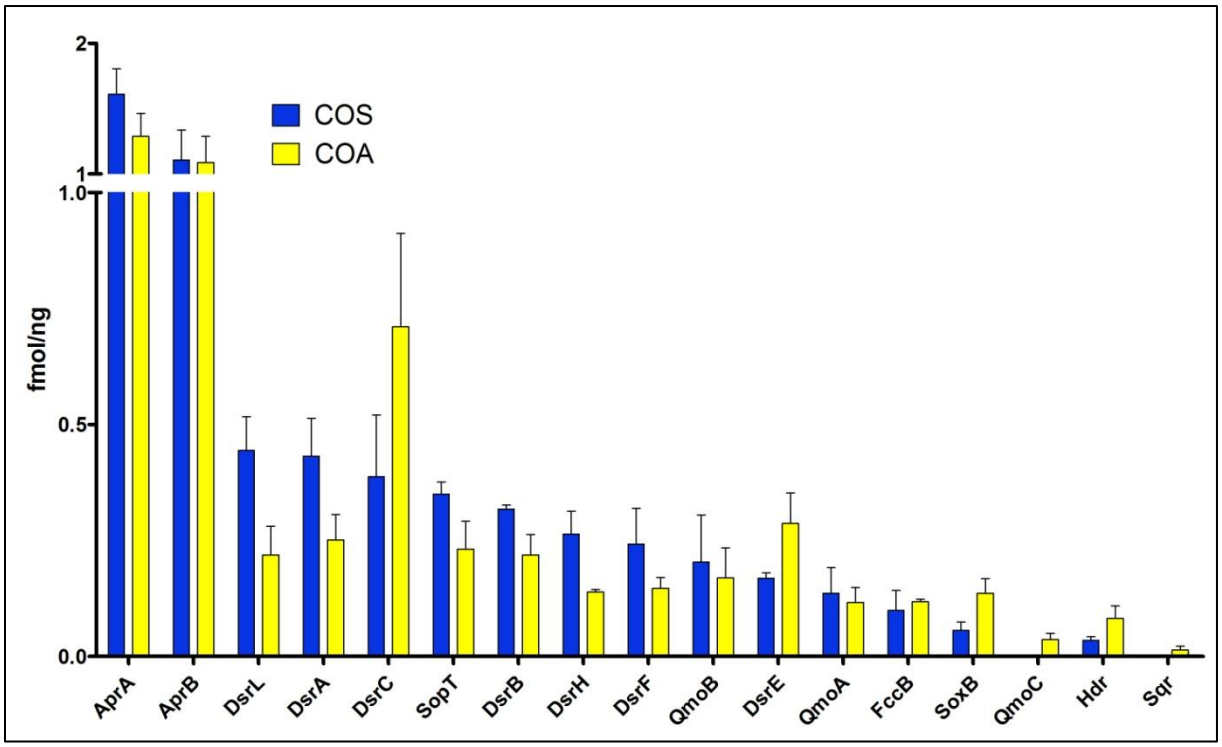

Figure 3.26. Identified proteins in the gel free approach, related to the sulfur metabolism in freshly *Codakia* symbionts. The data represent the mean ± standard deviation of three independent experiments each performed in three technical replicates measured with the Synapt G2 mass spectrometer. (blue - symbiont proteins of freshly collected *C. orbicularis*, yellow - symbiont proteins of freshly collected *C. orbiculata*). (AprAB - adenylylsulfate reductase, Dsr -dissimilatory sulfite reductase and associated proteins,SopT - ATP sulfurylase, FccAB - sulfide dehydrogenase, QmoAB - heterosulfide reducatse, putative QmoAB, Sox - sulfur oxidation proteins, Sqr sulfide quinone reductase)

The comparison of both *Codakia* symbionts show a more or less equal expression schema (figure 3.27). AprAB is in both symbionts abundant expressed. It seems that the Dsr pathway is a little bit more abundant expressed in the *C. orbicularis* symbiont.

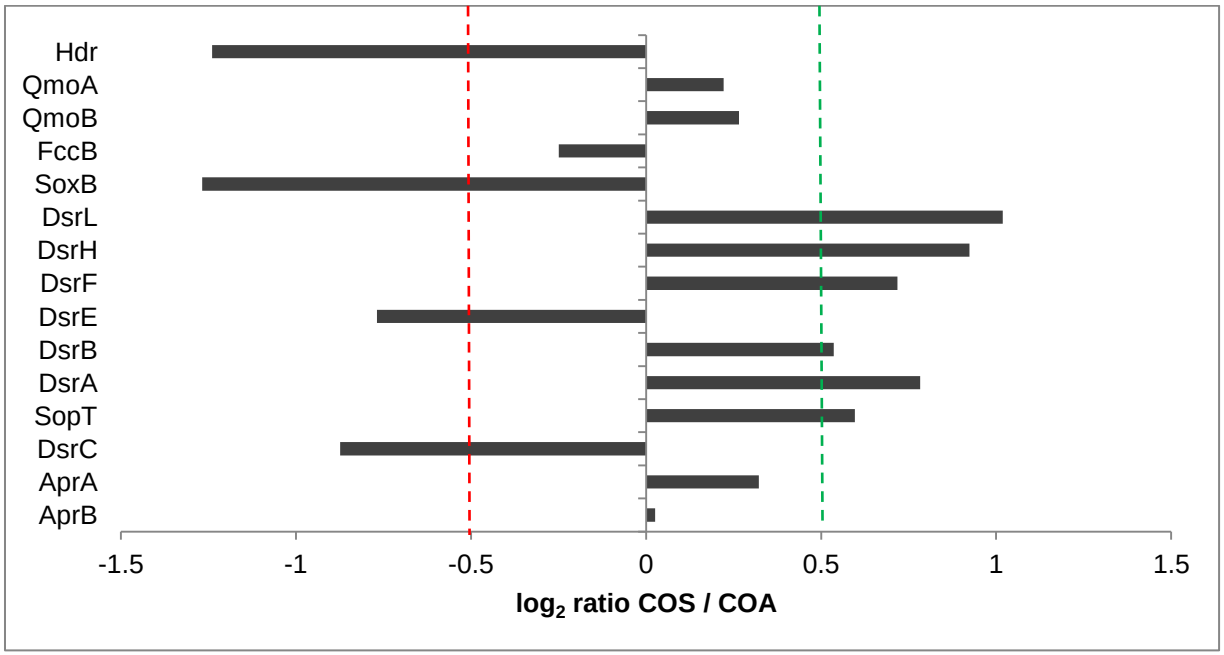

Figure 3.27. Comparison (log_2 ratio) of the identified sulfur oxidation metabolism proteins in *C. orbicularis and C. orbiculata* symbiont. Proteins with positive log_2 ratio are up-regulated in *C. orbicularis* (COS) symbiont,

proteins with negative log_2 ratio are up-regulated in *C. orbiculata* (COA) symbiont. Ratios more than -0,5/0,5 indicate a significant change of the protein expression (green/red line).

3.2.3.3 Nitrogen metabolism

Nitrate can be used as primary nitrogen source for biosynthesis and as well as a respiratory terminal electron acceptor. Interestingly in the genome genes for nitrate reduction to gaseous nitrogen and gaseous nitrogen fixation are encoded (table 3.10 table 6.5). These two metabolic pathways are working in the opposite direction, so it needs a stringent control for these pathways. Also the gaseous nitrogen fixation is an energy consuming pathway. In table 3.12 the putative regulatory proteins of the nitrogen metabolism are shown.

3.2.3.3.1 Nitrate reduction

The dissimilatory nitrate reduction uses different predicted membrane bound or periplasmic respiratory nitrate reductases. The genome of the *Codakia* symbiont encode several sets of enzymes for the dissimilatory nitrate reduction (table 3.10). In the proteome very few indications were found for a dissimilatory nitrate reduction under the investigated conditions. The genes for the dissimilatory nitrate reduction are organized in gene clusters in the genome. NarXL encodes for a nitrate responsive two component regulatory system and is part of one gene cluster.

Table 3.10. Putative dissimilatory nitrate reduction proteins encoded in the *C. orbicularis* symbiont genome. (Cello - subCELlular LOcalization prediction, CP - cytoplasm, PP - periplasm, IM - inner membrane, OM - outer membrane, "Proteomics" - number of identifications in the symbiont proteome). Genes organized in cluster are separated by double lines.

ID	Product	Label	IEP	MW (kDa)	Cello	TMH	Prote-omics
654	nitrous oxide reductase accessory protein NosL	NosL	7.7	20.6	CP	0	0
655	nitrous oxide reductase maturation protein NosF	NosF	6.3	33.0	CP	0	1
656	ferredoxin-type protein, NapH/MauN family	NapH	9.9	37	IM	4	0
657	ferredoxin-type protein	------	5	31	CP	1	0
658	periplasmic copper binding nitrous oxidase accessorry protein NosD	NosD	4.8	53	OM	0	0
662	nitrous oxide reductase, partial	Nor	5.7	83.8	PP	0	0
669	nitric oxide reductase activation protein NorD	NorD	6.1	68.8	CP	0	0
671	denitrification regulatory protein NirQ	NirQ	5.2	30.6	CP	0	0
673	nitric oxide reductase activation protein NorE	NorE	6.3	21.8	IM	5	0
674	NorE_like subfamily of heme-copper oxidase subunit III	NorE	6.8	22.5	IM	5	0
682	nitric oxide reductase, large subunit	Nor	6.1	53.9	IM	12	0
683	nitric-oxide reductase, subunit C	NorC	4.6	16.9	CP	1	0

ID	Product	Label	IEP	MW (kDa)	Cello	TMH	Prote-omics
684	nitric oxide -responding transcriptional regulator Dnr	Dnr	5.6	27.4	CP	0	0
2532	denitrification system component NirT/cytochrome c552	NirT	5.7	102	PP	1	0
2533	nitrite reductase	NirS	6.5	59.0	PP	0	0
2536	nitric-oxide reductase, subunit C	NorC	4.6	16.9	CP	1	0
2537	nitric oxide reductase, subunit B	NorB	6.1	53.9	IM	12	0
2550	nitric oxide reductase activation protein NorD	NorD	6.1	68.8	CP	0	0
2557	nitrous oxide reductase, NosZ	NosZ	5.7	83.8	PP	0	1
2564	nitrous oxide reductase maturation protein NosF	NosF	6.3	33.0	CP	0	0
2565	nitrous oxide reductase maturation protein, outer-membrane lipoprotein NosL	NosL	7.7	20.6	CP	0	0
3714	putative nitric oxide reductase activation protein NorQ	NorQ	5.1	29.3	CP	0	5
3715	putative nitric oxide reductase activation protein NorD	NorD	4.8	81.4	CP	0	1
3935	nitrate/nitrite response regulator protein NarL	NarL	5.3	23	CP	0	0
3936	nitrate/nitrite sensor protein, NarX	NarX	5.5	70	CP	2	0
3937	putative nitrate transporter, NarT	NarT	8.4	52	IM	11	0
3938	nitrate/nitrite antiporter	NarK	9.8	57	IM	14	0
3939	respiratory nitrate reductase 1 alpha chain	NarG	6.3	141.6	PP	0	0
3940	respiratory nitrate reductase beta chain	NarH	5.5	58.5	PP	0	0
3941	respiratory nitrate reductase, delta chain	NarJ	4.7	20.1	CP	0	0
3942	respiratory nitrate reductase, gamma chain	Narl	10.2	25.7	IM	5	0
3976	cytochrome C-type protein NapC	NapC	5.8	21.3	PP	1	7
4298	cytochrome c-type protein NapC	NapC	6.4	21.8	PP	1	0
4299	periplasmic nitrate reductase subunit NapB	NapB	9.7	16.2	PP	1	0
4300	periplasmic nitrate reductase subunit NapH	NapH	9.6	33.4	IM	4	0
4301	periplasmic nitrate reductase subunit NapG	NapG	4.9	31.3	CP	1	0
4302	nitrate reductase catalytic subunit	NapA	9.6	95.6	CP	0	5
4303	protein NapD	NapD	4.2	9.8	CP	0	0
4304	nitrate/nitrite response regulator protein NarL	NarL	5	24.4	CP	0	3
4305	nitrate/nitrite sensor protein NarX	NarX	5.7	70.3	CP	1	0

The assimilatory nitrate is used to form organic nitrogen compounds like amino acids. The assimilatory nitrate reductase is located in the cytoplasm and mostly NAD(P)H

dependent. In the genome the ass. nitrate reductase is organized in a gene cluster (table 3.11). Only few indications for an assimilatory nitrate reduction were found in the proteome (ID 3286/3285). In the gene cluster transporter for nitrate are encoded and expressed.

Table 3.11. Assimilatory nitrate reduction and related proteins encoded in the *C. orbicularis* symbiont genome. Genes organized in cluster are separated by double lines. (Cello - subCELlular LOcalization prediction, CP - cytoplasm, PP - periplasm, IM - inner membrane, "Proteomics" - number of identifications in the symbiont proteome)

ID	Product	Label	IEP	MW (kDa)	Cello	TMH	Proteomics
3284	assimilatory nitrate reductase, large subunit	NarA	6.6	98	CP	0	0
3285	nitrite reductase (NAD(P)H), small subunit	NirD/NasG	5.8	11.4	CP	0	2
3286	nitrite reductase (NAD(P)H), large subunit	NirB	4.9	90.5	CP	0	2
3287	nitrate transporter	NasA	5.2	50	IM	0	0
3288	nitrate transport ATP-binding subunits C and D	NrtCD	6.1	32	CP	0	0
3289	nitrate ABC transporter permease protein	NrtB	10.3	38	IM	6	1
3290	nitrate transporter	NasA	5.6	53	PP	1	9
3291	response regulator NasT	NasT	4.6	23	CP	0	0

3.2.3.3.2 Nitrogen fixation

All organisms need nitrogen for growth and biosynthesis, which means organisms need an exogenous source of nitrogen. The *C. orbicularis* symbiont the nitrogen fixation cluster comprises 44 genes, this cluster is shown with the particular best NCBI best blast hit in figure S6.5. The genes for the nitrogen fixation are mostly highly conserved. Figure 3.29 shows only the best blast hit but in several cases it exist more than one hit with the same eValule and identity.
The cluster comprises the nitrogenase genes (NifH and NifK) and several supporting genes like regulation electron transfer to the nitrogenase and cofactor synthesis genes. The key genes of nitrogen fixation (NifK/D/NifH) are highly conserved and show a very high similarity to related genes. The support genes for the nitrogen fixation are less conserved and show sometimes only low similarities.

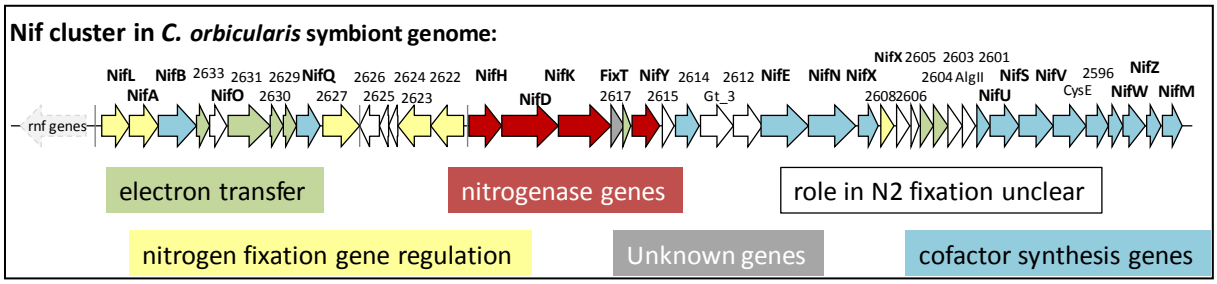

Figure 3.28. Nitrogen fixation related gene cluster encoded in the *C. orbicularis* symbiont genome.
The color of the genes symbolized the predicted molecular function for the nitrogen fixation. (the cluster was generated by RAST and modified from Dr. Stephanie Markert) (gene number = gene ID in database)

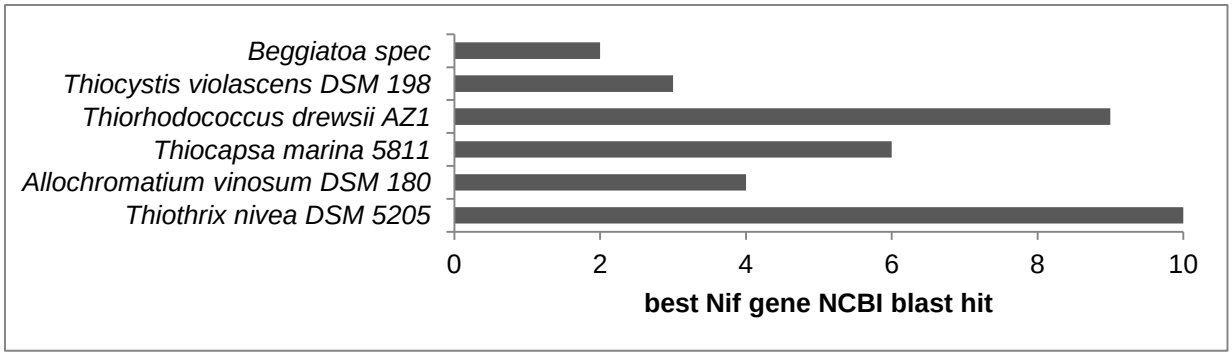

Figure 3.29. Overview about the best NCBI blast hit of the 44 nitrogen fixation related proteins in the *C. orbicularis* symbiont genome.

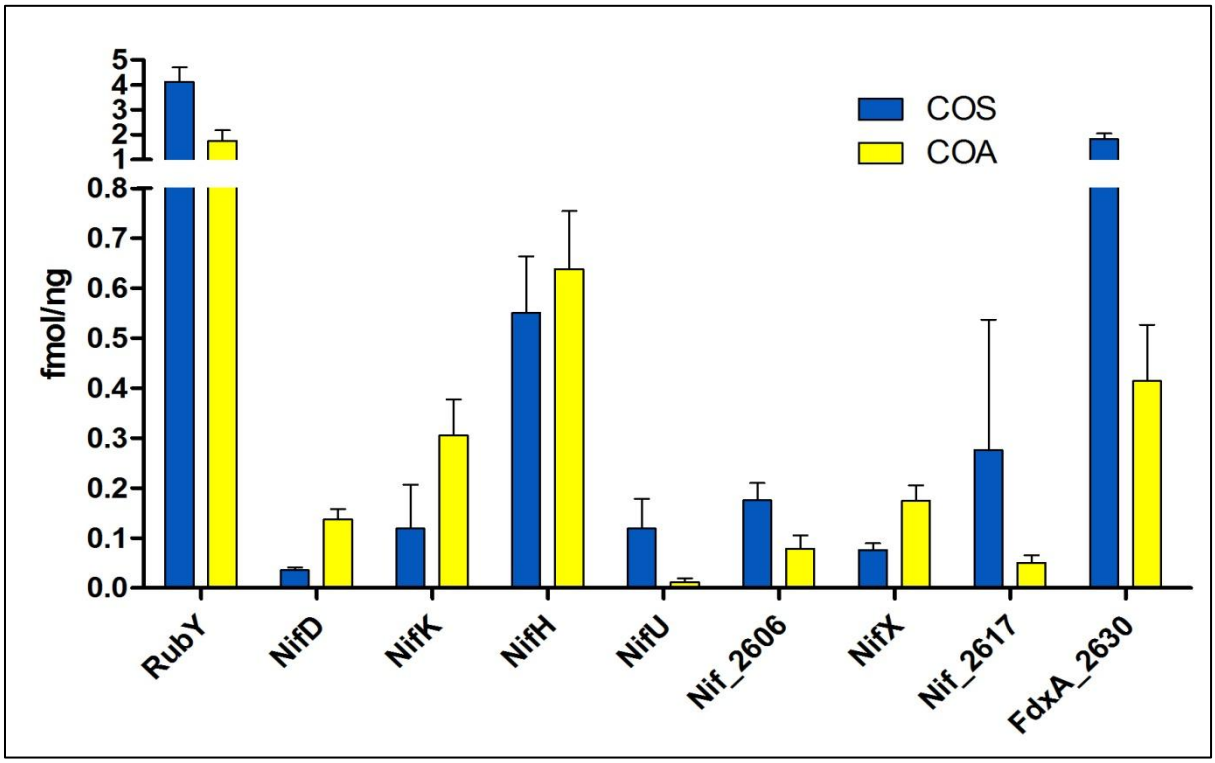

Figure 3.30. Identified nitrogen fixation genes in the *Codakia* symbiont proteome.
The data represent the mean ± standard deviation of three independent experiments each performed in three technical replicates measured with the Synapt G2 mass spectrometer. (blue - symbiont proteins of freshly collected *C. orbicularis* (COS), yellow - symbiont proteins of freshly collected *C. orbiculata* (COA)) (RubY - rubrerythrin, NifD - nitrogenase molybdenum-iron protein, NifD - nitrogenase molybdenum-iron protein subunit alpha, NifK - nitrogenase molybdenum-iron protein subunit beta, NifU - Fe-S cluster assembly protein NifU, Nif_2606 - Putative nitrogen fixation protein, NifX - nitrogen fixation protein, Nif_2617 - 4Fe-4S ferredoxin iron-sulfur binding domain-containing protein, FdxA_2630 - ferredoxin I)

In summary, in all proteomic investigations we were able to identifiy 26 Nif proteins in the *Codakia* symbiont (table 3.28, figur 3.30). Intrestingly the key genes NifK/D are more abundant in the *C. orbiculata* symbiont (figure 3.31).

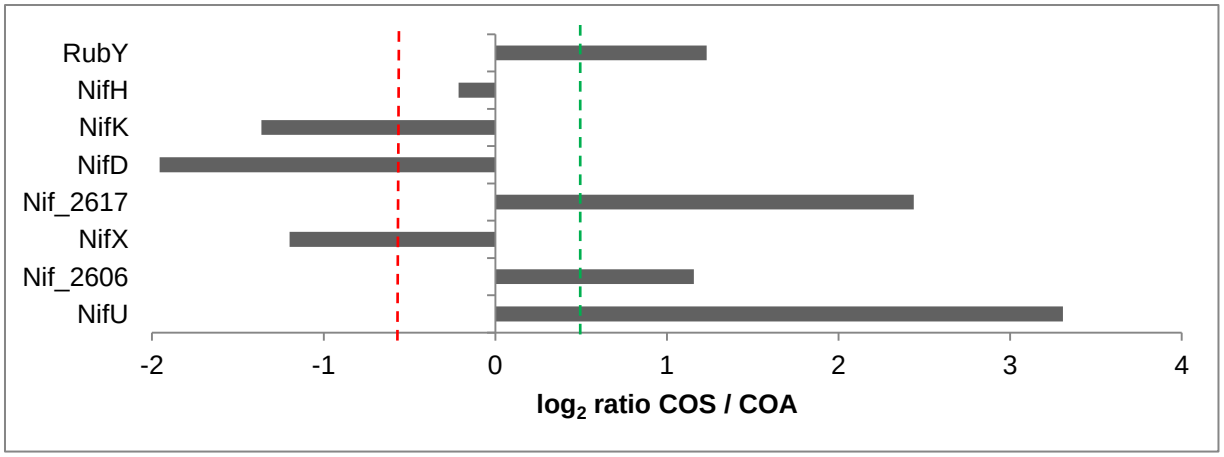

Figure 3.31. Comparison (log_2 ratio) of the identified nitrogen fixation proteins in *C. orbicularis and C. orbiculata* symbiont. Log_2 ratios of the identified nitrogen metabolism related proteins in gel free approach in both bivalves. Proteins with positive log_2 ratio are up-regulated in *C. orbicularis* (COS) symbiont, proteins with negative log_2 ratio are up-regulated in *C. orbiculata* (COA) symbiont. Ratios more than -0,5/0,5 indicate a significant change of the protein expression (green/red line).

Table 3.12. Putative nitrogen metabolism regulatory proteins in the *Codakia* symbiont genome.
(grey marked gene is co-localized in the symbiont genome)(Cello - subCELlular LOcalization prediction, CP - cytoplasm, PP - periplasm, IM - inner membrane, "Proteomics" - number of identifications in the symbiont proteome)

ID	Product	Label	IEP	MW (kDa)	Cello	TMH	Proteomics
995	ammonium transporter	AmtB	4.1	44	IM	11	0
996	nitrogen regulatory, protein P-II	GlnB/K	5	12.1	CP	0	6
1035	two-component response regulator, NtrC	NtrC	5.1	53.2	CP	0	5
1036	signal transduction histidine kinase, nitrogen specific, NtrB	NtrB	5.6	40.0	CP	0	4
3687	nitrogen regulation protein, NtrY	NtrY	9.3	81.5	IM	4	2
3688	nitrogen assimilation regulatory protein	NtrX	5.2	51.7	CP	0	2
4114	nitrogen regulatory protein, P-II family protein	GlnB	4.3	12.6	CP	0	0
4148	nitrogen regulatory protein, P-II	GlnB/K	5.3	12.2	CP	0	8

The nitrogenase is sensitive to oxygen furthermore the fixation is very energy expensive. Therefore a strict regulation of this pathway is assumed. In the genome seven putative nitrogen fixation metabolism regulatory genes are encoded (table 3.12). Three copies of the GlnB/K are encoded in the genome. One of this is co-located in the genome with an ammonium transporter (ID 995-996). A two component system nitrogen specific NtrC/B (ID 1035-1036) is encoded and co-localized in the genome. Furthermore proteins annotated as NtrY and NtrX are co-localized in the genome and less expressed in the proteome.
Two transporter for nitrate are encoded in the genome AmtB (ID 2000/995) (table 3.17). The ammonium transporter AmtB (ID995) and the nitrogen regulatory protein P-II are co-localized in the *C. orbicularis* symbiont genome. Two GnlB/K proteins were often identified in the proteome.

Table 3.13. Putative *rnf* gene cluster in *C. orbicularis* symbiont genome.
Protein identification in the membrane enrichment of the symbiont of *C. orbicularis* (COS) and *C. orbiculata* (COA) of the putative electron transport proteins (Rnf) to the nitrogenase. Genes organized in cluster are separated by double lines. (Cello - subCELlular LOcalization prediction, CP - cytoplasm, PP - periplasm, IM - inner membrane, "Proteomics" - number of identifications in the symbiont proteome)

ID	Label	IEP	MW (kDa)	Cello	TMH	COA 1D membrane	COS 1D membrane	Proteomics
1813	RnfB	4.2	30.8	CP	1			0
1815	RnfC	7.7	48.4	CP	0			0
1816	RnfD	5.6	37.2	IM	6			0
1817	RnfG	4.4	25.4	CP	1	X	X	2
1818	RnfE	5.7	23.3	IM	6			0
1819	RnfA	6.8	20.5	IM	6			0
Nif genes								
2637	RnfA	5	21.0	IM	6			0
2638	RnfB	4.6	18.6	CP	1	X		2
2641	RnfG	5.9	24.4	PP	1	X	X	2
2642	RnfE	9.8	24.5	IM	6	X	X	2
2463	RnfH	4.2	9.8	CP	0			0
2672	RnfA	9.5	20.8	IM	6			0
2673	RnfE	7.3	24.1	IM	5			0
2674	RnfG	8.7	21.6	PP	1			0
2675	RnfD	5.5	36.0	IM	7			0
2676	RnfC	9.5	49.0	CP	0			0
2678	RnfB	4.8	30.2	CP	1			0
4197	RnfE	6.6	24.7	IM	6			0
4198	RnfG	9.4	23.3	PP	1	X	X	3
4199	RnfD	9	38.6	IM	10			0
4201	RnfB	4.1	20.5	CP	1			0
4202	RnfA	5.1	20.7	IM	6			0

The *rnf* genes build a putative membrane bound complex to support the electron transport to the nitrogenase. The *rnf* encoding genes are shown in table 3.13. They are organized in gene clusters in the genome. The *rnf* gene cluster (ID 2637-2463) is closely located downstream to the 44 nitrogen fixation related gene cluster in the genome. In the proteome the Rnf proteins close to the nitrogen fixation cluster were most identified but not abundant. A lot of these Rnf proteins contain TMHs, which were only identified in the membrane protein enrichment.

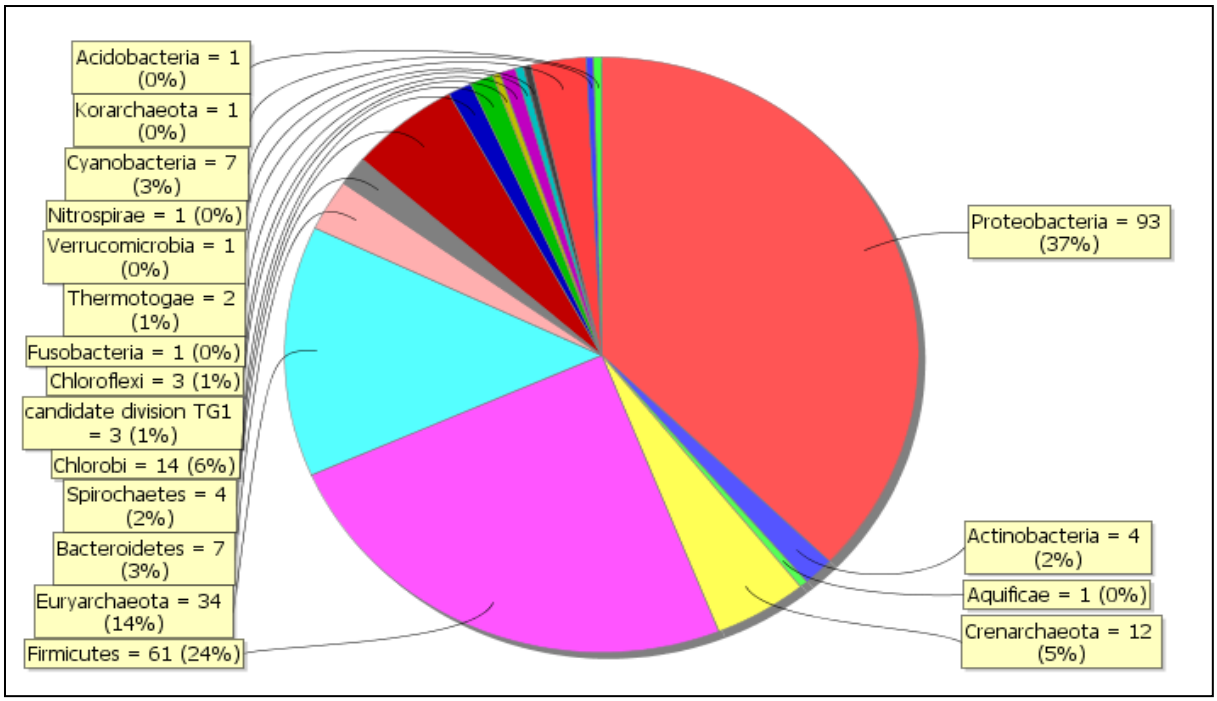

Figure 3.32. The phylogenetic affiliation, NCBI blast hits, of RubY encoded in *C. orbicularis*. (generated by Jcoast)(RubY - rubrerythrin ID 1096)

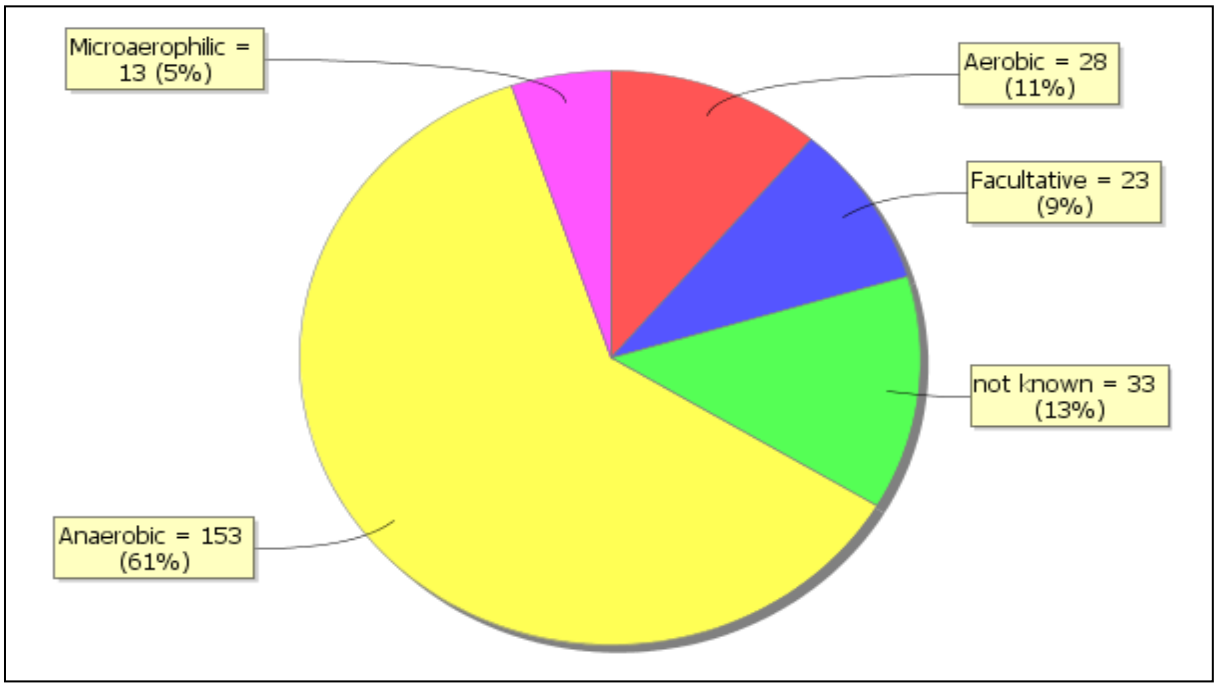

Figure 3.33. The NCBI blast hits of RubY encoded in *C. orbicularis* dependent on the oxygen requirements. (generated by Jcoast)(RubY - rubrerythrin ID 1096))

Rubrerythrin a putative nitrogenase protection protein was abundant identified in the proteome (table 3.29). Three rubrerythrin genes are encoded in the genome with no similarity to each other and only one rubrerythrin protein (ID 1096) could be identified in the proteome. The biggest amount of rubrerythrin was measured in the freshly collected *C. orbicularis* symbiont.

The phylogenetic species distribution of rubrerythrin is shown in figure 3.32, the main part belongs to the Gammaproteobacteria. The Cyanobacteria are also represented

but only with 3%. Figure 3.33 gives an overview about the distribution of rubrerythrin dependent on the oxygen requirements of the organisms. The biggest part of organisms which encode rubrerythrin live in an anaerobic habitat but microaerophile and aerobic habitats are also present. It seems rubrerythrin is phylogenetic and metabolic widespread which indicates a wide disturbed function.

3.2.3.4 Amino acid

The precursor of the amino acid metabolism are intermediate compounds of the glycolysis, the TCA-cycle and the PPP. In figure 3.34 important key enzymes of the amino acid metabolism are shown, in table 3.30 the gel free proteomic results are shown. All these key enzyme were abundant identified. Ammonium is assimilated via the glutamine synthetase (GlnA), these enzyme is one of the most abundant enzymes in the symbiont proteome. The high amount of this key enzymes is another indication that amino acids are one exchange compound between symbiont and host.

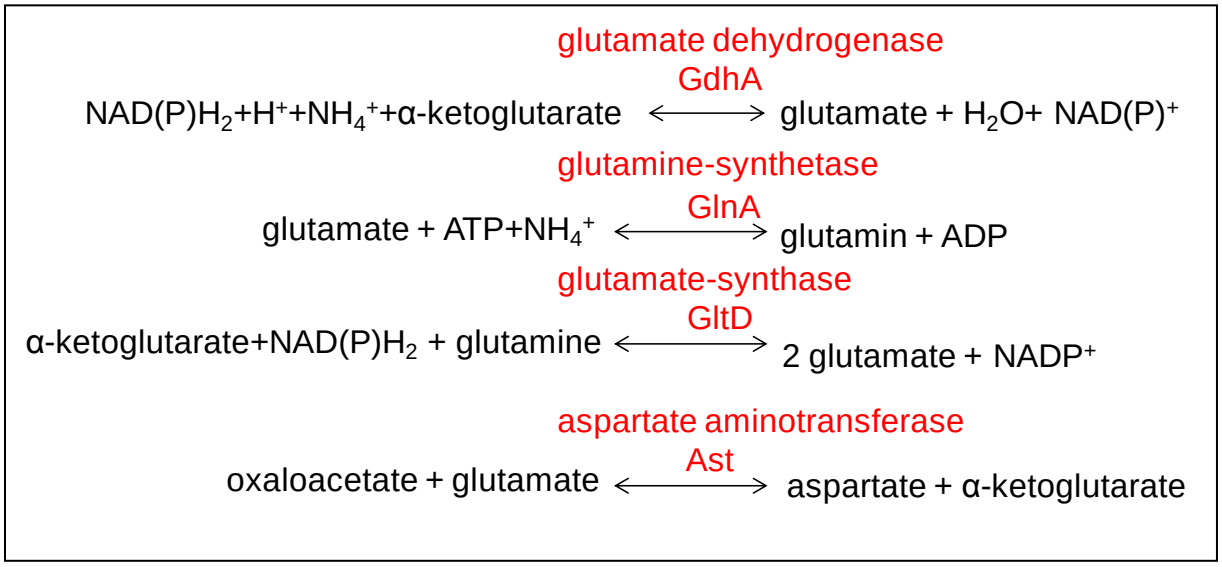

Figure 3.34. Key genes in amino acid metabolism encoded in the *Codakia* symbiont genome.
(red marked enzymes are identified in the proteome)

3.2.3.5 Stress-related proteins

In the proteome around 16 genes were found to be related to stress responses, among them are proteins for oxygen, starvation, temperature and copper stress (table 3.31). The oxygen stress related proteins, the superoxide dismutase (SodA) and the thiol peroxidase (Tpx) are induced after one week of incubation in oxygen saturated water whereas the catalse/peroxidase is not abundant expressed in any sample and not induced in the incubation sample. The SodA, Tpx and the cold shock protein (CspA) are the most abundant stress proteins in the proteome.

3.2.3.6 Motility / Chemotaxis

The genome encodes three gene cluster with in total 42 genes for flagella (table S6.7). In the proteome no indications for a flagellum was found under the given conditions. But it could be that the symbiont use these genes in a free living form to move inside the sediment. In the genome also five genes related to chemotaxis are encoded (CheY/Z/A/B/W), these genes are part of the motility gene clusters and were also not identified in the proteome. The presence of these genes indicates a free living form of the symbiont.

3.2.3.7 Environmental information processing

In the genome 554 genes are encoded for environmental information processing thereof 362 genes related to membrane transport, like ABC transporter, phosphotransferase systems (PTS), and protein secretion systems. In table S6.8 the ABC-transporter related proteins are shown. 162 genes are related to signal transduction like two component systems.

(1) TRAP - Tripartite ATP-independent periplasmic transporters

Several *TRAP*-type (TRipartite ATP-independent periplasmic) transporters are encoded in the genome (table 3.14). Four subunits of the TRAP transporter were identified in the proteome, all these proteins own no or one TMH. The other non identified protein encoded *TRAP* genes contain up to 13 TMHs.

Table 3.14. Encoded TRAP type proteins in the *Codakia* symbiont.
Genes organized in cluster are separated by double lines. (Cello - subCELlular LOcalization prediction, CP - cytoplasm, PP - periplasm, IM - inner membrane, "Proteomics" - number of identifications in the symbiont proteome) (grey marked proteins were identified in the proteome)

ID	Product	Label	IEP	MW (kDa)	Cello	TMH	Prote-omics
889	TRAP dicarboxylate transporter subunit DctM	DctM	5.9	48.7	IM	13	0
890	TRAP transporter, DctQ-like membrane protein	DctQ	7.7	42.9	IM	8	0
891	TRAP transporter subunit DctP	DctP	7.7	40.1	CP	0	9
1101	TRAP transporter solute receptor	DctP	9.4	39.6	PP	0	8
1991	TRAP dicarboxylate transporter- DctP subunit	DctP	9.9	37.3	IM	0	0
1993	TRAP-type C4-dicarboxylate transport system, small permease component	DctQ	10	18.5	IM	4	0
1994	TRAP-type C4-dicarboxylate transport system, large permease component	DctM	4.9	46.0	IM	12	0
2338	Tripartite ATP-independent periplasmic transporter DctQ component	DctQ	5.9	20.6	IM	4	0
2339	TRAP dicarboxylate transporter subunit DctM	DctM	4.7	48.9	IM	13	0

ID	Product	Label	IEP	MW (kDa)	Cello	TMH	Proteomics
2444	TRAP transporter solute receptor, TAXI family protein		9.8	34.8	PP	1	10
3002	TRAP transporter solute receptor, TAXI family precursor		9.3	34.2	CP	0	3
4663	C4-dicarboxylate transporter family protein, DctQ subunit	DctQ	9.5	23.9	IM	4	0
4664	TRAP dicarboxylate transporter subunit DctM	DctM	4.8	14.2	IM	3	0

(2) Type VI secretion system

In the *Codakia* symbiont genome a cluster of 15 type VI secretion system (*T6SS*) genes could be identified (table 3.15). Most of these genes are expressed and could be identified in the proteome.

Table 3.15. Type VI secretion system gene cluster in the *Codakia* symbiont genome.
Genes organized in cluster are separated by double lines. (Cello - subCELlular LOcalization prediction, CP - cytoplasm, PP - periplasm, OM - outer membrane, IM - inner membrane, "Proteomics" - number of identifications in the symbiont proteome)

ID	Product	strand	IEP	MW (kDa)	Cello	TMH	Proteomics
1840	type VI secretion system Vgr family protein	-	5.1	81.7	OM	0	2
1841	type VI secretion system protein	+	4.6	53.7	OM	0	0
1842	type VI secretion-associated protein, ImpA family	+	4	38.5	CP	0	2
1843	type VI secretion protein TssB1	+	4.4	19.3	CP	0	5
1844	type VI secretion protein EvpB	+	4.9	56.3	CP	0	7
1845	type VI secretion system protein	+	4.6	17.5	PP	0	8
1847	type VI secretion system lysozyme-related protein	+	5.8	19.5	CP	0	0
1848	type VI secretion protein	+	5.2	70.4	CP	0	0
1849	type VI secretion protein	+	5.9	39.3	CP	0	0
1850	type VI secretion system family ATPase, ClpV1	+	4.7	101	CP	0	4
1851	type VI secretion system, FHA-type	+	4.1	56.1	PP	0	0
1852	type VI secretion lipoprotein/ VasD	+	5.2	17.5	CP	1	5
1853	type VI secretion protein	+	4.8	50.1	CP	0	1
1854	type VI secretion system OmpA/MotB family protein	+	8.6	48.7	CP	1	3
1855	type VI secretion system protein ImpL	+	9	134	IM	3	6
1856	type VI secretion-associated protein	+	4.4	30.0	CP	0	1

(3) Sec/Tat-systems translocation

Sec-dependent and Tat-dependent (twin arginine translocation) pathways are responsible for the initial transport of proteins to the periplasma. The Tat-dep. secretion pathway allows the transport of folded proteins, whereas the Sec-dep. secretion pathway is only able to transport unfolded proteins. The proteins of the Tat-dep. pathway exhibit a N-terminal signal sequence with a highly conserved twin arginine motif, which is eponymous. In the *Codakia* symbiont genome several Sec-dep. and Tat-dep. pathway related genes are encoded, most of these genes were identified in the proteome (table 3.16).

Table 3.16. Sec/Tat- related secretion proteins encoded in the *Codakia* symbiont genome.
(Cello - subCELlular LOcalization prediction, CP - cytoplasm, PP - periplasm, IM - inner membrane, EC - extracellular, "Proteomics" - number of identifications in the symbiont proteome)

ID	Product	Label	IEP	MW (kDa)	Cello	TMH	Prote-omics
1009	bacterial protein export chaperone	SecB	4.4	17.7	CP	0	7
1633	preprotein translocase subunit	SecF	4.6	33.7	IM	6	8
1634	preprotein translocase subunit	SecD	5.1	66.2	IM	5	9
1635	preprotein translocase subunit	YajC	7.3	11.8	CP	1	9
1644	twin-arginine translocation pathway signal protein		4.4	55.5	EC	0	0
2028	twin-arginine protein translocation system	TatC	4.1	41.8	IM	5	1
2029	twin-arginine translocation protein	TatB	4.9	12.6	CP	1	1
2030	twin-arginine translocation protein	TatA	6.9	9.1	CP	1	7
2658	preprotein translocase SecG subunit	SecG	4.1	16.4	PP	2	4
3050	preprotein translocase subunit SecA	SecA	4.9	105.8	CP	0	8
3874	twin-arginine translocation pathway signal		7.7	48.0	PP	0	0
4069	preprotein translocase subunit SecE	SecE	9.4	13.6	IM	3	2
4102	preprotein translocase subunit SecY	SecY	10	48.7	IM	10	4

(4) Miscellaneous transporter

In *C. orbiculata* and *C. orbicularis* symbiont a "porine Gram-negative type" is encoded. Porine in Gram-negative bacteria are proteins which cross the membrane and act as a pore through which molecules can be transported by passive diffusion. The similarity of both genes is high ($9e^{-142}$, 65% identity). But it did not achieve the general protein criteria for the database creation so both proteins from both organism are in the database. Both proteins are abundant expressed in 2D gels and in the membrane enrichment samples (table 3.18). The proteomics results have to handle with care because of the high similarity of both proteins. Also in the endosymbiont of *Tevnia jerichonana* genome a PorT gene is encoded with less similarity to the *Codakia* symbiont PorT ($1.0e^{-30}$, 39% identity)(figure 3.35).

Furthermore transporter for phosphate and cobalamin could be identified in the proteome. In the genome a phosphotransferase system (PTS) is encoded but were not identified on the proteome level. Transporter for nitrate and ammonium are encoded in the genome but only low expressed under the given conditions. An overview about the important miscellaneous transporter in the *Codakia* symbiont is shown in table 3.17.

Table 3.17. Miscellaneous transporter encoded in the *Codakia* symbiont genome.
Genes organized in gene cluster are separated by double lines. (Cello - subCELlular LOcalization prediction, CP - cytoplasm, PP - periplasm, OM - outer membrane, IM - inner membrane, "Proteomics" - number of identifications in the symbiont proteome)

ID	Product	Label	IEP	MW (kDa)	Cello	TMH	Prote-omics
4756	porin, Gram-negative type	PorT	4	35.2	PP	1	8
910	porin, Gram-negative type	PorT	4.1	35.2	OM	1	7
	Nitrogen uptake						
2000	ammonium transporter	AmtB	5.1	43.7	IM	11	1
995	ammonium Transporter family protein	AmtB	4.1	44.2	IM	11	0
3287	nitrate transporter		5.2	50.4	IM	0	0
3288	nitrate transport ATP-binding, subunits C and D		6.1	31.8	CP	0	0
3289	nitrate ABC transporter permease protein		10	38.1	IM	6	1
3290	nitrate transporter	NasA	5.6	52.9	PP	1	9
	B12 uptake						
3672	vitamin B12 ABC transporter, B12-binding component BtuF	BtuF	9.6	30.6	PP	0	2
3673	TonB-dependent vitamin B12 receptor, BtuB	BtuB	4.5	72.4	CM	0	8
	Phosphate uptake						
3312	phosphate ABC transporter ATP-binding protein		5.2	30.4	CP	0	7
3313	phosphate ABC transporter ATPase		4.8	30.0	CP	0	5
3314	phosphate ABC transporter inner membrane subunit PstA	PstA	10	31.2	IM	6	2
3315	phosphate ABC transporter, inner membrane subunit PstC	PstC	4.4	34.0	IM	6	3
3316	phosphate uptake regulator PhoU	PhoJ	4.6	39.9	CP	0	10
3317	putative phosphate ABC transporter substrate binding protein		8.8	31.0	PP	0	8
	Miscellaneous						
2926	efflux transporter, RND family, MFP subunit		7.4	41.8	PP	0	9
3459	sodium/solute symporter family protein		8.4	63.7	IM	12	9
163	PEP-CTERM, putative polysaccharide export protein		4.8	19.9	PP	0	8
2925	transporter, AcrB/D/F family	Arc	7	149	IM	13	8
2928	Mg2+/Co2+ transporter	Cor	6.5	55.8	CP	2	8
3372	RND family efflux transporter MFP subunit		4.6	40.8	CP	1	8
909	major facilitator superfamily transporter		5.1	48.0	IM	12	3
	Phosphotransferase system (PTS)						

ID	Product	Label	IEP	MW (kDa)	Cello	TMH	Prote-omics
3129	HPr kinase		6.8	42.6	CP	0	1
3757	phosphotransferase system, phosphoenolpyruvate-dependent sugar	EIIA 2	4.3	17.2	CP	0	2
3758	HPr kinase/phosphorylase		5.3	35	CP	0	2
3759	putative nucleotide-binding protein		5.5	32	CP	0	2
3760	PTS system fructose subfamily IIA component	IIA	4.9	14.1	CP	0	0
3761	phosphotransferase system, phosphocarrier protein HPr	HPr	5.5	10.1	CP	0	0
3762	phosphoenolpyruvate-protein phosphotransferase		4.6	64	CP	0	3
4283	phosphoenolpyruvate-protein phosphotransferase PtsP	PtsP	5	82.9	CP	0	0

Table 3.18. NSAF quantification in the membrane enrichments of the "porine Gram-negative proteins".
(measured with the Orbitrap in the 1 D membrane enrichment approach) (Cello - subCELlular LOcalization prediction, PP - periplasm, OM - outer membrane, NSAF - normalized spectral abundance factor, Proteomics - number of identifications in the symbiont proteome)

							NSAF		
ID	source	Label	IEP	MW (kDa)	TMH	Cello	COA 1D membr.	COS 1D membr.	Proteomics
910	COS	PorT	4.1	35.0	1	OM	0.0053	0.0223	7
4756	COA	PorT	4	35.0	1	PP	0.1015	0.0671	8

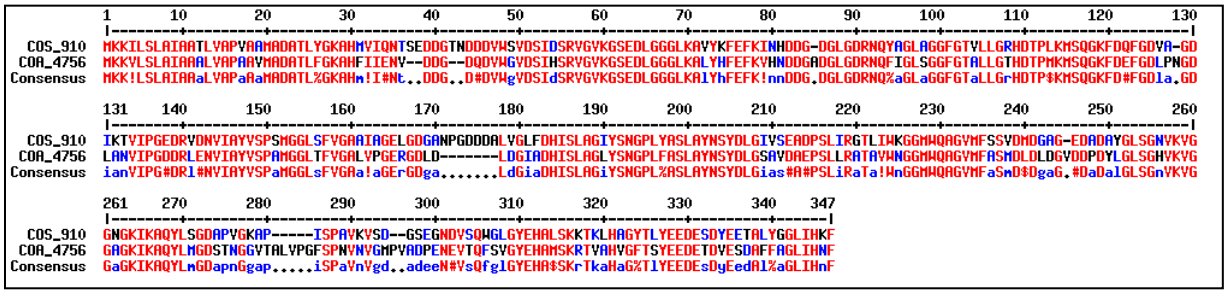

Figure 3.35. Alignment of the "porin Gram-negative type", PorT proteins.
(red - identical nucleotide, blue - different nucleotide, - missing nucleotide)

3.2.4 Symbiont proteome after one week of *C. orbicularis* incubation

The freshly collected bivalves could be incubated up to 6 months in sterile sea water tanks without organic particles and reduced sulfur compounds as food source. Under these conditions the bivalves together with the bacteria are starving. It has been shown that under these conditions the symbiont consume their stored elemental sulfur globules (Caro et al., 2009). After one week of starvation the bacterial separation is still quite clean (described above) and its long enough that the symbiont can adapt the metabolism to the new growth conditions. This implies that the bacteria

change the proteome and thus can handle the new growth conditions That is the reason we decided to use one week as maintenance duration of *C. orbicularis*. For the maintenance experiments only *C. orbicularis* was used because a single bivalve proteomics is possible.

The bivalve lives in burrows in the sea grass sediment at the interface between oxic and anoxic conditions. It is thinkable that the bivalve has to manage temporary low oxygen conditions and particle and sulfides starvation. To investigate how the symbiont handle temporary low oxygen conditions we incubated the bivalves under low oxygen (< 0.5 mg/l) (-[O_2]) and saturated oxygen (+[O_2]) conditions.

During the incubation it was visible that the bivalves are more "open", probably because of an increased water pumping. The bivalve lethality was under both conditions at the same level with around 3-5% per week.

In figure 3.36 the 1D gel of the membrane enrichment of the both oxygen conditions is shown, only low differences are visible. In table 3.19 the overview about the identified proteins in the *C. orbicularis* incubation with different oxygen concentrations is presented.

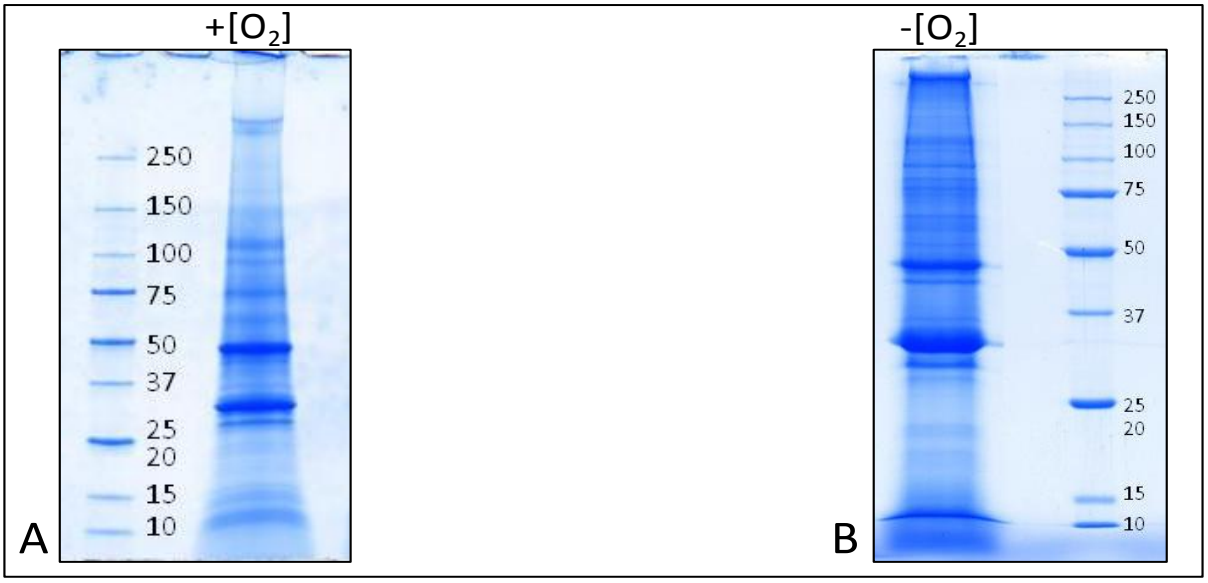

Figure 3.36. 1D SDS PAGE of membrane protein fraction of the *C. orbicularis* symbiont incubation with saturated (+[O_2]) and low oxygen (-[O_2]) conditions. A) one week of starvation with saturated oxygen concentration and B) one week of starvation with low oxygen concentration. The proteins were stained with Coomassie Brillant Blue. (marker in kDa)

Table 3.19. Overview about the sure identified proteins of the *C. orbicularis* symbiont after incubation. Therefore the bivalves were incubated in saturated oxygen concentration and low oxygen concentration for one week of sulfide starvation. The gel free approach was measured with the Synapt G2 and the 1D approach with a Orbitrap mass spectrometer. (COS - C. orbicularis)

	Gel free soluble	1D membrane
COS +[O_2]	221	707
COS -[O_2]	326	398

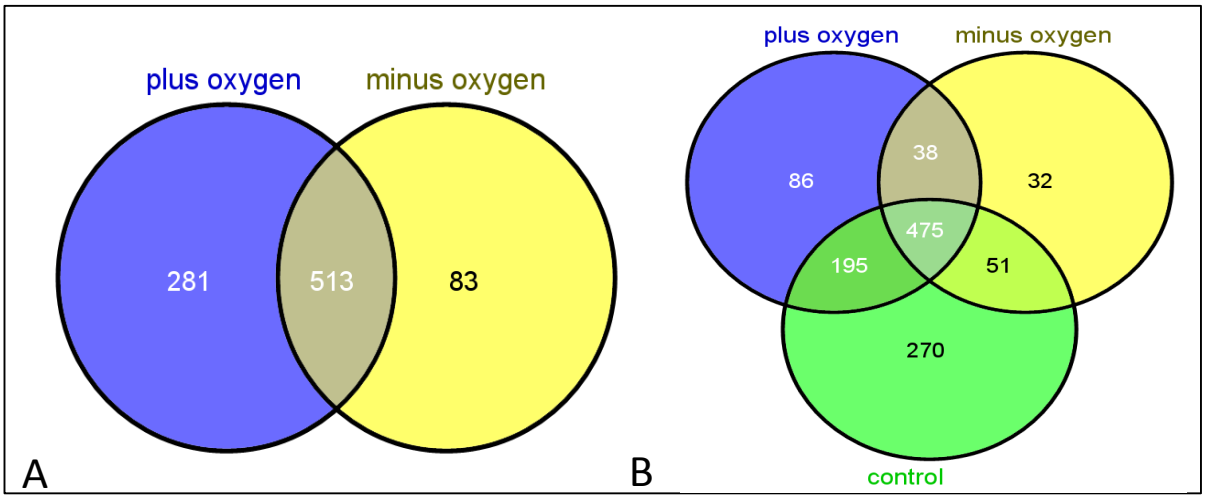

Figure 3.37. Venn diagram comparison of the identified symbiont proteins of the *C. orbicularis* incubation. A - Comparison of the proteome of one week of starvation with saturated (plus) and low (minus) oxygen concentrations. B - Comparison of symbionts extracted from freshly collected bivalves and extracted symbionts after one week of incubation of the bivalves under low and saturated oxygen conditions and sulfide starvation.

Venn diagrams reveal the overlap of proteins expressed under different conditions. Figure 3.37-A show that the main part of the identified proteins under both conditions is the comparable. With the help of Venn diagrams proteins can be identified which are either in one of both conditions or under both conditions expressed. It is also possible to compare more than two conditions with the Venn diagram. In figure 3.37 - B a Venn diagram comparison of the two different oxygen conditions and the control (freshly collected specimens) is shown. The identified proteins in the different Venn groups are provided in the electronic supplementary material in the exel file: Venn diagrams.

These analysis reveals that the phosphotransferase system, IIA-like nitrogen-regulatory protein PtsN (ID 3757), are exclusively present in the *C. orbicularis* symbiont under low oxygen conditions.

After one week of symbiont starvation under both oxygen conditions seven enzymes were exclusively identified related to the glycolysis and TCA-cycle pathways (-[O_2]: ID 112/1116/1245/1805, +[O_2]: ID 113/1775/3005) (table 3.21).

Interestingly is that four Sox system (sulfur oxidation enzymes) proteins (ID SoxZ 563, SoxY 564, SoxX 771,Sox(AX) 772) (table 3.25) were exclusively identified in the one week of starvation sample under both oxygen conditions. It is imaginable that the Sox system is related with the elemental sulfur globule utilization in case of reduced sulfur compounds starvation.

Five gaseous nitrogen fixation related genes were exclusively identified in the control sample.

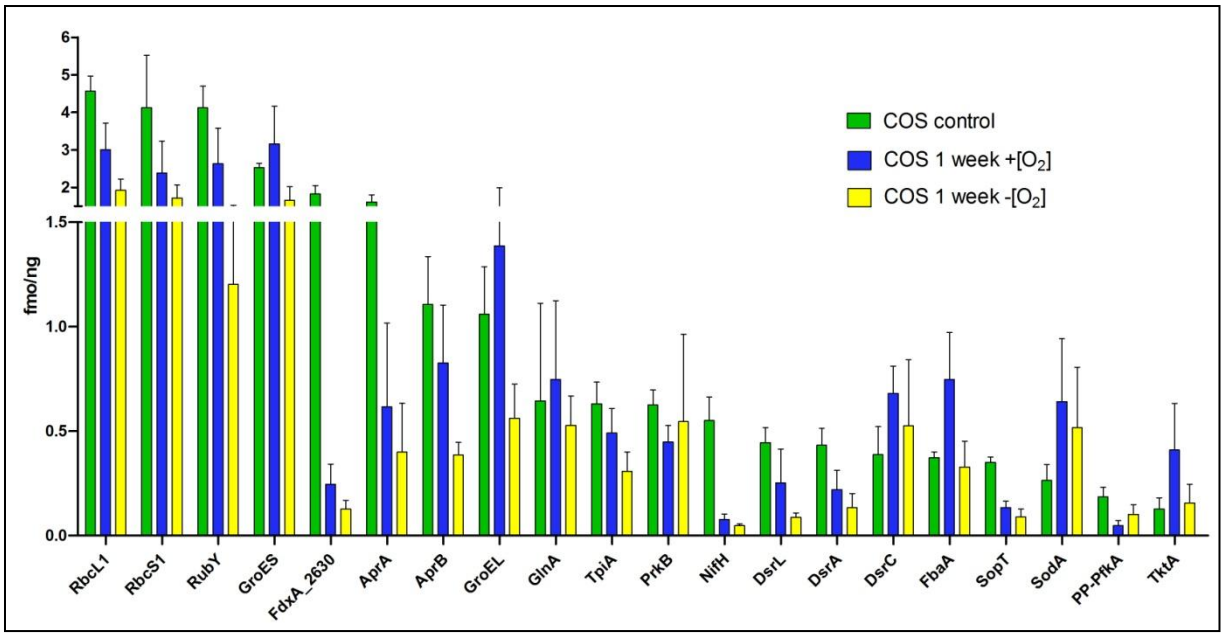

Figure 3.38. Comparison of the most abundant metabolic key proteins of *C. orbicularis* incubation symbiont. The data represent the mean ± standard deviation of three independent experiments each performed in three technical replicates measured with the Synapt G2 mass spectrometer. (from freshly collected *C. orbicularis* (control)(COS) and after one week of starvation with saturated (+[O_2]) and low oxygen (-[O_2]) concentration)(RbcL1-ribulose bisophosphate carboxylase large subunit, RbcS1- ribulose-1,5-bisphosphate carboxylase/oxygenase small subunit, RubY- rubrerythrin, GroEL/ES- chaperone, FdxA_2630- ferrodoxin, AprAB- adenylylsulfate reductase subunit alpha/beta, GlnA-glutamine synthetase, type I, TpiA- triosephosphate isomerase, PrkB- phosphoribulokinase, NifH- nitrogenase, Dsr- dissimilatory sulfite reductase and associated proteins, FbaA- fructose-1,6-bisphosphate aldolase, SopT- ATP sulfurylase, SodA-superoxide dismutase (Fe), PP-PfkA- PP dep. 6-phosphofructokinase, TktA- transketolase).

The 20 abundant key proteins in the control sample of the *C. orbicularis* symbiont are shown in figure 3.38. The two subunits of RuBisCO proteins form I are the most abundant expressed proteins. The sulfur oxidizing enzymes AprA and AprB are down-regulated under starvation conditions and at least expressed under sulfide starvation with low oxygen concentrations. The third one is rubrerythrin a putative nitrogenase protection protein, whereas the nitrogenase (NifH) is strict down-regulated under starvation conditions. The rubrerythrin is also down-regulated but much less strict regulated than the nitrogenase proteins.

Noteable is the high amount of GroEL/ES even under freshly conditions. Additional GroEL/ES is down-regulated but only under starvation conditions without oxygen, in starvation conditions with oxygen the chaperon GroEL/ES is up-regulated, the reasons are not known.

3.2.5 Protein functions in *C. orbicularis* incubation

3.2.5.1 Carbon metabolism

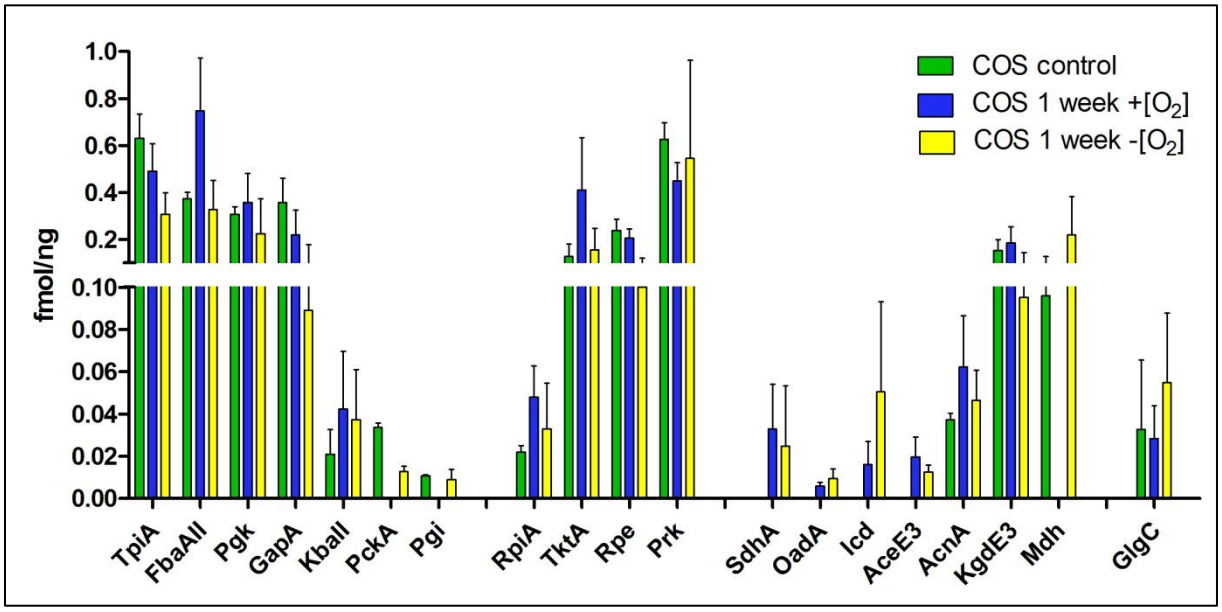

Figure 3.39. Identified carbon metabolism proteins in the *C. orbicularis* symbiont proteome.
The data represent the mean ± standard deviation of three independent experiments each performed in three technical replicates measured with the Synapt G2 mass spectrometer. The proteins are divided in pathway groups, glycolysis, PPP, TCA-cycle and glycogen metabolism. (TpiA - Triosephosphate isomerase, , FbaAII - fructose-1,6-bisphosphate aldolase, Pgk - phosphoglycerate kinase GapA - glyceraldehyde-3-phosphate dehydrogenase, type I, KbaII - ketose-bisphosphate aldolase, class-II,PckA - phosphoenolpyruvate carboxykinase, Pgi - glucose-6-P isomerase, , RpiA - ribose-5-P isomerase, TktA - transketolase, Rpe - ribulose-P 3-epimerase, Prk - phosphoribulokinase, SdhA - succinate DHG, OadAA - oxalacetate decarboxylase,Icd - isocitrate DHG, AceE3 - Pyruvat-DHG, AcnA - aconitate hydratase, KgdE3 - dihydrolipoamide DHG, MdhA - malate dehydrogenase, GlgC - Glucose-1-phosphate adenylyltransferase)

3.2.5.1.1 Carbohydrate metabolism

(1) Glycolysis/Gluconeogenesis / Tricarboxylic acid cycle (TCA-cycle) / reverse Tricarboxylic acid cycle (rTCA-cycle)

Only five enzymes related to the glycolysis, the triosephosphate isomerse (TpiA), the fructose-1,6-bisphosphat aldolase (FbaAII), pyruvate kinase (Pyk) and the glyceraldehyd-3-phosphat dehydrogenase (GapA), were identified in all incubation samples (figure 3.39).

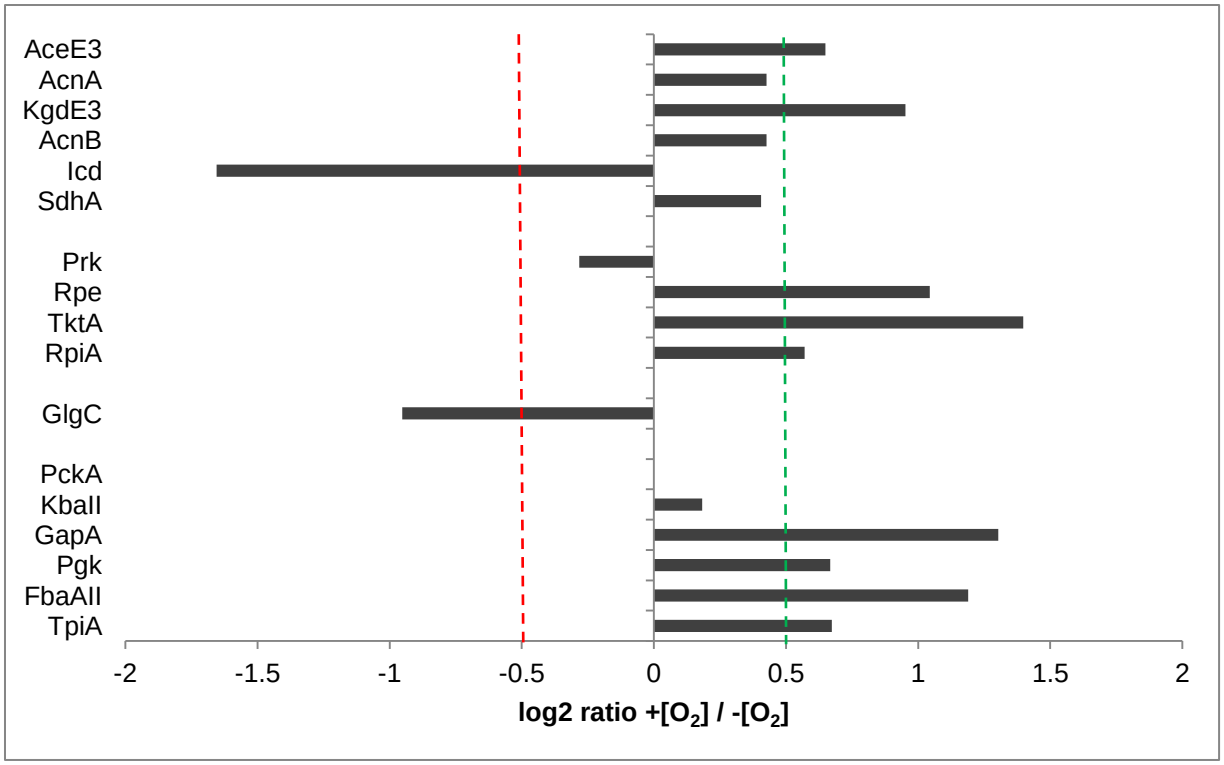

Figure 3.40. Comparison (log_2 ratio) of the identified carbon metabolism proteins in *C. orbicularis* symbiont incubation experiment (+[O_2]/-[O_2]). The proteins are divided in pathway groups, glycolysis, PPP, TCA-cycle and glycogen metabolism. Proteins with positive log_2 ratio are up-regulated in saturated oxygen concentrations, proteins with negative log_2 ratio are up-regulated in low oxygen conditions. Ratios more than -0,5/0,5 indicate a significant change of the protein expression (green/red line).

The comparison between the different oxygen concentrations showed a significant up-regulation of four (TpiA, FbaAII, PgkA, GapA) glycolysis proteins in saturated oxygen conditions (figure 3.40). Three proteins related to the PPP are also significant up-regulated under high oxygen conditions. Most of the TCA-cycle related enzymes were measured in both samples in the same amount only the Pyruvate-Dehydrogenase (KgdE3) is significant up-regulated. The glucose-1-phosphate adenylyltransferase (GlgC) is under low oxygen conditions highly up-regulated.

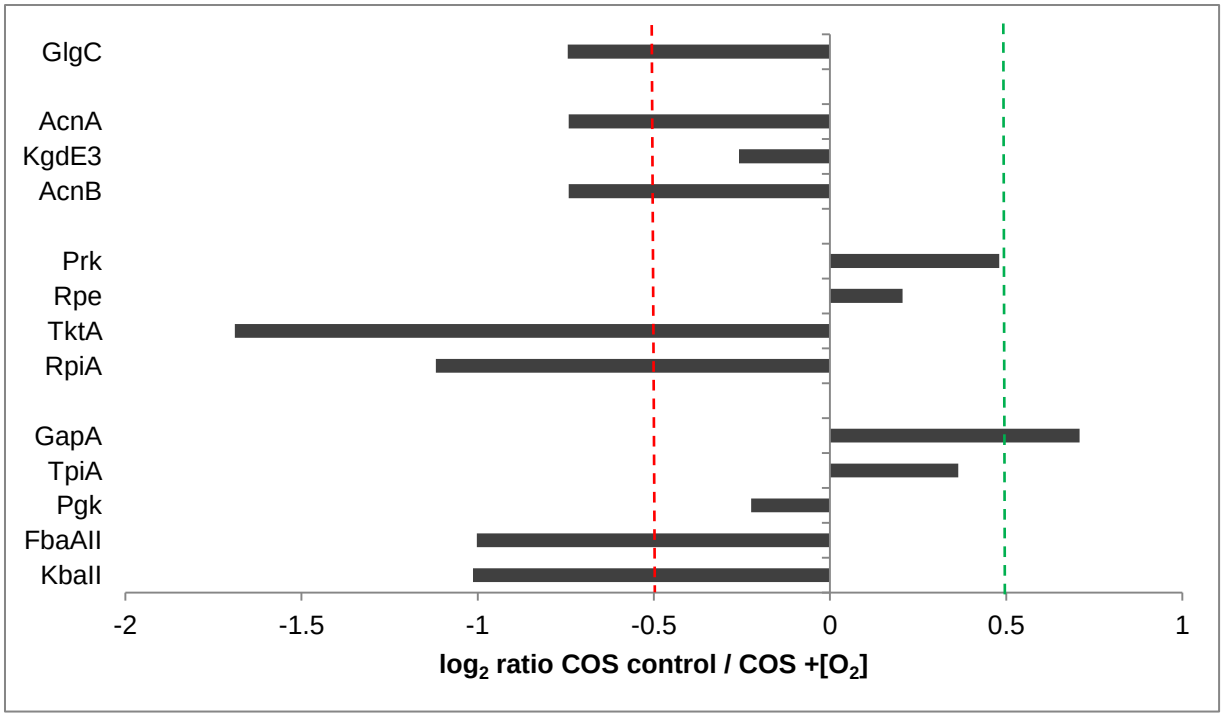

Figure 3.41. Comparison (log_2 ratio) of the identified carbon metabolism proteins in *C. orbicularis* symbiont control and incubation +[O_2] sample. The proteins are divided in pathway groups, glycolysis, PPP, TCA-cycle and glycogen metabolism. Proteins with positive log_2 ratio are up-regulated in the control sample, proteins with negative log_2 ratio are up-regulated in starvation with saturated oxygen concentrations. Ratios more than -0,5/0,5 indicate a significant change of the protein expression (green/red line).

Under reduced sulfur compounds starvation it seems that the TCA-cycle is more expressed (figure 3.34, 3.41). Unfortunately only less ratios could calculate of freshly collected and starvation samples to compare the TCA-cycle expression. But an indication is that several enzymes could only identified in the one week of starvation samples with saturated and low oxygen conditions.
The transketolase (TktA) and the ribose-5-phosphate isomerase (RpiA) are highly up-regulated in starvation oxygen saturated samples. Only the glyceraldehd-3-phosphate (GapA) is up-regulated in the control sample, this enzyme is aslo part of the suggested unconventional CBB.
The biggest amount of glucose-1-phosphate adenylyltransferase (GlgC) was found after one week starvation with low oxygen concentrations; the two other conditions were in the same expression level. Three proteins related to the glycogen metabolism (ID 39/41/49) were exclusively identified in the one week of starvation samples under both oxygen conditions. The glycogen degradation enzymes are really less expressed also under starvation and low oxygen conditions.

3.2.5.1.2 Calvin Benson Bassham Cycle (CBB)

The RuBisCO subunits are in all samples abundant expressed and measured in a high amount (figure 3.42). The activator proteins CbbQ and CbbO were identified but in a really low amount, therefore no ratios were calculated.

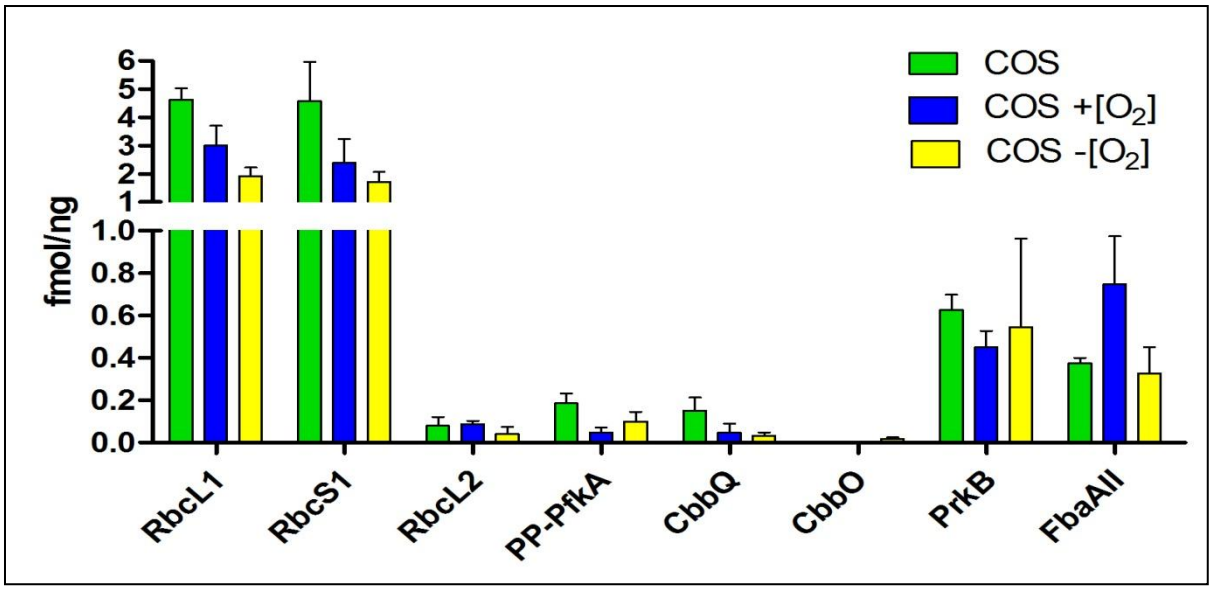

Figure 3.42. Identified CBB proteins in the *C. orbicularis* symbiont proteome.
The data represent the mean ± standard deviation of three independent experiments each performed in three technical replicates measured with the Synapt G2 mass spectrometer. (Rbcl1 - ribulose bisophosphate carboxylase, RbcS1 - ribulose-1,5-bisphosphate carboxylase/oxygenase small subunit, RbcL2 - ribulose 1,5-bisphosphate carboxylase large subunit, CbbQ/CbbO - RuBisCO activation protein, PP-PfkA - 6-phosphofructokinase, PrkB - phosphoribulokinase, FbaAII - fructose-1,6-bisphosphate aldolase, class II)(COS - *C. orbicularis*)

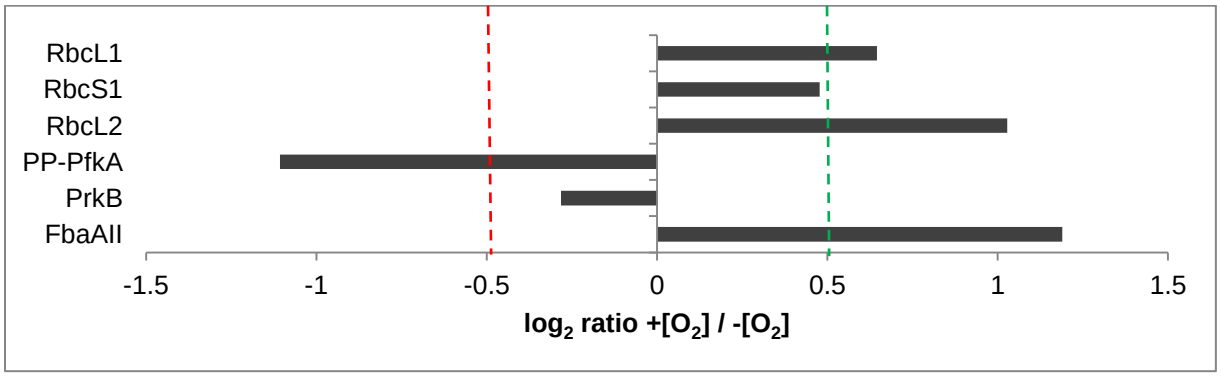

Figure 3.43. Comparison (log_2 ratio) of the identified CBB proteins in *C. orbicularis* symbiont incubation experiment (+[O_2]/-[O_2]). Proteins with positive log_2 ratio are up-regulated in saturated oxygen concentrations, proteins with negative log_2 ratio are up-regulated in low oxygen conditions. Ratios more than -0,5/0,5 indicate a significant change of the protein expression (green/red line).

Two proteins are significant up-regulated (FbaAII, RbcL1/2) under saturated oxygen conditions, but the RbcL2 were measured in a very low amount (figure 3.43). The PP-PfkA is significant up-regulated under low oxygen conditions.

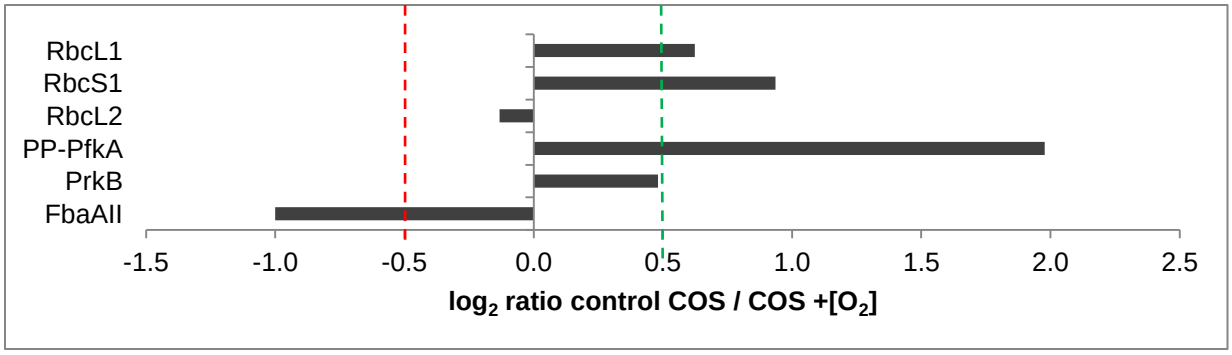

Figure 3.44. Comparison (log_2 ratio) of the identified CBB proteins in *C. orbicularis* symbiont control and incubation +[O_2] sample. Proteins with positive log_2 ratio are up-regulated in the control sample, proteins with negative log_2 ratio are up-regulated in starvation conditions with saturated oxygen concentrations. Ratios more than -0,5/0,5 indicate a significant change of the protein expression (green/red line).

The carbon dioxide fixation is down-regulated in the sulfide starvation sample; only the fructose-1,6-bisphosphate aldolase, class II (FbaAII) is up-regulated (figure 3.44). The carbon fixation is an energy consuming pathway, probably because of that it is down-regulated if the symbiont have no reduced sulfur compounds for energy generation.

3.2.5.2 Energy metabolism

3.2.5.2.1 Oxidative phosphorylation

The expression of the ATP-synthase is down-regulated under starvation conditions in comparison to the control conditions (table 3.24). Interestingly, only one subunit is up-regulated in the starvation sample with saturated oxygen conditions. The proteomic data indicate that the low oxygen conditions (0.5 mg/l) are buffered by the bivalve trought increased pumping of water.

3.2.5.2.2 Sulfur metabolism

After one week of reduced sulfur starvation the amount of sulfur oxidizing enzymes decrease. The lowest sulfur oxidizing enzymes amount were measured in one week of starvation with low oxygen conditions. In summary 19 sulfur oxidizing enzymes were identified (figure 3.45), this reflect the importance of this pathway. The SopT/Sat, AprAB and the Dsr pathway were identified in a high amount.

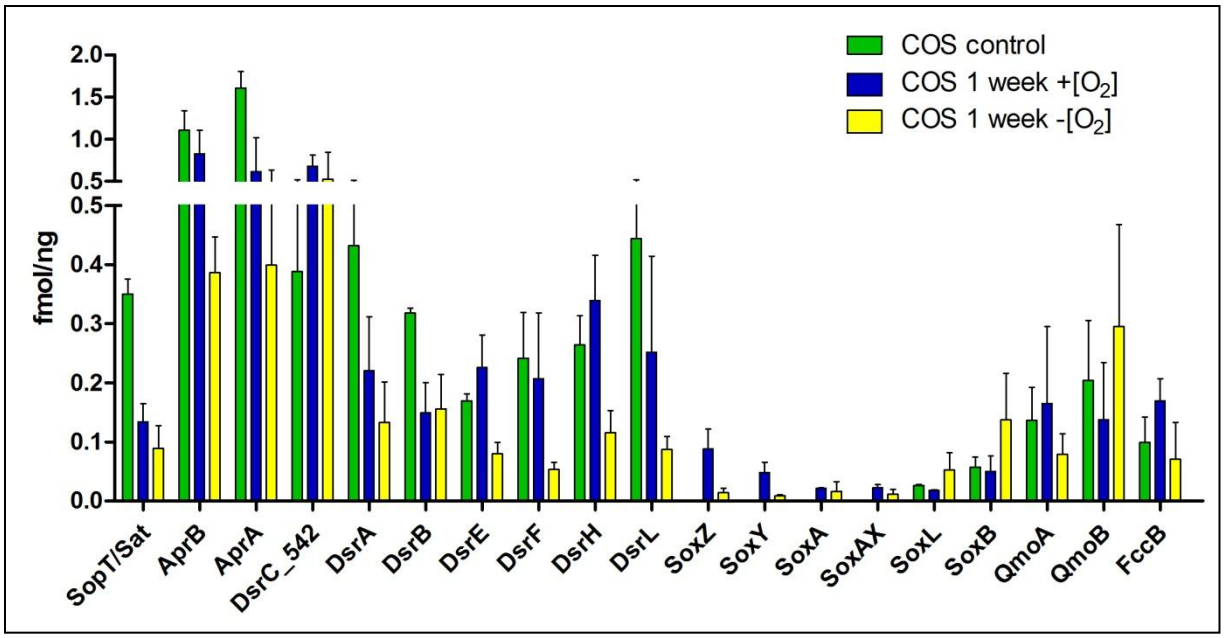

Figure 3.45. Overview about the identified sulfur metabolism related proteins in incubated *C. orbicularis* symbionts. The data represent the mean ± standard deviation of three independent experiments each performed in three technical replicates measured with the Synapt G2 mass spectrometer. (SopT - ATP sulfurylase, AprAB - adenylylsulfate reductase, Dsr -dissimilatory sulfite reductase and associated proteins, FccAB - sulfide dehydrogenase, QmoAB - heterosulfide reducatse, putative QmoAB, Sox - sulfur oxidation proteins, Sqr - sulfide quinone reductase)

The Sox related proteins were only in a very low amount measured and up-regulated under starvation conditions. Four Sox, SoxZYX(AX), proteins were only identified after one week of starvation under both oxygen conditions. Noticeable is that the SoxZ and SoxY proteins were significant up-regulated under oxygen saturated conditions, while SoxL and SoxB are up-regulated in low oxygen conditions (figure 3.46). The symbiont store elemental sulfur, it is imaginable that the Sox system is involved by the degradation of sulfur globules in starvation conditions

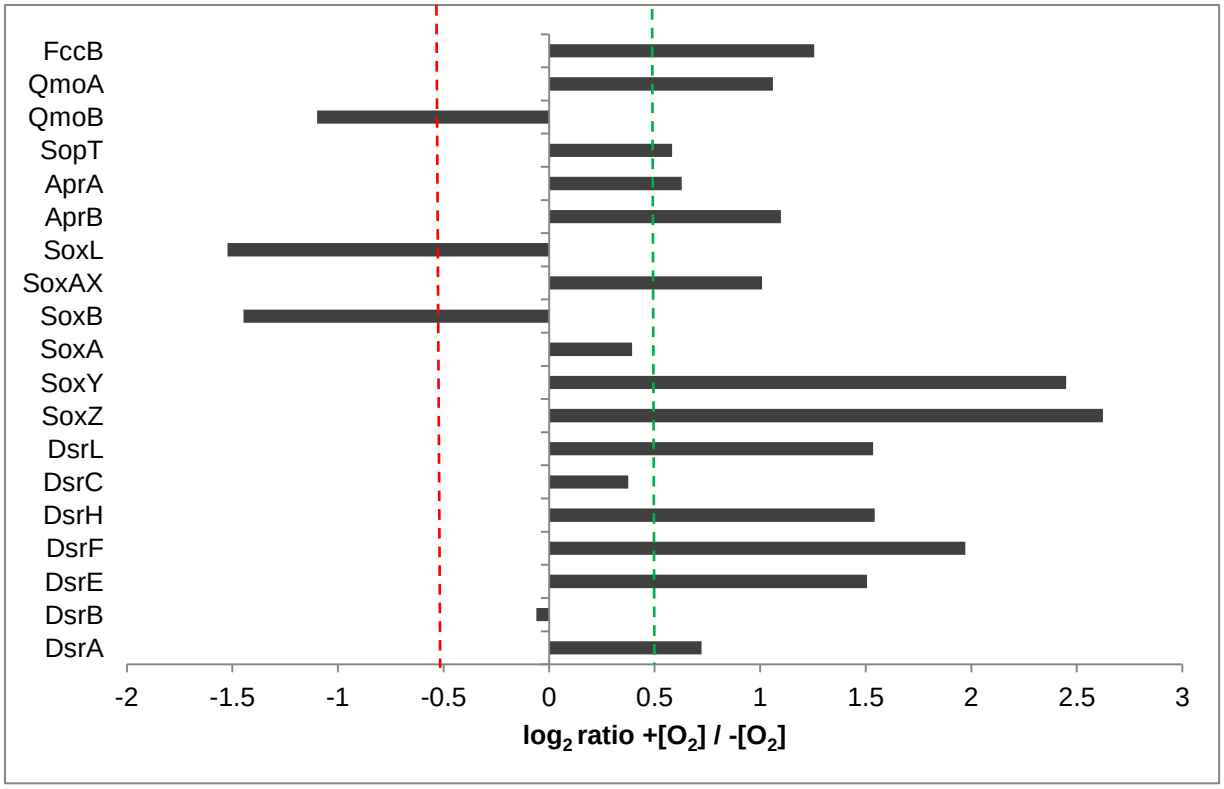

Figure 3.46. Comparison (log_2 ratio) of the identified sulfur metabolism proteins in *C. orbicularis* symbiont incubation experiment (+[O_2]/-[O_2]). Proteins with positive log_2 ratio are up-regulated in saturated oxygen concentrations, proteins with negative log_2 ratio are up-regulated in low oxygen conditions. Ratios more than -0,5/0,5 indicate a significant change of the protein expression (green/red line).

Most of the sulfur metabolism related proteins are significant up-regulated under saturated oxygen conditions (figure 3.46).
The symbiont use oxygen as terminal electron acceptor, therefore its logical that the sulfur oxidizing proteins are up-regulated under high oxygen conditions.

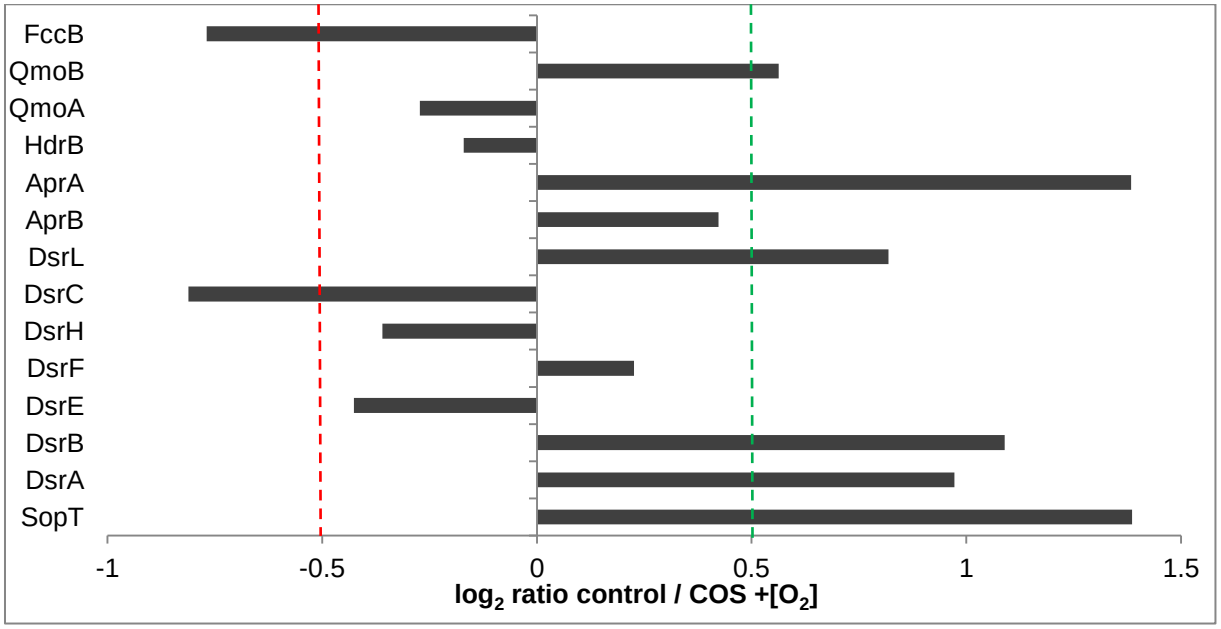

Figure 3.47. Comparison (log_2 ratio) of the identified sulfur metabolism proteins in *C. orbicularis* symbiont control and incubation +[O_2] sample. Proteins with positive log_2 ratio are up-regulated in the control sample, proteins with negative log_2 ratio are up-regulated in starvation conditions with saturated oxygen concentrations. Ratios more than -0,5/0,5 indicate a significant change of the protein expression (green/red line).

Most of the sulfur metabolism proteins are more expressed in the control sample. Only two proteins, DsrC and FccB are significant up-regulated in the starvation sample (figure 3.47). In total the sulfur metabolism is significant down-regulated under sulfide starvation conditions but the proteins are still present in a quite high amount. This indicate that after one week of sulfide starvation the symbiont still oxidize sulfide from the elemental sulfur storage inside the periplasmic space.

3.2.5.3 Nitrogen fixation

The gaseous nitrogen fixation is energy expensive so it is obvious that this pathway is strictly regulated. The proteome data show that this pathway is under starvation conditions independent of the oxygen conditions strong down-regulated (figure 3.48).

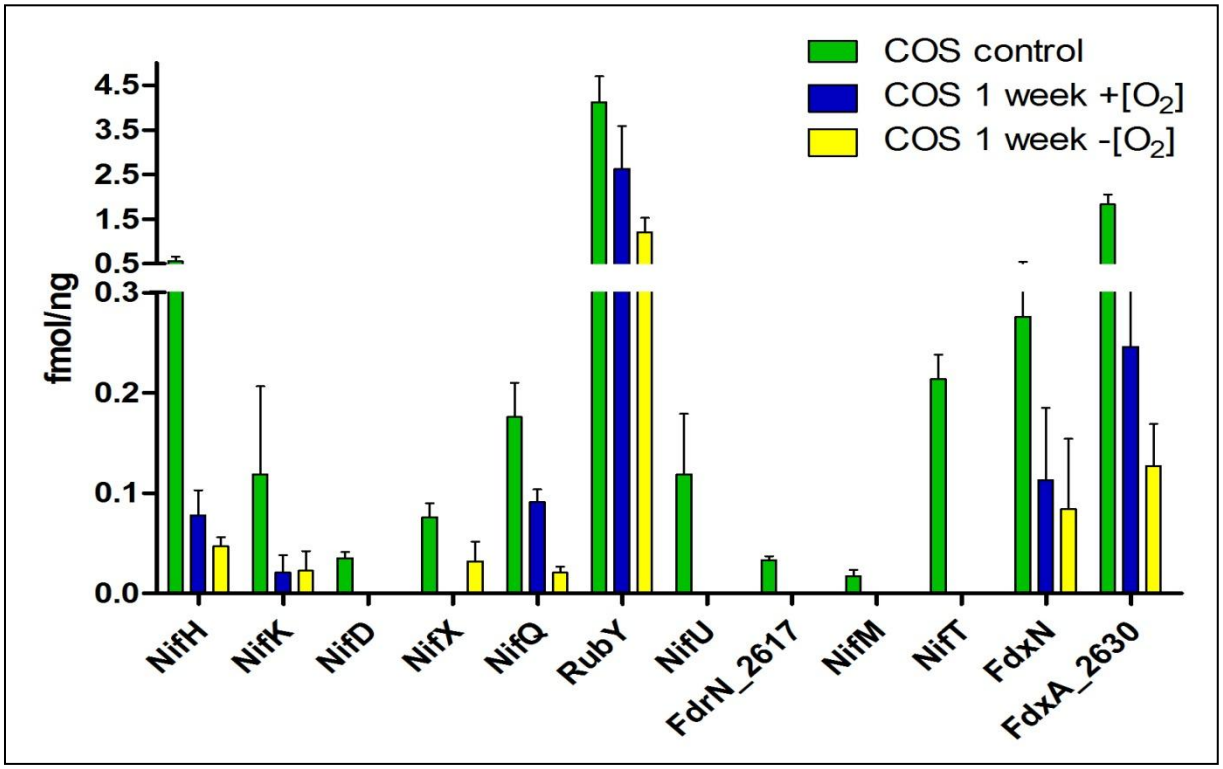

Figure 3.48. Identified carbon nitrogen fixation proteins in the *C. orbicularis* symbiont proteome.
The data represent the mean ± standard deviation of three independent experiments each performed in three technical replicates measured with the Synapt G2 mass spectrometer. (RubY - rubrerythrin, NifH - nitrogenase molybdenum-iron protein, NifD - nitrogenase molybdenum-iron protein subunit alpha, NifK - nitrogenase molybdenum-iron protein subunit beta, NifU - Fe-S cluster assembly protein, Nif_2606 - Putative nitrogen fixation protein, NifX - nitrogen fixation protein, Nif_2617 - 4Fe-4S ferredoxin iron-sulfur binding domain-containing protein, NifQ - nitrogenase FeMo cofactor synthesis molbdenum delivery protein, NifM - nitrogen fixation protein, FdxN - ferrodoxin, NifT - nitrogen fixation protein, FdxA_2630 - ferrodoxin I) (COS - *C. orbicularis*)

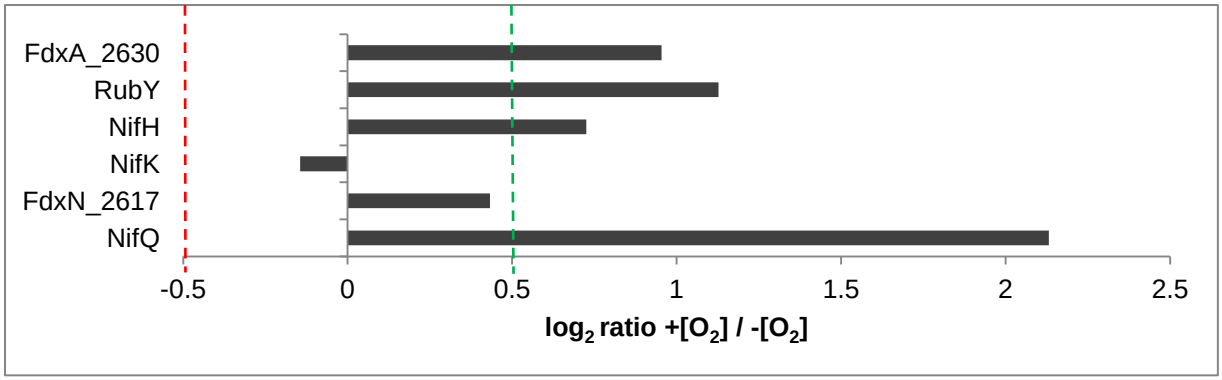

Figure 3.49. Comparison (log_2 ratio) of the identified nitrogen fixation proteins in *C. orbicularis* symbiont incubation experiment (+[O_2]/-[O_2]). Proteins with positive log_2 ratio are up-regulated in saturated oxygen concentrations, proteins with negative log_2 ratio are up-regulated in low oxygen conditions. Ratios more than -0,5/0,5 indicate a significant change of the protein expression (green/red line).

The nitrogen fixation proteins are significant up-regulated after one week of starvation under saturated oxygen concentration (figure 3.49). Under low oxygen and starvation conditions RubY is up-regulated.

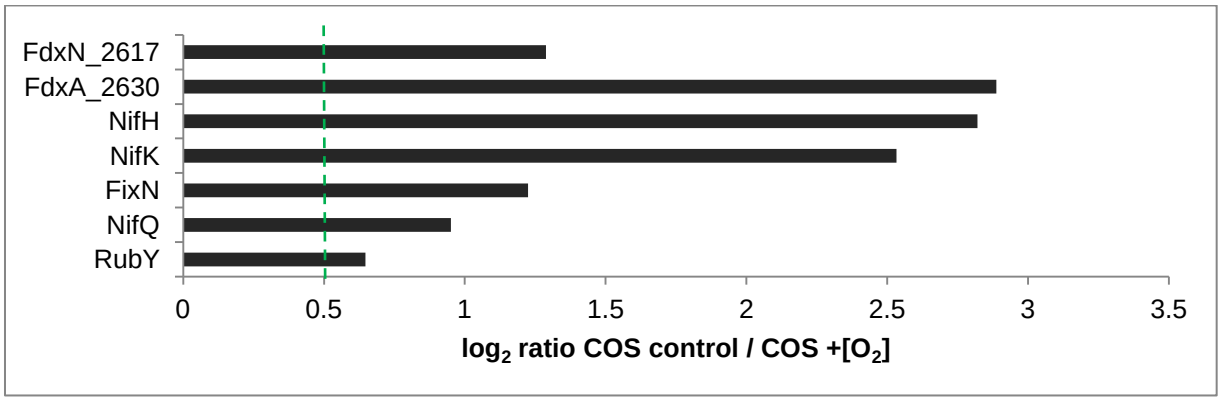

Figure 3.50. Comparison (log_2 ratio) of the identified nitrogen fixation proteins in *C. orbicularis* symbiont control and incubation +[O_2] sample. Proteins with positive log_2 ratio are up-regulated in the control sample, proteins with negative log_2 ratio are up-regulated in starvation conditions with saturated oxygen concentrations. Ratios more than -0,5/0,5 indicate a significant change of the protein expression (green line).

All genes related to the gaseous nitrogen fixation are much more expressed in the freshly collected *C. orbicularis* symbiont than in the starvation sample (figure 3.50). Two nitrogen fixation proteins (NifH/NifK/D) of them are around 100 times more expressed, it seems that the Nif gene expression is regulated very strict. One reason for that is that the gaseous nitrogen fixation is an energy expensive pathway. The rubrerythrin expression follows the expression schema of the nitrogenase but seems to be not so strict regulated like the nitrogenase.

3.2.5.4 Stress-related proteins

The stress related proteins are shown in table 3.31 and figure 3.51. Interestingly three of four stress related proteins the most amount were measured in the one week

of starvation sample with saturated oxygen conditions. The highest amount of rubrerythrin was measured in the control sample. This could indicate that this protein not function as a normal stress protein. Because it is assumed that the stress level after one week starvation in the lab is higher than in freshly collected bivalves. This indicate also the other compared stress proteins, their amount increase during starvation.

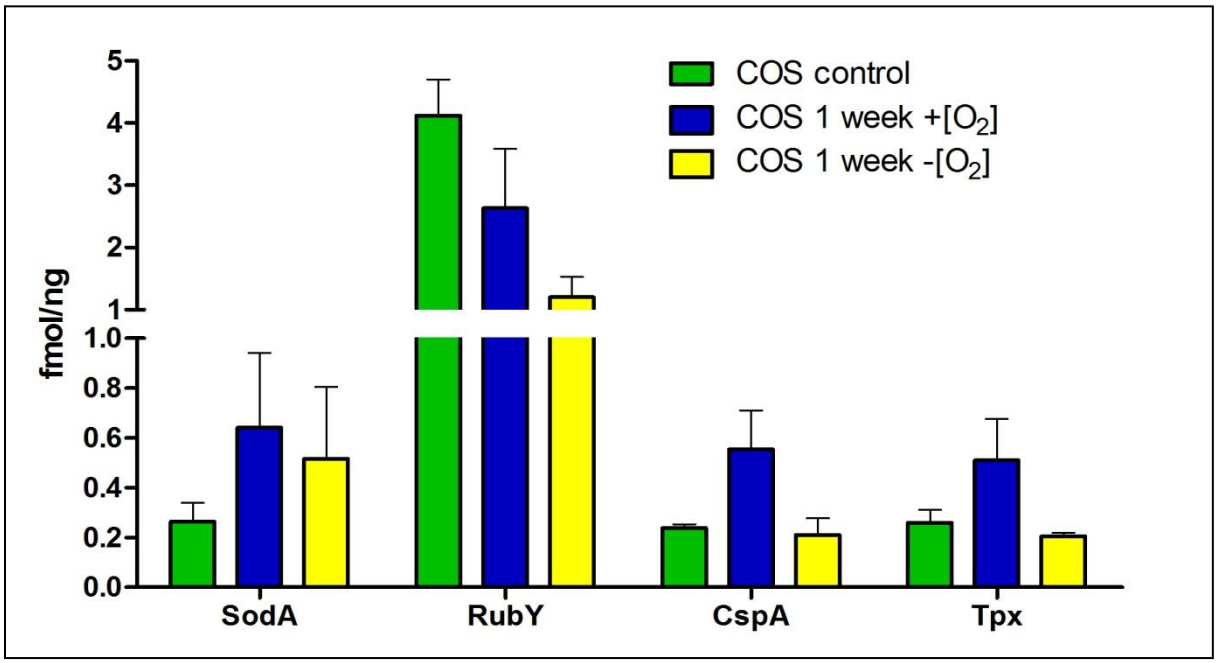

Figure 3.51. Identified stress related proteins in the *C. orbicularis* symbiont proteome.
The data represent the mean ± standard deviation of three independent experiments each performed in three technical replicates measured with the Synapt G2 mass spectrometer. (SodA - superoxide dismutase (Fe), RubY - rubrerythrin, CspA - cold shock protein, Tpx - peroxidase) (COS - C. orbicularis)

3.2.6 Proteome of only symbiont incubation for 24 h

The symbiont is until now not cultivable without the host. However the detection of a free living stock of bacteria in the sea grass sediment proof that the Gammaproteobacteria are not obligate symbionts. In the lab it is possible to keep the bacteria alive for up to two days but not longer (personal communication O. Gros). Because we think that the bivalve buffers the chemical conditions we decided to maintain only the extracted bacteria for 24 hours under saturated oxygen and low oxygen concentrations with added nitrate. The physiological adaption of the symbiont proteome to the new living conditions should be finished after 24 h. The two 1D gels show a different distribution of the protein bands under the different oxygen concentrations in symbiont 24 h maintenance. For example, a protein band with the size of 60 kDa is visible in saturated oxygen conditions and decreased in the low oxygen conditions (figure 3.52- A/B). In comparison to freshly collected *Codakia* symbionts the protein line contain less protein bands.

We were able to identify in the symbiont sample with saturated oxygen concentrations 282 proteins and in the sample with low oxygen concentrations 356 proteins.
We were not able to cultivate the bacteria longer than 2-3 days that means the bacteria were not satisfied under the used lab conditions. It is an extreme stress situation for the bacteria, consideration must be given to this fact for the data interpretation.

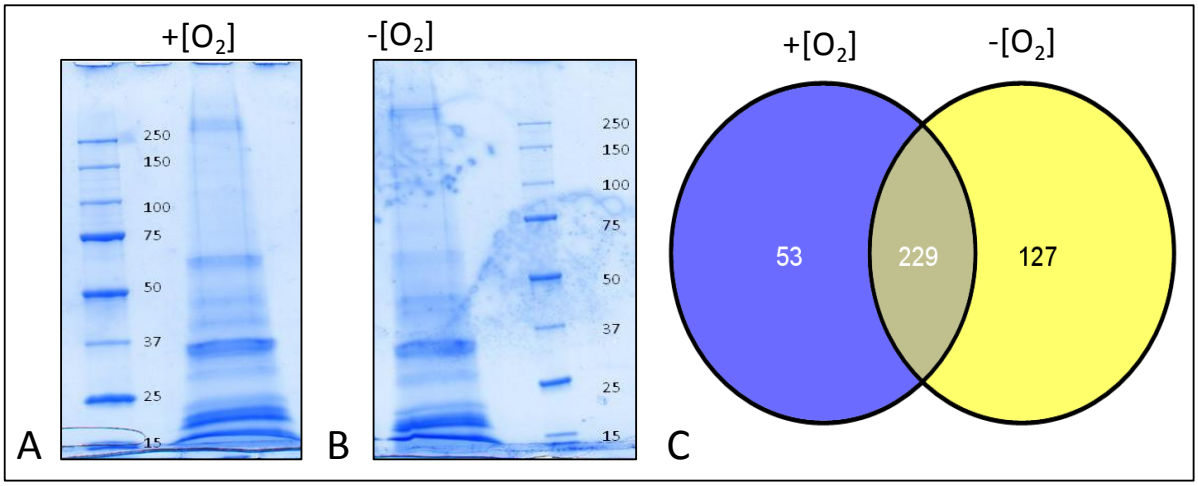

Figure 3.52. Incubation of isolated bacterial *C. orbicularis* symbiont outside of the host for 24 hours.
A/B 1D SDS PAGE of soluble symbiont protein fraction after 24 h incubation with (A) saturated oxygen concentrations (+[O_2]) and (B) with low oxygen concentrations (-[O_2]), stained with Coomassie Brillant Blue (marker in kDa) (C) - Venn diagram comparison of the sure identified proteins in the symbiont in bothe oxygen conditions.

The Venn diagram comparison shows that the main part of the identified proteins were found under both oxygen conditions (figure 3.52-C). In the low oxygen sample several energy metabolism related proteins were exclusively identified, like several NADH-quinon oxidoreductase proteins but also unexpectly ATP-synthase subunits. Two proteins, NapC and NapA, related to the dissimilatory nitrate reduction were only identified under low oxygen conditions but these proteins were also identified in freshly conditions. A table of the Venn group proteins is provided in the electronic supplementary material (exel file: "Venn diagrams").

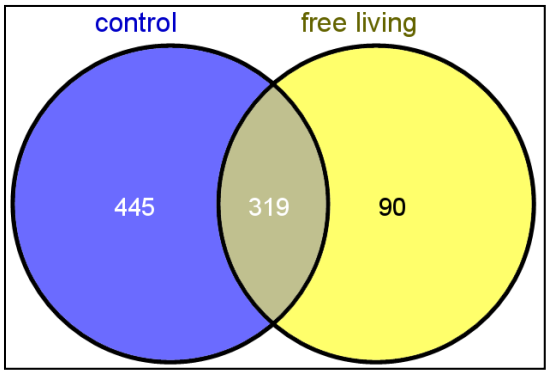

Figure 3.53. Venn diagram comparison of the soluble proteome of *C. orbicularis* symbiont fraction from freshly collected bivalves and "free living" symbiont incubation (+[O_2]).

A comparison between the 24 hours maintenance of only the symbiont without the bivalve and the freshly collected sample of the *C. orbicularis* symbiont shows that only 90 proteins were exclusively identified in the "free living" symbiont sample (figure 3.53). Interestingly, 24 proteins related to membrane export and only five stress related proteins were only identified in the only symbiont incubation. Interestingly several TonB related proteins were exclusively identified in the "free living" sample It was expected that stress response related proteins are more abundant by the maintenance of only the bacteria. In summary, 319 proteins were identified in control and incubation of only bacteria. The big amount of the proteins only identified in the control sample is not remarkable because for this control sample much more mass spectroscopy measurements were done. In this data no clear indications to low oxygen conditions or free living form proteins were not detected. The reasons for that are dubious and further analysis are necessary.

3.2.7 Supplementary material from the gel free analysis

Table 3.20. Identified glycolysis and related enzymes in *C. orbicularis* and *C. orbiculata* symbiont proteome.
C. orbicularis bivalves were incubated for one week of starvation in sterile sea water with saturated oxygen (+[O_2]) and low oxygen (-[O_2]) conditions. The data represent the mean ± standard deviation of three independent experiments each performed in three technical replicates measured with the Synapt G2 mass spectrometer. The protein amount is measured in fmol/ng. Ratios more than -0,5/0,5 indicate a significant change of the protein expression. Proteins with significant positive log_2 ratio are up-regulated (green), proteins with significant negative log_2 ratio are down-regulated (red). Lower ratios are marked depending of the up- and down-regulation in lighter green and red colors. ("Proteomics" - number of identifications in the symbiont proteome) (COS - *C. orbicularis*, COA - *C. orbiculata*)

			Freshly (control)				COS incubation				Ratio log_2			
ID DB	Product	Label	COA	+/- SD	COS	+/- SD	COS +[O_2]	+/- SD	COS -[O_2]	+/- SD	COS/COA	+[O_2]/-[O_2]	COS control /COS +[O_2]	Prote-omics
40	glucose-6-phosphate isomerase	Pgi	0.012	0.003	0.011	0.001			0.009	0.005	-0.140			8
197	phosphoglucomutase	GpmI	0.023	0.019										7
324	2-phosphoglycerate kinase	Pgk												3
478	fructose-bisphosphate aldolase class I	FbaAI												1
895	ketose-bisphospahte aldolase, class II	KbaII	0.049	0.017	0.021	0.012	0.042	0.027	0.037	0.024	-1.220	0.018	-1.010	6
1006	2,3-bisphosphoglycerate-independent phosphoglycerate mutase	GpmI												4
1023	phosphoenolpyruvate carboxylase	Ppc												4
1099	glucokinase	Glk												0
1107	6-phosphofructokinase	ATP-PfkA												4
1227	phosphoenolpyruvate carboxykinase (GTP)	PckA	0.074	0.012	0.034	0.002			0.013	0.003	-1.136			9
1245	glyceraldehyde-3-phosphate dehydrogenase, type I	GapA	0.014	0.006					0.003	0.001				8
1631	inositol monophosphatase/fructose-1,6-bisphosphatase family protein	FbpA												4
1820	phosphoenolpyruvate synthase	PpsA	0.008	0.002										6
2659	triosephosphate isomerase	TpiA	0.116	0.059	0.631	0.104	0.490	0.118	0.307	0.092	2.448	0.674	-0.364	10
3037	fructose-1,6-bisphosphate aldolase, class II	FbaAII	0.635	0.148	0.373	0.027	0.747	0.226	0.327	0.124	-0.770	1.190	1.002	12
3038	pyruvate kinase	Pyk	0.013	0.008										9
3039	phosphoglycerate kinase	Pgk	0.245	0.126	0.306	0.033	0.358	0.123	0.225	0.148	0.322	0.667	0.224	12
3040	glyceraldehyde-3-phosphate dehydrogenase, type I	GapA	0.482	0.111	0.357	0.104	0.219	0.105	0.089	0.088	-0.434	1.303	-0.708	12
3895	pyruvate kinase	Pyk												4
4327	enolase	Eno	0.030	0.006	0.015	0.004					-1.039			7

			Freshly (control)				COS incubation				Ratio log_2			
ID DB	Product	Label	COA	+/- SD	COS	+/- SD	COS $+[O_2]$	+/- SD	COS $-[O_2]$	+/- SD	COS/COA	$+[O_2]/-[O_2]$	COS control /COS $+[O_2]$	Prote-omics
38	glucose-1-phosphate adenylyltransferase	GlgC	0.0495	0.0004	0.0327	0.0267	0.0284	0.0155	0.0548	0.0155	-0.5970	-0.9520	-0.7443	10
49	UTP-glucose-1-phosphate uridylyltransferase	GalU	0.0173	0.0110			0.0156	0.0077	0.0182	0.0077		-0.2185		4

Table 3.21. Identified TCA-cycle and related proteins in *C. orbicularis and C. orbiculata* symbiont proteome.
C. orbicularis bivalves were incubated for one week of starvation in sterile sea water with saturated oxygen (+[O_2]) and low oxygen (-[O_2]) conditions. The data represent the mean ± standard deviation of three independent experiments each performed in three technical replicates measured with the Synapt G2 mass spectrometer. The protein amount is measured in fmol/ng. Ratios more than -0,5/0,5 indicate a significant change of the protein expression. Proteins with significant positive log_2 ratio are up-regulated (green), proteins with significant negative log_2 ratio are down-regulated (red). Lower ratios are marked depending of the up- and down-regulation in lighter green and red colors. ("Proteomics" - number of identifications in the symbiont proteome) (COS - *C. orbicularis*, COA - *C. orbiculata*)

				Freshly (control)				COS incubation				Ratio log_2			
ID DB	Product	Label	TMH	COA	+/- SD	COS	+/- SD	COS +[O_2]	+/- SD	COS -[O_2]	+/- SD	COS/COA	+[O_2]/-[O_2]	COS control/COS +[O_2]	Prote-omics
112	succinate dehydrogenase and fumarate reductase iron-sulfur protein	SdhB	0	0.004	0.0021					0.0092	0.0077				7
113	succinate dehydrogenase flavoprotein subunit	SdhA	0	0.0111	0.0075			0.0329	0.0211	0.0248	0.0285		0.4059		8
114	succinate dehydrogenase, hydrophobic membrane anchor	SdhD	3												0
115	succinate dehydrogenase, cytochrome b556 subunit	SdhC	3												0
469	oxaloacetate decarboxylase beta chain	OadB	9												4
470	oxaloacetate decarboxylase subunit alpha	OadA	0	0.0018	0.0023			0.0059	0.0017	0.0094	0.0046		-0.6750		10
471	oxaloacetate decarboxylase gamma chain	OadG	1												0
1113	succinyl-CoA synthetase subunit alpha	SucD	0												4
1114	succinyl-CoA synthetase subunit beta	SucC	0	0.0006	0.0050										6
1116	citrate synthase	GltA	0	0.0004	0.0082					0.0255	0.0187				9
1775	isocitrate dehydrogenase, NADP-dependent	Icd	0	0.0007	0.0113			0.0160	0.0110	0.0505	0.0426		-1.6554		9
1800	malate dehydrogenase	Mdh	0	0.026	0.0111	0.0960	0.0314			0.2187	0.1648				8
1805	isocitrate lyase	AceA	0							0.0224	0.0213				3
1806	malate synthase A	AceB	0												0
2979	2-oxoglutarate dehydrogenase E1 component	KgdE1	0												6
2980	dihydrolipoyllysine-residue succinyltransferase, E2 component of oxoglutarate dehydrogenase (Succinyl-transferring) complex	KgdE2	0												8
2981	fumarate hydratase class I	FumB	0	0.0001	0.0202										7
3005	pyruvate dehydrogenase complex E3 component, dihydrolipoamide dehydrogenase	AceE3	0	0.0001	0.0054			0.0196	0.0095	0.0125	0.0032		0.6497		8
3009	dihydrolipoyllysine-residue acetyltransferase component of pyruvate dehydrogenase complex	AceE2	0												6
3010	pyruvate dehydrogenase subunit E1	AceE1	0	0.0004	0.0029										7

				Freshly (control)				COS incubation				Ratio log_2			
ID DB	Product	Label	TMH	COA	+/- SD	COS	+/- SD	COS $+[O_2]$	+/- SD	COS $-[O_2]$	+/- SD	COS/COA	$+[O_2]/-[O_2]$	COS control/COS $+[O_2]$	Prote-omics
3556	bifunctional aconitate hydratase 2/2-methylisocitrate dehydratase	AcnA	0	0.0004	0.0163	0.0373	0.0031	0.0623	0.0241	0.0464	0.0142	-6.4209	0.4265	-0.7415	11
4022	dihydrolipoamide dehydrogenase	KgdE3	0	0.036	0.0071	0.1537	0.0452	0.1839	0.0698	0.0951	0.0489		0.9518	-0.2583	7
4030	succinyl-CoA synthetase subunit beta	SucC	0	0.0003	0.0021										6
4031	succinyl-CoA synthetase subunit alpha	SucD	0	0.0003	0.0026										5

Table 3.22. Identified PPP enzymes in the symbiont proteome *C. orbicularis* and *C. orbiculata* bivalves.
C. orbicularis bivalves were incubated for one week of starvation in sterile sea water with saturated oxygen (+[O2]) and low oxygen (-[O2]) conditions. The data represent the mean ± standard deviation of three independent experiments each performed in three technical replicates measured with the Synapt G2 mass spectrometer. The protein amount is measured in fmol/ng. Ratios more than -0,5/0,5 indicate a significant change of the protein expression. Proteins with significant positive log2 ratio are up-regulated (green), proteins with significant negative log2 ratio are down-regulated (red). Lower ratios are marked depending of the up- and down-regulation in lighter green and red colors. ("Proteomics" - number of identifications in the symbiont proteome) (COS - *C. orbicularis*, COA - *C. orbiculata*)

			Freshly (control)				COS incubation				Ratio log2			
ID D	Product	Label	COA	+/- SD	COS	+/- SD	COS $+[O_2]$	+/- SD	COS $-[O_2]$	+/- SD	COS/COA	$+[O_2]/-[O_2]$	COS control /COS $+[O_2]$	Prote-omics
9	6-phosphogluconate dehydrogenase, decarboxylating	Gnd												0
10	6-phosphogluconolactonase	Pgl												1
11	glucose-6-phosphate 1-dehydrogenase	Zwf												3
3808	ribose-5-phosphate isomerase A	RpiA	0.006	0.003	0.022	0.003	0.048	0.015	0.033	0.022	1.955	0.571	-1.118	11
3041	transketolase, bacterial-like	TktA	0.419	0.107	0.127	0.053	0.410	0.223	0.156	0.090	-1.722	1.398	-1.690	12
3732	ribulose-phosphate 3-epimerase	Rpe	0.094	0.033	0.238	0.048	0.206	0.038	0.100	0.020	1.336	1.044	0.206	9
309	phosphoribulokinase	Prk	0.270	0.102	0.626	0.071	0.449	0.077	0.545	0.418	1.212	-0.282	0.481	12

Table 3.23. Identified CBB related proteins in *C. orbicularis and C. orbiculata* symbiont proteome.
C. orbicularis bivalves were incubated for one week of starvation in sterile sea water with saturated oxygen (+[O_2]) and low oxygen (-[O_2]) conditions. Data represent the mean ± standard deviation of three independent experiments each performed in three technical replicates measured with the Synapt G2 mass spectrometer. The protein amount is measured in fmol/ng. Ratios more than -0,5/0,5 indicate a significant change of the protein expression proteins with significant positive log_2 ratio are up-regulated (green), proteins with significant negative log_2 ratio are down-regulated (red). Lower ratios are marked depending of the up- and down-regulation in lighter green and red colors. ("Proteomics" - number of identifications in the symbiont proteome, gene clusters are indicated by double line) (COS - *C. orbicularis*, COA - *C. orbiculata*)

			Freshly (control)				COS incubation				Ratio log_2			
ID	Product	Label	COA	+/- SD	COS	+/- SD	COS +[O_2]	+/- SD	COS -[O_2]	+/- SD	COS/COA	+[O_2]/-[O_2]	COS control/COS +[O_2]	Prote-omics
312	RuBisCO operon transcriptional regulator CbbR	CbbR												0
313	ribulose bisophosphate carboxylase	RbcL1	0.638	0.070	4.631	0.403	3.006	0.703	1.921	0.300	2.860	0.646	0.624	12
314	ribulose-1,5-bisphosphate carboxylase/oxygenase small subunit	RbcS1	0.841	0.124	4.570	1.399	2.388	0.847	1.715	0.352	2.443	0.477	0.936	12
315	CbbQ	CbbQ	0.046	0.005	0.151	0.062	0.047	0.042	0.032	0.015	1.699	0.552	1.683	12
316	RuBisCO activation protein CbbO	CbbO	0.010	0.007					0.017	0.008				8
319	ribulose 1,5-bisphosphate carboxylase large subunit	RbcL2	0.016	0.007	0.079	0.042	0.087	0.016	0.042	0.032	2.296	1.028	-0.132	10
320	RuBisCO activation protein CbbQ	CbbQ												0
321	RuBisCO activation protein CbbO	CbbO												0
984	6-phosphofructokinase	PP-PfkA	0.250	0.059	0.186	0.046	0.047	0.024	0.102	0.045	-0.427	-1.107	1.977	12
985	vacuolar-type H(+)-translocating pyrophosphatase, putative HppA	HppA	0.039	0.011										9
986	inorganic diphosphatase	PpA	0.019	0.005			0.310	0.247	0.078	0.068		1.996		7
309	phosphoribulokinase	PrkB	0.270	0.102	0.626	0.071	0.449	0.077	0.545	0.418	1.212	-0.282	0.481	12
3037	fructose-1,6-bisphosphate aldolase, class II	FbaAII	0.635	0.148	0.373	0.027	0.747	0.226	0.327	0.124	-0.770	1.190	-1.002	12
none specific enzymes														
1631	inositol monophosphatase/fructose-1,6-bisphosphatase family protein	Fbp												4
3808	ribose-5-phosphate isomerase A	RpiA	0.006	0.003	0.022	0.003	0.048	0.015	0.033	0.022	1.955	0.571	-1.118	11
3041	transketolase, bacterial-like	TktA	0.419	0.107	0.127	0.053	0.410	0.223	0.156	0.090	-1.722	1.398	-1.690	12
3040	glyceraldehyde-3-phosphate dehydrogenase, type I	GapA	0.482	0.111	0.357	0.104	0.219	0.105	0.089	0.088	-0.434	1.303	0.708	12
2659	triosephosphate isomerase	TpiA	0.116	0.059	0.631	0.104	0.490	0.118	0.307	0.092	2.448	0.674	0.364	10
3039	phosphoglycerate kinase	PgkA	0.245	0.126	0.306	0.033	0.358	0.123	0.225	0.148	0.322	0.667	-0.224	12

Table 3.24. Identified proteins of the ATP synthase in *C. orbicularis and C. orbiculata* symbiont proteome.
C. orbicularis bivalves were incubated for one week of starvation in sterile sea water with saturated oxygen (+[O_2]) and low oxygen (-[O_2]) conditions. The data represent the mean ± standard deviation of three independent experiments each performed in three technical replicates measured with the Synapt G2 mass spectrometer. The protein amount is measured in fmol/ng. Ratios more than -0,5/0,5 indicate a significant change of the protein expression. Proteins with significant positive log_2 ratio are up-regulated (green), proteins with significant negative log_2 ratio are down-regulated (red). Lower ratios are marked depending of the up- and down-regulation in lighter green and red colors. ("Proteomics" - number of identifications in the symbiont proteome, gene clusters are indicated by double line) (COS - *C. orbicularis*, COA - *C. orbiculata*)

							Freshly (control)				COS incubation				Ratio log_2			
ID DB	Product	Label	IEP	MW (kDa)	Cello	TMH	COA	+/- SD	COS	+/- SD	COS +[O_2]	+/- SD	COS -[O_2]	+/- SD	COS/COA	+[O_2]/-[O_2]	COS control/ COS +[O_2]	Proteomics
1067	ATP synthase subunit B	AtpB	5.3	17.4	CP	1	0.014	0.003										9
1068	ATP synthase subunit delta	AtpD	4.2	19.2	CP	0	0.024	0.011	0.184	0.030	0.060	0.023	0.062	0.069	2.934	-0.054	1.628	8
1069	F0F1 ATP synthase subunit alpha	AtpA	4.8	55.6	CP	0	0.154	0.026	0.136	0.012	0.055	0.019	0.065	0.040	-0.181	-0.237	1.292	12
1070	F0F1 ATP synthase subunit gamma	AtpG	7	31.9	CP	0	0.036	0.010					0.014	0.006				9
1071	F0F1 ATP synthase subunit beta	AtpB	4.6	49.4	CP	0	0.200	0.019	0.257	0.080	0.191	0.033	0.194	0.215	0.359	-0.019	0.427	12
1072	ATPase, F1 complex, delta/epsilon subunit	AtpD	4.9	15.3	CP	0	0.024	0.004	0.065	0.009	0.082	0.020	0.052	0.020	1.440	0.653	-0.324	6
1940	V-type ATP synthase subunit E	AtpE	8.8	24.1	CP	0												7

Table 3.25. Identified proteins of the Sox system in *C. orbicularis and C. orbiculata* symbiont proteome.
C. orbicularis bivalves were incubated for one week of starvation in sterile sea water with saturated oxygen (+[O_2]) and low oxygen (-[O_2]) conditions. The data represent the mean ± standard deviation of three independent experiments each performed in three technical replicates measured with the Synapt G2 mass spectrometer. The protein amount is measured in fmol/ng. Ratios more than -0,5/0,5 indicate a significant change of the protein expression. Proteins with significant positive log_2 ratio are up-regulated (green), proteins with significant negative log_2 ratio are down-regulated (red). Lower ratios are marked depending of the up- and down-regulation in lighter green and red colors. ("Proteomics" - number of identifications in the symbiont proteome, gene custers are indicated by double lines) (COS - *C. orbicularis*, COA - *C. orbiculata*)

				Freshly (control)				COS incubation				Ratio log_2			
ID	Product	Label	TMH	COA	+/- SD	COS	+/- SD	COS +[O_2]	+/- SD	COS -[O_2]	+/- SD	COS/COA	+[O_2]/-[O_2]	COS control/COS +[O_2]	Proteomics
563	sulfur oxidation protein SoxZ	SoxZ	0	0.119	0.067			0.088	0.034	0.014	0.007		2.624		9
564	sulfur oxidation protein SoxY	SoxY	0	0.031	0.008			0.048	0.017	0.009	0.002		2.450		6
770	sulfur oxidation protein SoxX	SoxX	0	0.025	0.004										3
771	sulfur oxidation protein SoxA	SoxA	0	0.041	0.008			0.021	0.001	0.016	0.016		0.393		6
772	putative SoxXA-binding protein	SoxAX	0	0.060	0.021			0.022	0.006	0.011	0.008		1.009		4
773	sulfur transferase	SoxL	0	0.031	0.021	0.026	0.002	0.018	0.001	0.052	0.030	-0.239	-1.523	0.506	10
1318	sulfur/thiosulfate oxidation protein SoxB	SoxB	0	0.137	0.031	0.057	0.017	0.050	0.026	0.137	0.079	-1.268	-1.447	0.171	9
2485	putative thioredoxin SoxW	SoxW	0												0
4159	sulfur oxidation protein SoxZ	SoxZ	0												0
4160	sulfur oxidation protein SoxY	SoxY	0												0

Table 3.26. Identified proteins of the Dsr- sulfur oxidizing pathway in *C. orbicularis and C. orbiculata* symbiont proteome.
C. orbicularis bivalves were incubated for one week of starvation in sterile sea water with saturated oxygen (+[O_2]) and low oxygen (-[O_2]) conditions. The data represent the mean ± standard deviation of three independent experiments each performed in three technical replicates measured with the Synapt G2 mass spectrometer. The protein amount is measured in fmol/ng. Ratios more than -0,5/0,5 indicate a significant change of the protein expression. Proteins with significant positive log_2 ratio are up-regulated (green), proteins with significant negative log_2 ratio are down-regulated (red). Lower ratios are marked depending of the up- and down-regulation in lighter green and red colors. ("Proteomics" - number of identifications in the symbiont proteome, gene clusters are indicated by double line) (COS - *C. orbicularis*, COA - *C. orbiculata*)

			Freshly (control)				COS incubation				Ratio log_2			
ID	Product	Label	COA	+/- SD	COS	+/- SD	+[O_2]	+/- SD	-[O_2]	+/- SD	COS/COA	+[O_2]/-[O_2]	COS control/COS +[O_2]	Prote-omics
533	DsrC like protein	DsrC	0.053	0.006			0.078	0.012	0.048	0.041		0.707		5
537	dissimilatory sulfite reductase, alpha subunit	DsrA	0.251	0.055	0.432	0.082	0.220	0.091	0.133	0.068	0.782	0.722	0.972	12
538	dissimilatory sulfite reductase, beta subunit	DsrB	0.219	0.043	0.318	0.008	0.149	0.051	0.156	0.058	0.536	-0.061	1.089	12
539	sulfur oxidation protein DsrE, TusD-like sulfur relay protein	DsrE	0.287	0.066	0.169	0.012	0.226	0.055	0.080	0.019	-0.768	1.506	-0.426	10
540	sulfur oxidation protein DsrF, TusC-like sulfur relay protein	DsrF	0.147	0.023	0.242	0.077	0.207	0.111	0.053	0.012	0.718	1.972	0.226	9
541	sulfite reduction-associated protein DsrH	DsrH	0.139	0.005	0.264	0.050	0.339	0.077	0.116	0.037	0.924	1.542	-0.360	6
542	dissimilatory sulfite reductase (desulfoviridin) gamma subunit	DsrC	0.711	0.202	0.264	0.133	0.681	0.131	0.524	0.318	-1.429	0.376	-1.367	
543	sulfur oxidation protein DsrM	DsrM												8
544	sulfur oxidation protein DsrK	DsrK	0.017	0.003										8
545	sulfite reduction-associated protein DsrL	DsrL	0.219	0.061	0.444	0.073	0.252	0.162	0.087	0.022	1.018	1.535	0.819	12
546	sulfite reduction-associated complex DsrMKJOP iron-sulfur protein DsrO	DsrO	0.006	0.004										9
547	sulfite reduction-associated complex DsrMKJOP protein DsrP	DsrP												0
548	sulfite reduction-associated protein DsrN	DsrN												0
549	sulfite reduction-associated protein DsrR	DsrR												4
550	sulfur oxidation-associated protein DsrS	DsrS												0

Table 3.27. Identified sulfur related proteins in *C. orbicularis and C. orbiculata* symbiont proteome.
C. orbicularis bivalves were incubated for one week of starvation in sterile sea water with saturated oxygen (+[O_2]) and low oxygen (-[O_2]) conditions. The data represent the mean ± standard deviation of three independent experiments each performed in three technical replicates measured with the Synapt G2 mass spectrometer. The protein amount is measured in fmol/ng. Ratios more than -0,5/0,5 indicate a significant change of the protein expression. Proteins with significant positive log_2 ratio are up-regulated (green), proteins with significant negative log_2 ratio are down-regulated (red). Lower ratios are marked depending of the up- and down-regulation in lighter green and red colors. ("Proteomics" - number of identifications in the symbiont proteome, gene clusters are seperated by double lines) (COS - *C. orbicularis*, COA - *C. orbiculata*)

			Freshly (control)				COS incubation				Ratio log_2			
ID	Product	Label	COA	+/- SD	COS	+/- SD	+[O_2]	+/- SD	-[O_2]	+/- SD	COS/ COA	+[O_2]/ -[O_2]	COS control/ COS +[O_2]	Prote-omics
145	sulfate adenylyltransferase	SopT/Sat	0.232	0.0607	0.35	0.0252	0.134	0.0306	0.089	0.0382	0.597	0.584	1.386	11
2280	putative sulfide-quinone reductase	Sqr	0.014	0.0079										3
3699	adenylylsulfate reductase, beta subunit	AprB	1.088	0.1979	1.108	0.2277	0.826	0.2776	0.386	0.0605	0.026	1.098	0.423	12
3700	adenylylsulfate reductase, alpha subunit	AprA	1.288	0.1742	1.610	0.1939	0.617	0.4005	0.399	0.2346	0.322	0.628	1.384	12
3709	heterodisulfide reductase, subunit B	Hdr	0.082	0.0277	0.035	0.0078	0.039	0.0085	0.023	0.0060	-0.674	0.765	-0.17	10
3710	heterodisulfide reductase, putative QmoC	QmoC	0.037	0.0126			0.093	0.0884	0.018	0.0056		2.411		9
3711	heterodisulfide reductase, putative QmoB	QmoB	0.169	0.0648	0.204	0.1018	0.138	0.0961	0.295	0.1723	-0.221	-1.099	0.564	12
3712	heterodisulfide reductase, putative QmoA	QmoA	0.117	0.0322	0.136	0.0556	0.165	0.1302	0.079	0.0352	-0.234	1.061	-0.272	12
3220	sulfide dehydrogenase (flavocytochrome), flavoprotein subunit	FccB												8
3221	putative cytochrome subunit of sulfide dehydrogenase	FccA												8
3973	sulfide dehydrogenase (flavocytochrome), flavoprotein subunit	FccB	0.118	0.0054	0.099	0.0431	0.169	0.0375	0.071	0.0622	-0.249	1.250	-0.769	12
3974	putative cytochrome subunit of sulfide dehydrogenase	FccA												6

Table 3.28. Identified nitrogen fixation proteins in *C. orbicularis and C. orbiculata* symbiont proteome.
C. orbicularis bivalves were incubated for one week of starvation in sterile sea water with saturated oxygen (+[O_2]) and low oxygen (-[O_2]) conditions. The data represent the mean ± standard deviation of three independent experiments each performed in three technical replicates measured with the Synapt G2 mass spectrometer. The protein amount is measured in fmol/ng. Ratios more than -0,5/0,5 indicate a significant change of the protein expression. Proteins with significant positive log_2 ratio are up-regulated (green), proteins with significant negative log_2 ratio are down-regulated (red). Lower ratios are marked depending of the up- and down-regulation in lighter green and red colors. ("Proteomics" - number of identifications in the symbiont proteome) (COS - *C. orbicularis*, COA - *C. orbiculata*)

			Freshly (control)				COS incubation				Ratio log_2			
ID	Product	Label	COA	+/- SD	COS	+/- SD	+[O_2]	+/- SD	-[O_2]	+/- SD	COS/ COA	+[O_2]/ -[O_2]	COS control/ COS +[O_2]	Prote-omics
2593	nitrogen fixation protein NifM	NifM												0
2594	NifZ family protein	NifZ												1
2595	nitrogenase-stabilizing/protective protein nifW	NifW												0
2596	hypothetical protein	HP_2596												3
2597	serine O-acetyltransferase	CysE												0
2598	homocitrate synthase	NifV												0
2599	cysteine desulfurase NifS	NifS	0.02	0.0055	0.033	0.0026					0.725			6
2600	Fe-S cluster assembly protein NifU	NifU	0.012	0.0069	0.119	0.0602					3.308			6
2601	iron-sulfur cluster assembly accessory protein	2601	0.075	0.0134			0.084	0.0254	0.035	0.0215		1.279		6
2602	hypothetical protein, putative AlG2	HP_2602												0
2603	hypothetical protein	HP_2603												0
2604	ferredoxin 2Fe-2S	2604												0
2605	ferredoxin III 4[4Fe-4S] nif-specific	2605			0.033	0.0037								1
2606	putative nitrogen fixation protein	2606	0.079	0.0260	0.176	0.0340	0.091	0.0127	0.021	0.0058	1.157	2.132	0.951	4
2607	nitrogen fixation protein	NifX	0.174	0.0309	0.076	0.0138			0.032	0.0195	-1.199			11
2608	positive regulator of sigma(E) RseC/MucC	2608												0
2609	dinitrogenase iron-molybdenum cofactor biosynthesis protein	NifX												3
2610	nitrogenase molybdenum-iron cofactor biosynthesis protein NifN	NifN	0.008	0.0014										5
2611	nitrogenase iron-molybdenum cofactor biosynthesis protein NifE	NifE	0.004	0.0012										3
2612	conserved hypothetical protein	HP_2612												3
2613	glycosyl transferase family 3	Gt_3	0.011	0.0045										6
2614	LRV FeS4 cluster domain-containing protein	2614												0
2615	hypothetical protein	HP_2615												0
2616	dinitrogenase iron-molybdenum cofactor biosynthesis protein	NifY			0.017	0.0064								3

			Freshly (control)				COS incubation				Ratio log_2			
ID	Product	Label	COA	+/- SD	COS	+/- SD	+[O_2]	+/- SD	- [O_2]	+/- SD	COS/ COA	+[O_2]/ -[O_2]	COS control/ COS +[O_2]	Prote-omics
2617	4Fe-4S ferredoxin iron-sulfur binding domain-containing protein	FdxN	0.051	0.0143	0.276	0.2614	0.113	0.0721	0.084	0.0701	2.438	0.433	1.283	4
2618	putative nitrogen fixation protein FixT	FixT/NifT			0.213	0.0246								1
2619	nitrogenase molybdenum-iron protein subunit beta	NifK	0.305	0.0725	0.119	0.0877	0.021	0.0177	0.023	0.0199	-1.362	-0.144	2.533	12
2620	nitrogenase molybdenum-iron protein subunit alpha	NifD	0.137	0.0209	0.035	0.0060					-1.957			10
2621	nitrogenase molybdenum-iron protein	NifH	0.638	0.1161	0.55	0.1137	0.078	0.0250	0.047	0.0088	-0.214	0.726	2.821	9
2622	NAD(+)-dinitrogen-reductase ADP-D-ribosyltransferase	2622												0
2623	conserved hypothetical protein	HP_2623												0
2624	conserved hypothetical protein	HP_2624	0.011	0.0018										6
2625	amidohydrolase 2	2625												0
2626	hemerythrin metal-binding domain-containing protein	2626												0
2627	ADP-ribosyl-(dinitrogen reductase) hydrolase	2627												1
2628	NifQ	NifQ												0
2629	ferredoxin	2629	0.238	0.0641	0.073	0.0125	0.041	0.0139	0.017	0.0076	-1.707	1.272	0.815	7
2630	ferredoxin	FdxA_2630	0.414	0.1128	1.827	0.2263	0.246	0.0957	0.127	0.0419	2.143	0.954	2.890	7
2631	flavodoxin/nitric oxide synthase	2631												2
2632	nitrogenase-associated protein	NifO												0
2633	4Fe-4S ferredoxin iron-sulfur binding domain-containing protein	2633												0
2634	nitrogenase cofactor biosynthesis protein NifB	NifB												0
2635	transcriptional regulator NifA subfamily. Fis Family	NifA	0.005	0.0024										3
2636	PAS/PAC sensor signal transduction histidine kinase	NifL												4

Table 3.29. Putative nitrogenase protection proteins of the nitrogenase in *C. orbicularis and C. orbiculata* symbiont proteome.
C. orbicularis bivalves were incubated for one week of starvation in sterile sea water with saturated oxygen (+[O_2]) and low oxygen (-[O_2]) conditions. The data represent the mean ± standard deviation of three independent experiments each performed in three technical replicates measured with the Synapt G2 mass spectrometer. The protein amount is measured in fmol/ng. Ratios more than -0,5/0,5 indicate a significant change of the protein expression. Proteins with significant positive log_2 ratio are up-regulated (green), proteins with significant negative log_2 ratio are down-regulated (red). Lower ratios are marked depending of the up- and down-regulation in lighter green and red colors. (CELLO - subCELlular LOcalization prediction, CP - cytoplasm, IM - inner membrane, EC - extracellular, PP - periplasm, OM - outer membrane, "Proteomics" - number of identifications in the symbiont proteome) (COS - *C. orbicularis*, COA - *C. orbiculata*)

							Freshly (control)				COS incubation				Ratio log_2			
ID	Product	Label	IEP	MW (kDa)	Cello	TMH	COA	+/- SD	COS	+/- SD	+[O_2]	+/- SD	-[O_2]	+/- SD	COS/ COA	+[O_2]/ -[O_2]	COS control / COS +[O_2]	Prote-omics
130	rubrerythrin	RubY	4.1	20.2	CP	0												1
826	rubrerythrin	RubY	4.8	19.2	CP	0												0
1095	Fe-S oxidoreductase-like protein in Rubrerythrin cluster		6.3	50.1	CP	0	0.0508	0.0125	0.008	0.0029			0.013	0.0048	-2.672			8
1096	rubrerythrin	RubY	4.4	15.2	CP	0	1.755	0.4254	4.119	0.5809	2.629	0.9529	1.203	0.3237	1.230	1.128	0.6477	12
1282	cytochrome bd-type quinol oxidase subunit 1-like protein		9.1	68.5	IM	14												0
3897	superoxide dismutase (Fe)	SodA	4.6	21.4	EC	0	0.1639	0.0202	0.264	0.0770	0.6411	0.3012	0.516	0.2896	0.6851	0.313	-1.282	10
399	catalase/peroxidase HPI	KatG	4.4	75.1	CP	0	0.012	0.0021										2

Table 3.30. Key proteins of the amino acid metabolism in the *C. orbicularis and C. orbiculata* symbiont proteome.
C. orbicularis bivalves were incubated for one week of starvation in sterile sea water with saturated oxygen (+[O_2]) and low oxygen (-[O_2]) conditions. The data represent the mean ± standard deviation of three independent experiments each performed in three technical replicates measured with the Synapt G2 mass spectrometer. The protein amount is measured in fmol/ng. Ratios more than -0,5/0,5 indicate a significant change of the protein expression. Proteins with significant positive log_2 ratio are up-regulated (green), proteins with significant negative log_2 ratio are down-regulated (red). Lower ratios are marked depending of the up- and down-regulation in lighter green and red colors. (CELLO - subCELlular LOcalization prediction, CP - cytoplasm, PP - periplasm, "Proteomics" - number of identifications in the symbiont proteome) (COS - *C. orbicularis*, COA - *C. orbiculata*)

							Freshly (control)				COS incubation				Ratio log_2			
ID	Product	Label	IEP	MW (kDa)	Cello	TMH	COA	+/- SD	COS	+/- SD	+[O_2]	+/- SD	-[O_2]	+/- SD	COS/ COA	+[O_2]/-[O_2]	COS control/ COS +[O_2]	Prote-omics
809	aspartate aminotransferase	Ast	4.7	42	PP	0	0.1355	0.0313	0.1137	0.0119	0.1795	0.0414	0.1088	0.0106	-0.2526	0.7223	-0.6589	12
825	glutamate synthase [NADPH] small chain	GltD	5.4	60	CP	0												2
1038	glutamine synthetase, type I	GlnA	4.5	51.8	CP	0	0.5728	0.0565	0.6452	0.4673	0.7466	0.3784	0.5263	0.1416	0.1718	0.5044	-0.2105	12
1350	glutamate dehydrogenase, partial	GdhA	5.8	51.7	CP	0	0.0059	0.0043										3
3581	glutamate synthetase	GltX	5.9	60.9	CP	0												1

Table 3.31. Identified stress proteins in *C. orbicularis and C. orbiculata* symbiont proteome.
C. orbicularis bivalves were incubated for one week of starvation in sterile sea water with saturated oxygen (+[O_2]) and low oxygen (-[O_2]) conditions. The data represent the mean ± standard deviation of three independent experiments each performed in three technical replicates measured with the Synapt G2 mass spectrometer. The protein amount is measured in fmol/ng. Ratios more than -0,5/0,5 indicate a significant change of the protein expression. Proteins with significant positive log_2 ratio are up-regulated (green), proteins with significant negative log_2 ratio are down-regulated (red). Lower ratios are marked depending of the up- and down-regulation in lighter green and red colors. ("Proteomics" - number of identifications in the symbiont proteome, Cello - subCELlular LOcalization prediction, CP - cytoplasmic, PP - periplasma, IM - inner membrane, EC - extracellular, gene clusters are indicated by double line)(COS - *C. orbicularis*, COA - *C. orbiculata*)

							Freshly (control)				COS incubation				Ratio log_2			
ID	Product	Label	IEP	MW (kDa)	TMH	Cello	COA	+/- SD	COS	+/- SD	COS +[O_2]	+/- SD	COS - [O_2]	+/- SD	COS/COA	+[O_2]/-[O_2]	COS control/COS +[O_2]	Prote-omics
3781	co-chaperonin GroES	GroES	4.7	10.4	0	CP	0.749	0.094	2.529	0.112	3.160	1.003	1.661	0.356	1.755	0.927	-0.322	8
3782	molecular chaperone GroEL	GroEL	4.5	57.8	0	CP	0.499	0.046	1.059	0.227	1.388	0.604	0.561	0.165	1.085	1.306	-0.390	12
2089	stringent starvation protein A	SspA	4.9	23.8	0	CP	0.031	0.011	0.065	0.015	0.042	0.011	0.041	0.014	1.083	0.052	0.627	11
3897	superoxide dismutase (Fe)	SodA	4.6	21.4	0	EC	0.164	0.020	0.263	0.077	0.641	0.301	0.516	0.290	0.685	0.313	-1.283	10
976	heat shock protein HslVU, ATPase subunit HslU	HslU	4.8	49.4	0	CP	0.014	0.013										8
1868	thiol peroxidase	Tpx	4.1	17.6	0	PP	0.554	0.057	0.260	0.051	0.510	0.167	0.205	0.013	-1.091	1.313	-0.971	7
2761	cold shock, CspA	CspA	5.5	7.7	0	CP	0.184	0.081	0.238	0.015	0.554	0.157	0.210	0.068	0.374	1.399	-1.218	7
58	GrpE nucleotide exchange factor	GrpE	4.2	23.2	0	CP	0.026	0.010			0.036	0.004	0.013	0.005		1.443		5
2165	UspA domain-containing protein	UspA	10.2	18.7	0	CP												5
2818	cold shock, CspA	CspA	5.4	7.8	0	CP	0.083	0.049	0.081	0.011	0.308	0.117	0.084	0.037	-0.025	1.864	-1.921	4
4671	protein containing UspA domain	UspA	5.3	20.0	0	CP												4
4758	multidrug resistance protein AcrB/AcrD/AcrF	Acr	5.7	115.5	11	IM												4
51	putative heat shock protein	Hsp	6.1	45.7	0	CP												3
100	copper-translocating P-type ATPase	CtpA	5	104.3	7	IM												3
3615	multidrug resistance protein	MdtL	5	114.5	11	IM												3
399	catalase/peroxidase HPI	KatG	4.4	75.1	0	CP	0.012	0.002										2
1790	multidrug resistance protein MdtB	MdtB	5	106.3	9	IM												2
3299	copper resistance protein CopB	CopB	4.1	24.5	0	CP												2
3630	cysteine desulfurase activator complex subunit SufB	SufB	4.8	53.3	0	CP												2

3.3 Endosymbiont digestion in the bivalves *Codakia* - Cytochemistry

During starvation, the gills of the bivalves changed visibly from thick and to thin and dark brown. The mortality of the bivalves in the sterile sea water tanks was low: around one to five percent per week. Although the lethality was higher during the first three weeks compared to the later five months of starvation. On transmission electron micrographs, positive enzymatic reactions appeared as black areas, resulting from the reaction of acid phosphatase and arylsulfatase with lead and barium, respectively.

In freshly collected *C. orbiculata,* gill filaments showed weak positive reaction: only a few black areas were visible in a couple of bacteriocytes (figure 3.54 A/B) while most of the bacteriocytes throughout the gill filaments are free of lysosomal enzyme activity (data not shown). This data suggests that adult individuals in their natural habitat can digest symbionts. After two weeks of starvation, the acid phosphatase activity was detected close to symbionts, while other gill cells were free of enzyme activity (fig 3.54 C/D). Likewise, the activity of the arylsulfatase increased during starvation time, after two months most of the symbionts inside the bacteriocytes are affected by an arylsulfatase activity (fig 3.54 E). After the beginning of bacterial degradation in the gills, the lysosomal activity was observed in all bacteriocytes of the lateral zone. It seems that it is a global process in the bacteriocytes of the gill in case of starvation. The further the course of starvation in both bivalves went, the more the amount of black areas increased until most of the bacteria were degraded. Nevertheless, for long period of starvation (5 months and more), the lysosomal activity globally decreases as less bacteriocytes are present through the lateral zone which now contains mostly granule cells. However, in the few remaining bacteriocytes which contain only few bacterial symbionts, there is still a lysosomal digestion occurring. No black areas were observed in granule cells and intercalary cells.

For *C. orbicularis* the story looked similar. After three weeks of starvation, no positive signal was detected for arylsulfatase and acid phosphatase in the gills of *C. orbicularis*, indicating very weak to no lysosomal activity (figure 3.55 A/B). In *C. orbicularis,* both enzyme activities were detected after six weeks of starvation (figure 3.55 C/D/E). The longer the starvation period, the more extensive was the area corresponding to the enzyme activity, until most symbionts seemed affected (data not shown). Similarly to *C. orbiculata*, the enzymatic activity in *C. orbicularis* increased during the first months of starvation, then progressively decreased after 4-5 months of starvation when the number of bacteriocytes and bacterial gill-endosymbionts have been strongly reduced through the lateral zone. The EDX analysis confirmed that the black areas contained barium and lead demonstrating that arylsulfatase and acid phosphatase enzymes were efficient (figure 3.54F/3.55F). In the negative controls, no barium or lead was detectable (not shown).

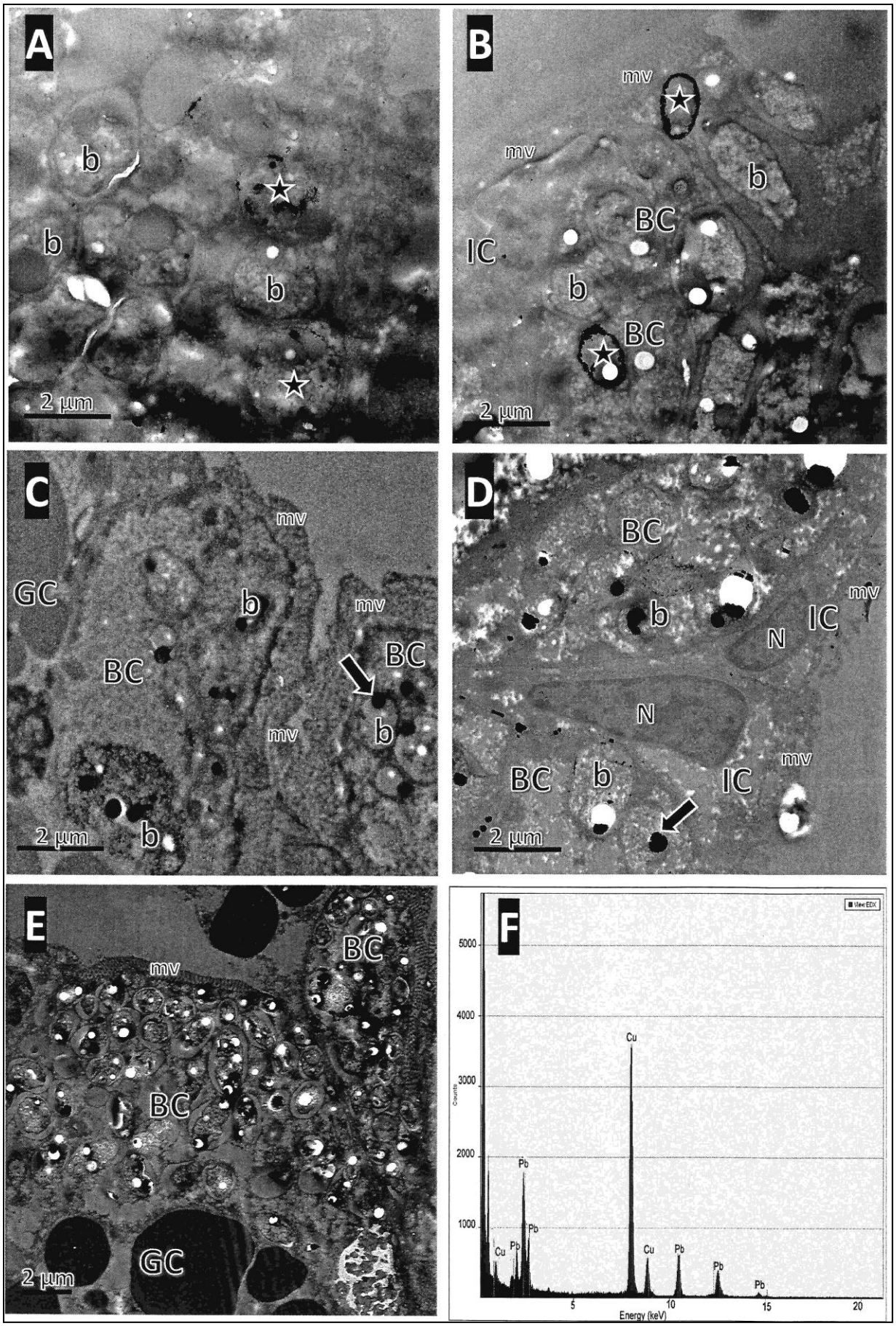

Figure 3.54. Cytochemical detection of lysosomal enzymes arylsulfatase and acid phosphatase in thin gill sections of *C. orbiculata.*

Figure 3.54. A-B: Very few bacteriocytes from freshly collected individuals presented a couple of bacterial endosymbionts (stars) partially degraded due to the activity of arylsulfatase (A) and to acid phosphatase (B). C-D: In two week starved individuals, the number of bacteria degraded by lysosomal enzymes like acid phosphatase increases with numerous bacteriocytes concerned. In each bacteriocyte (BC), several bacteria (b) are degraded by this enzyme suggesting a high lysosomal activity from host. In the same time, other gill cells (intercalary cell (IC) and granule cell (GC)) devoid of bacterial symbionts present no lysosomal activity. E: In two month starved bivalves, most of the bacteria inside the bacteriocytes appear to be concerned by arylsulfatase degradation. F: EDX spectrum obtained from cytochemical detection of acid phosphatase by Lead precipitation from black areas located inside bacterial endosymbionts indicated by black arrows in pictures C and D. Cupper comes from the grid bars supporting the thin section. Figures were took and labeled by O. Gros. (BC - bacteriocytes; GC - granule cells; IC - intercalary cells; mv - microvilli; N - nucleus)

Figure 3.55. A-B: Conversely to *C. orbiculata*, lysosomal activity in *C. orbicularis* was weak or absent in freshly collected individuals and even after three weeks of starvation, no activity was detected either for arylsulfatase (A) or for acid phosphatase (B). C-D: Acid phosphatase activity from six weeks starved individuals concerns a large number of bacteria within most of the bacteriocytes composing each gill filament. In each bacteriocyte (BC), several bacteria (b) are degraded simultaneously by this enzyme confirming a strong lysosomal activity from the host cell. In the same time, granule cell (GC), which are normally devoid of bacterial endosymbionts, present no lysosomal activity. E: In six weeks starved bivalves, all the symbionts seemed to be concerned by arylsulfatase degradation as shown by the numerous black areas observed through the cytoplasm of the bacteriocytes. F: EDX spectrum obtained from cytochemical detection of arylsulfatase activity by Barium precipitation from black areas (black arrows) in pictures C-E. Cupper comes from the grid bars supporting the thin section. Figures were took and labeled by O. Gros. (BC - bacteriocytes; GC - granule cells; IC - intercalary cells; mv - microvilli; N - nucleus)

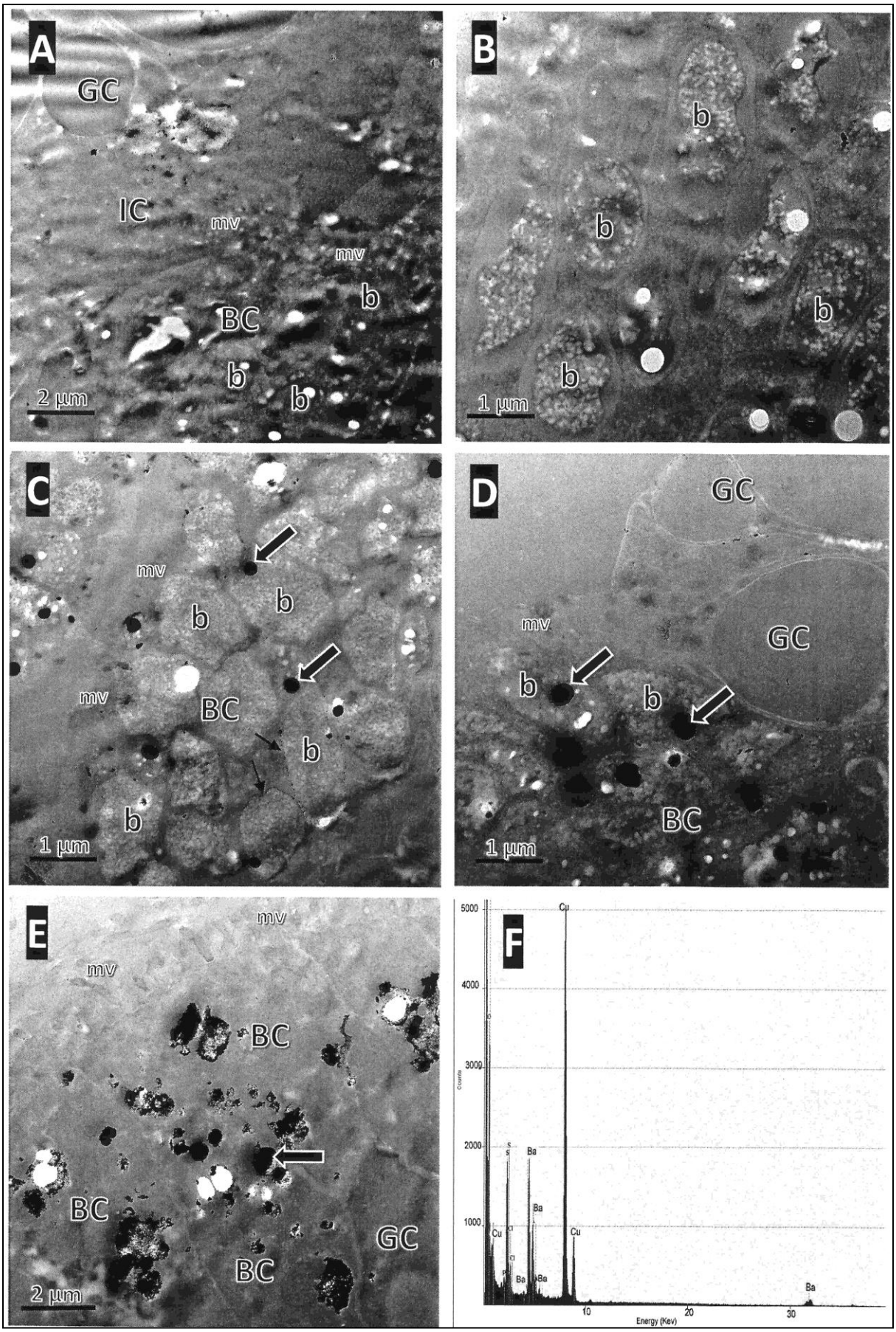

Figure 3.55. Cytochemical detection of lysosomal enzymes arylsulfatase and acid phosphatase in thin gill sections of *C. orbicularis.*

4. Discussion

The tropical bivalves *Codakia orbiculata* (Montagu, 1802) and *C. orbicularis* (Linnè 1792) live in the shallow-water sediment of the seagrass *Thalassium testudinum*. The seagrass ecosystem is a very productive system of organic matter. The anaerobic degradation of dead organic matter produces reduced sulfur compounds which are toxic for most plants and animals (Dorman et al., 2002). The bivalves *Codakia* burrow in the sediment to gain access to the sulfur pockets and pump particle-laden, oxygen saturated surface water (Dufour et al., 2003). The bivalves posses a reduced digestive tract, but particle feeding is still part of the feeding regime (Reid, 1990). Chemosynthetic symbiotic bivalves obtain main part of nutrients from symbiotic bacteria (Distel, 1998). The aim of this study was the investigation of the physiology of the symbiont and the symbiotic interaction with the bivalve.
Two features characterize this symbiotic model: (i) Bivalves live in shallow water, the sampling of bivalves is easy and cheap. (ii) Bivalves together with the symbionts can be cultivated in the lab for up to six months in normal sea water tanks. These advantages make bivalve *Codakia* a good model organism for marine chemoautotrophic symbioses. Another interesting aspect is the ecological significance of seagrass for the global climate (Orth et al., 2006). Lucinidae bivalves were found in 97% of the tropical seagrass sites (van der Heide et al., 2012).
The bivalve *Codakia* is able to adapt to biochemical parameters (i.e. increase pore-water nutrients and decrease sulfide concentration by oxygen introduction and symbiont sulfur oxidation) of the pore water of seagrass sediment and thus bivalves detoxify the habitat of sulfides (Dorman et al., 2002; Reynolds et al., 2007).
A positive correlation between the chemoautotrophic symbionts harboring bivalves and the seagrass *T. testudinum* has already been described (van der Heide et al., 2012). The authors hypothesize that the main reason for the positive correlation is the detoxification of sulfides by the sulfur oxidizing symbiont.

4.1 Habitat characterization of the seagrass bed *T. testudinum*

The burrowing depth of *C. orbiculata* is around two to five cm and five to ten cm for *C. orbicularis*, depending on the length of the burrowing foot. Usually the clams were found and collected in these depths but occasionally bivalves are located deeper in the sediment (personal communication O. Gros). The bivalves permanently pump oxygen saturated surface water. Bivalve can also move inside the sediment to search for the best growth conditions.
Nutrient concentration in the seagrass beds are dependent on the sediment depth and fluctuates very much, e. g. the oxygen concentration (Elisabeth, 2011). Therefore it is necessary to observe the geochemical parameters at the surface of the sediment and up to 10 cm depth.

Table 4.1. Field geochemical data from the seagrass bed *Thalassium testudinum* in Guadeloupe.

	Surface water	Pore water 10 cm	Reference
temperature	25 - 30°C variations over the year		this study
salinity	34‰		this study
oxygen O_2	at daytime up to 200% saturated		(Jackson, 1972)
nitrate NO_3^-	0	0 - 3.5 µM	(Duplessis et al., 2004a)
nitrite NO_2^-	0	0	(Duplessis et al., 2004a)
ammonium NH_4^+	≈ 8 µM	39 µM	(Grimonprez 2014)
hydrogen sulfide H_2S	10 µM	30 µM	(Maurin, 2009)

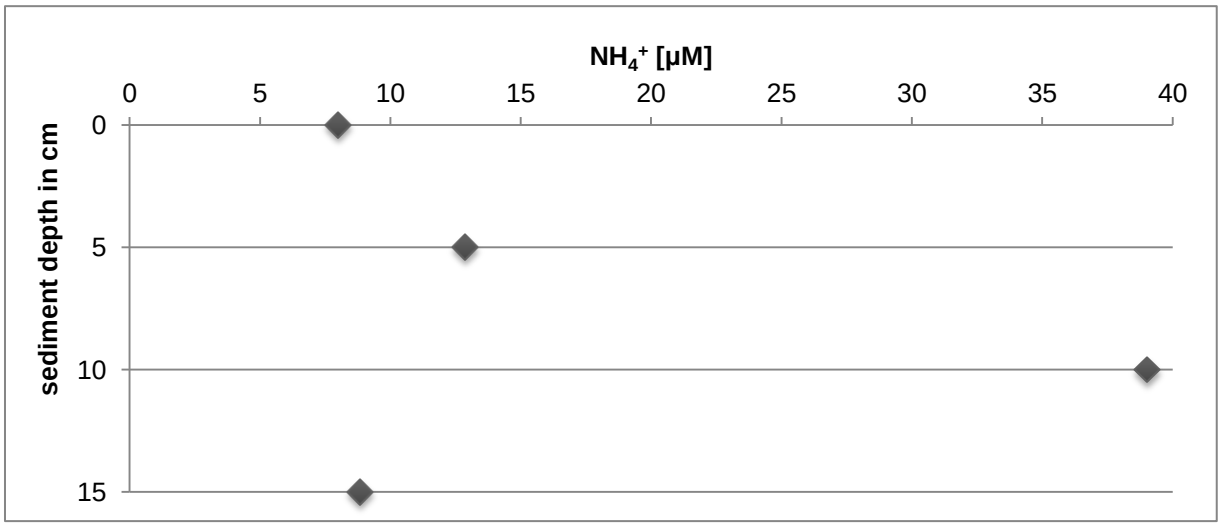

Figure 4.1. Ammonium concentration in the seagrass sediment of *T. testudinum*.
Measured at the "Ilet Cochons" in Guadeloupe at the 20.04.2014, the ammonium measurement was done by Adrien Grimonprez (University des Antilles et de la Guyane).

The geochemical parameters of the seagrass bed in Guadeloupe are summarized in table 4.1. Most of the inorganic nitrogen in the seagrass bed is present in form of ammonia (Touchette et. al, 2000). The seagrass bed close to Guadeloupe is described as a very nitrate poor environment (Duplessis et al., 2004a). At the surface, only low ammonium concentrations were measured while in around 10 cm sediment depth the ammonium concentration is increased to a maximal amount of 39 µM (figure 4.1). Such correlation of the sediment depth and the ammonium concentration is similarly described for a tropical sea grass bed in Indonesia where the concentration fluctuates between 40 µM at the surface and 50-90 µM in 10 cm sediment depth (Erftemeijer et al., 1993; Stapel et al., 1996). In the oxygenic zone, ammonium is oxidized to nitrate by nitrification. In the anoxic zone an anaerobic dissimilatory nitrate reduction (denitrification) can probably produce ammonium (An and Gardner, 2002; Bertics et al., 2010). This could be one reason why the ammonium concentration increases in the anaerobic sediment zone. It seems that nitrogen is biologically available in the habitat of *Codakia* and should be not a limiting growth factor for the symbionts or the bivalves.

The tropical chemoautotrophic symbiont bearing clam *Solemya reidi* use ammonium as nitrogen source. The measured ammonium concentration in the sulfide rich

sediment is 54-64 µM (Lee et al., 1992). Most of the symbioses between chemoautotrophic bacteria and invertebrates have access to ammonium and use this as nitrogen source. In contrast, the *Riftia pachyptila* symbionts assimilate nitrate instead of exogenous ammonium (Lee et al., 1994). At the hydrothermal vent environmental nitrate is available in concentration around 40 µM (Johnson et al., 1988).
In seagrass beds of *Zostera marina* in the northern hemisphere the ammonium concentration in 2 cm rhizopheric sediment depth is 100 to 650 µM. Interestingly, nitrogen fixation was observed independently of the ammonium concentration (McGlathery et al., 1998). The nitrogen fixation was typically higher under low oxygen or anaerob conditions. An interaction between heterotrophic nitrogen-fixing bacteria and the eelgrass is suggested, involving a bidirectional exchange of material (McGlathery et al., 1998), It is shown that the leaves of seagrass are able to take up ammonium (Stapel et al., 1996). The gaseous nitrogen fixation in presence of ammonium is also described in chemoautotrophic communities in cave ecosystems (Desai et al., 2013).

4.2 Genomics

The combined application of different -omics techniques allows a detailed insight into the physiology and interaction of organisms. Genomics is a suitable technique to investigate the potential physiology of bacteria, especially if they are not cultivable. In the marine environment only 0.001-0.1% bacteria are cultiviable (Amann et al., 1995). Symbiotic bacteria can also be found environmentally, however none of them is not cultivable without their host (Baumann et al., 1997; Gros et al., 2003a). An adult aposymbiotic *Codakia* bivalve has never been collected from a seagrass bed, therefore it seems that the symbiont is obligate for the host (personal communication with O. Gros). In the laboratory, it is possible to incubate the bivalve together with the symbiont for a long time and under different growth conditions, like sulfide or oxygen concentration. The extraction of *Codakia* symbionts from the bivalves tissue based on the high density because of the stored sulfur globules. Similar techniques are used to extract related symbiont bearing organims like *R. pachyptila* or *Lucinoma aequizonata* (Distel et al., 1988; Markert et al., 2011).
At the beginning of this thesis very little was known about the physiology of the *Codakia* symbiont (Berg et al., 1984; Duplessis et al., 2004a)(figure 1.8). In this study, the genome of the symbiont in *C. orbiculata* and *C. orbicularis* was sequenced to determine the genomic potential of these bacteria. The sequencing quality of the *C. orbicularis* symbiont genome is with 179 contigs much better than the symbiont genome of *C. orbiculata* with around 904 contigs (table 3.1). Both genomes are still draft genomes and not closed. The genome of the endosymbiont *R. pachyptila* is fragmented in around 184-414 contigs, the genome size is with a predicted size of 3.6-3.7 Mb smaller than that of the *Codakia* symbiont genome with a predicted size of 4.8-4.3 Mb. The GC content also differs 58.4% in the *R. pachyptila* symbiont genome and 52% in the *Codakia* symbiont genome (Gardebrecht et al., 2011).

The symbionts of *C. orbicularis* and *C. orbiculata* were determined as Gammaproteobacteria and chemoautotrophic sulfur oxidizer (Distel et al., 1988). Cross-infection experiments supported the hypothesis that *C. orbicularis* and *C. orbiculata* are colonized by the same and a single bacterial species (Gros et al., 2003b). It was shown that *Lucina pectinella and Linga pensylvanica,* a large symbiont harbouring tropical Lucinidae are also colonized by the same symbiont species (Durand P. et al., 1996). Later studies have demonstrated that the whole family of *Lucinidae* seems to be colonized by only three bacterial symbiont species (Brissac et al., 2011). The 16S rDNA sequence is a common molecular marker for bacterial phylogeny studies. The sequenced 16S rDNAs of *C. orbicularis* and *C. orbiculata* in the present study show a similarity of 98-99% with an NCBI eValue of 0.0 (table 3.2, figure 6.1). The alignment shows that the small differences in the sequences are probably due to sequencing mistakes, which can be found mainly at the end of the sequences. In our case the compared 16S rDNA sequence has a length of around 1000 bp. The 16S rDNA taxonomic classification has resolution limits. On a species level it occurs that the resolution of the 16S rDNA is not sufficient to discriminate between closely related species (Janda et al., 2007). To compare the whole genome, we aligned key genes and used the phylogenetic tool Jspecies, which measures the probability of two genomes belonging to the same species (Richter et al., 2009). In comparison, key genes show a very high similarity (figure 3.2). The calculated probability with Jspecies indicates that both symbionts belong to the same bacterial species (table 3.3). The hypothesis of one symbiont species in both *Codakia* bivalves is furthermore supported by the results of the bidirectional NCBI blast of the encoded proteins in both symbionts genomes (figure 3.2). Most of the encoded proteins (3578 proteins) is homolog in both symbionts and have a similarity above 95%. Only 1551 proteins of *C. orbiculata* and 580 proteins of *C. orbicularis* showed a lower similarity or different blast hits*.* Many of these proteins with low similarities are hypothetical small proteins or/and augmented incomplete sequenced proteins. Nevertheless, both symbiont genomes are still draft genomes and a definite statement concerning distinct genomic differences of both symbiont fractions can only be gained after closing the two genomes.

A phylogenetic analysis of the 16S rDNA of the *Codakia* symbiont showed a close relationship to the endosymbiont of *R. pachyptila* (Distel et al., 1988; Brissac et al., 2011; Imhoff et al., 2003). A comparison of the gene sequences of the adenylylsulfate reductase subunits *aprAB* showed that both symbionts belong to the *aprAII* lineage in the sulfur oxidizing bacteria clade, together with the endosymbiont of *R. pachyptila* (Brissac et al., 2011). We uploaded the genome to the SEED Viewer (http://rast.nmpdr.org/seedviewer.cgi) to have an additional tool to browse the genome. The closest neighbor of the *Codakia* symbiont in the SEED database is *Allochromatium vinosum* DSM 180, a thiosulfate oxidizing anaerobe purple sulfur bacterium in fresh water and marine habitats with a score of 524. The second is *Methylophaga thiooxidants* DMSO10 with a score of 289, a chemolithoheterotrophic halophilic bacterium (figure 6.2).

The environmental transmission requires a free living bacterial stock in the habitat (Bright et al., 2010; Gros et al., 1996). A free living stock of bacteria was found in the

seagrass bed sediment (Gros et al., 2003a). In the free living (life) stage flagellar motility is necessary for movement in the sediment. In the *Codakia* symbiont genome three gene clusters of flagellar genes (table 6.7) and additional chemotaxis genes are encoded. The chemotaxis genes *cheY/Z/A/B/W/R* are part of one flagellar motility gene cluster. The symbiont of *R. pachyptila* also encodes a functional flagellum and genes that are involved in chemotaxis (Robidart et al., 2008). It is believed that the symbiont is, similar to the *Codakia* symbiont, environmentally transmitted (Nussbaumer et al., 2006). In the genome of the *R. pachyptila* endosymbiont genes for a heterotrophic lifestyle, like a phosphotransferase system and ABC transporter, indicate that the symbiont can live autotroph as well heterotroph (Robidart et al., 2008). In the *Codakia* symbiont genome genes for heterotrophic pathways like glycolysis and TCA-cycle are encoded. Furhtermore, several subunits of ABC transporter, TRAP type transporter and a phosphotransferase system were found. The presence of these genes indicates that the *Codakia* symbiont can live autotroph as well as heterotroph.
In general, the endosymbiont genome of *R. pachyptila* shows high similarities to the *Codakia* symbiont. When comparing for example metabolic pathways like the sulfur metabolism, genes of the *R. pachyptila* symbiont are often the closest homolog to the *Codakia* symbiont according to NCBI blast. Further similarities are the environmental transmission, the free living form of the symbiont (both encode flagellar and chemotaxis genes) and the genetic potential to live mixotrophically (Nussbaumer et al., 2004; Robidart et al., 2008).

4.3 Proteomics - Physiology of the endoymbiont

The proteomic approach gives an insight into the main metabolic pathways of the symbiont at a specific time point under certain conditions. The incubation of the host together with the symbiont allows an investigation of the symbiont physiology under different environmental conditions. The physiology of life is based on basic rules, such as: (i) life requires a source of energy and raw material to build up cell mass. (ii) Bacteria need to take up or synthesize amino acids, carbohydrates and lipids.
Berg et al. (1984) found enzyme activities for sulfur oxidation, carbon fixation (via RuBisCO) and nitrate reduction. They assumed that the symbiont oxidizes reduced sulfur compounds, fixes CO_2 and produces organic compounds (Berg et al., 1984). These genes indicate an autotrophic lifestyle, however, genomic data additionally indicate a heterotrophic lifestyle of the symbiont.

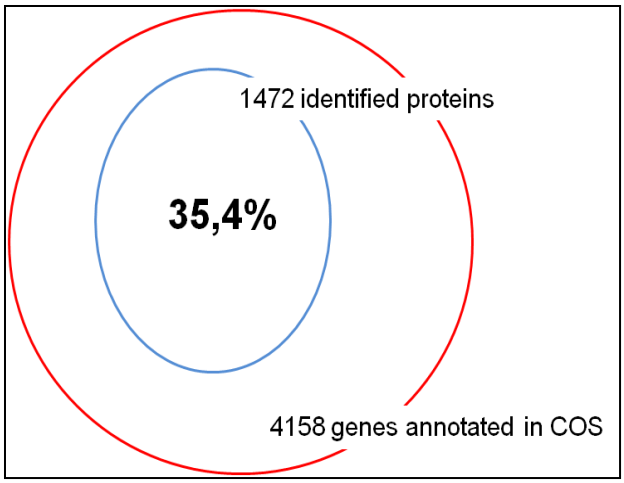

Figure 4.2. Overview about the "sure" identified proteins in the *Codakia* symbiont proteome.
(COS - *C. orbicularis* symbiont)

The proteome analysis of the present study led to the identification of 1472 symbiont proteins, which make up around 35% of all encoded proteins in the *C. orbicularis* symbiont genome. Further studies will probably identify even more *Codakia* symbiont proteins. The use of different proteomic approaches like the enrichment of periplasma proteins or small proteins are conceivable and could increase the number of protein identifications. Proteome analysis identified until now about 24% of the predicted proteins in the endosymbiont *R. pachyptila* (Markert et al., 2011). In *Bacillus subtilis* about 40% of the vegetative proteome has been identified (Eymann et al., 2004).

The direct quantitative comparison of the proteome of *C. orbicularis* and *C. orbiculata* has to be handled with care because both are mobile environmental organisms. Inside the seagrass sediment the geochemical parameters can fluctuate very much (Dufour et al., 2006). It becomes apparent that the gills of different bivalves can look very different. Frequently, the gill is thick and has a lightly beige color, but sometimes the gill can be brown and thin. The color depends on the symbiont population and their content of elementary sulfur therefore it mirrors the feeding situation of the symbionts and the host.

The high amounts of the molecular chaperone GroEL (ID 3782) and co-chaperon GroES are remarkable (ID 3781)(table 3.31). In the *C. orbicularis* symbiont control sample the identified amount of GroEL was higher than the sulfur oxidizing enzymes AprAB. In the endosymbiont of *R. pachyptila* GroEL were identified in the proteome, but in much lower amounts than sulfur oxidizing enzymes. In the deep sea, mussel *Bathymodiolus* the measured amount of GroEL/ES was also very high (personal communication R. Ponnudurai). Chaperones can function as heat shock proteins, to prevent non-functional protein structures and to buffer environmental changes (Fares et al., 2004). In freshly collected bivalves the heat stress level should be low, the water temperature in the seagrass bed is around 25-28°C. It is described that GroEL is also over-expressed in the endosymbiotic bacteria of aphids, GroEL constitutes 10% of the total protein (Baumann et al., 1996). Obligate endosymbionts have a small population size, which could cause the fixation of deleterious mutations. GroEL is predicted to mask the effects of such harmful mutations (Fares et al., 2002). The

aphid endosymbiont *Buchnera* is obligate endosymbiotic and vertically transmitted (Fares et al., 2004), in contrast the *Codakia* symbiont is facultative endosymbiotic and environmental transmitted. GroEL show a high functional flexibility (Fares et al., 2004) and further studies are necessary to explain the high GroEL amount in the *Codakia* symbiont.
Presumably, the physiology of the free living form of the symbiont greatly differs compared to the physiology as endosymbiont inside the gills of the bivalve. To get an insight into the free-living physiology, the symbionts were extracted from the bivalves gills and kept alive for 24 hours. Unfortunately, no free living-specific proteins, like motility proteins, were identified.

4.3.1 Carbon metabolism

The genome encodes the main enzymes for glycolysis, tricarboxylic acid cycle (TCA-cycle), pentose phosphate pathway (PPP) and Calvin Benson Bassham cycle (CBB). The main carbon catabolic pathways, like glycolysis, PPP and TCA-cycle are also necessary for anabolic processes, like amino acid metabolism and for the generation of pre-coursers of the CBB.

4.3.1.1 Carbohydrate metabolism

(1) Glycolysis, gluconeogenesis and glycogenesis

All enzymes of the glycolysis and gluconeogenesis were identified in the *Codakia* symbiont proteome. Only the glucose-6-phosphatase for the last step of the gluconeogenesis is not encoded in the genome. The first step of the glycolysis, the glucokinase, was not identified in the proteome (table 3.20). The enzymes fructose-1,6-bisphosphate aldolase class II (FbaAII, ID 3037), pyruvate kinase (Pyk, ID 3038)), phosphoglycerate kinase (Pgk ID 3039) and glyceraldehyde-3-phosphate dehydrogenase, type I (GapA ID 3040) are organized in a gene cluster. The other genes are widespread on the symbiont genome. The FbaAII (ID 3037) is predicted involved by the CBB but the glycolysis subtype FbaAI (ID 478) were only very less identified in the proteome. A ketose-bisphosphate aldolase (KbaII) was identified in the proteome. The FbaAII is part of a gene cluster for glycolysis genes, probably the FbaAII or the KbaII can undertake the function of the FbaAI in the glycolysis pathway. The enzymes of the glycolysis pathway are less expressed, which makes the exact interpretation of these proteome data difficult. The only enzymes abundantly expressed were part of the CBB.
Glycogen is a polysaccharide of glucose, the main chains are connected by α1→4 glycosidic bonds, the branches are attached by a α1→6 glycosidic linkages. Glycogen is a multi-branched polysaccharide, which is not transportable. In the genome all genes required for the glycogenesis are encoded. Three genes are organized in a gene cluster. The genes required for the glycogen degradation are

widespread over the genome. In the proteome, the proteins for glycogen synthesis and degradation were identified except for the glucokinase (table 3.5). This enzyme catalyzes the activation of glucose, the first step of the glycogen synthesis. One explanation could be that the glucose for the glycogen synthesis is produced by the gluconeogenesis and ends with glucose-6-phosphate and no glucose-6-phosphatase is encoded in the genome. This can be directly converted to glucose-1-phosphate with the phosphoglucomutase. Two different pathways for further activation steps are conceivable with ATP and UTP. For the ATP activation pathway all required enzymes are encoded and identified, the glucose-1-phosphate adenyltransferase (GlgC) and the glycogen synthase (GlgA), for the UTP way only the first enzyme was found in the in the current draft genome, the UTP-glucose-1-phosphate (GalU). The endosymbiont of *R. pachyptila* and *T. jerichonana* encodes only the ATP activation way (Gardebrecht et al., 2011). The glucose-1-phosphate adenylyltransferase GlgC (ID 38), a protein related with glycogen synthese via the ATP-activation way, was detected with the highest amount in the starvation sample with low oxygen conditions (figure 3.18.). Interestingly, the glycogen synthase (GlgA) was not abundantly identified at any condition. Likewise, the glycogen degradation enzyme, the glycogen phosphorylase (GlgP), was identified only once in the proteome. A light microscopy study showed that the symbionts are positive in the periodic acid-schiff reaction, which stains glycogen particles (figure 4.3) (unpublished data of O. Gros). This supports the genomic and proteomic data, which clearly indicate that the symbionts use glycogen as a storage compound. In the micoroscopic picture it is visible that the content of glycogen in the symbiont population differs (reflected by the intensity of the staining). Symbionts close to the space of two adjacent filaments are less strong stained than symbionts deeper in the bacteriocytes. Caro et al. (2007) found up to seven sub-populations of symbionts in one gill, differ in size, genome copies and amount of elemental sulfur. It is possible that they also differ in their glycogen content.

It seems that the glycolysis and gluconeogenesis pathways do not play a key role in the physiology of the symbiont inside the gills of freshly collected bivalves. The symbiont requires the energy from the oxidation of reduced sulfur compounds, Furthermore the symbiont needs these pathways to produce precursors for anabolic processes, like amino acids. The proteomic data indicate that the autotrophic enzymes of the glycolysis, PPP, TCA-cycle and the glyoxylate way are more expressed during starvation conditions. But no indications for an increased glycogen degradation was found under starvation conditions, rather a glycogen synthesis protein is up-regulated under starvation conditions. In *R. pachyptila* host and symbiont contain glycogen granules (Sorgo et al., 2002). For the symbiont of *R. pachyptila* it is assumed that under energy limiting conditions the symbiont can either grow heterotrophically, possibly by the degradation of glycogen (Markert, 2008), or mixotrophically meaning autotroph and heterotroph at the same time (Robidart et al., 2008). In contrast, the symbiont genome of *Calyptogena magnifica,* a chemoautotrophic sulfur oxidizer encodes all glycolytic enzymes, the non oxidative branch of the PPP and the TCA-cycle enzymes, except for the a-ketoglutarate dehydrogenase (Newton et al., 2007). This indicates an obligate autotrophic lifestyle

of the symbiont (Wood et al., 2004). For the *Codakia* symbiont the genomic and proteomic data suggest a mixotrophically lifestyle even as endosymbiont.
Codakia can survive low oxygen conditions (0.5 mg/l O_2) for more than five weeks (this study). The lethality was not dependent on the oxygen concentration during the incubation (personal observations). The bivalve needs oxygen as terminal electron acceptor for the oxidative phosphorylation, just as the symbiont both compete for oxygen (Girguis, 2006; Childress et al., 2011). Glycogen can be used to survive temporary anoxic conditions, as glycogen degradation can produce energy via the substrate-linked phosphorylation (Brayant, 1991; Arndt, 1998). One hypothesis is that under low oxygen conditions the symbiont decreases the sulfur oxidation and uses glycogen as energy source. A second hypothesis is that the host decreases the oxidative phosphorylation and uses instead glycogen for substrate-linked phosphorylation. The utilization of glycogen can reduce the oxygen competition between host and symbiont. In this respect it is interesting to note that the phosphotransferase system (PTS) is encoded but little expressed. The PTS translocates hexose sugars, a degradation product of glycogen (table XX). A PTS is also encoded in the endosymbiont of *R. pachyptila* genome (Robidart et al., 2008). The possibility of an interaction and glucose exchange is described below. Alternatively, the symbiont could overcome low oxygen conditions by the reduction of elementary sulfur or nitrate, which will be discussed later.

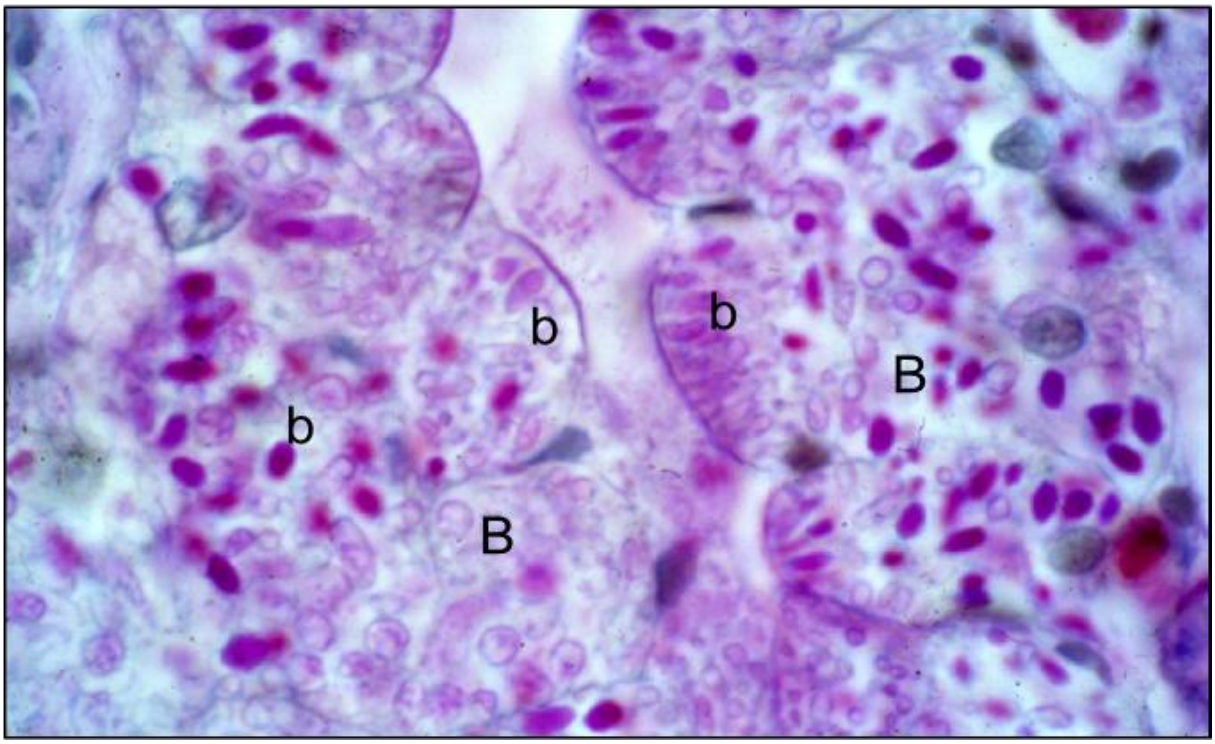

Figure 4.3. Light micrograph of freshly gills from *C. orbicularis* with stained glycogen particles.
Glycogen particle in the gills of *Codakia* were stained with the periodic acid schiff staining. The bacterial endosymbionts closed to the basal pole of the bacteriocytes are purple indicating a high content in glycogen-like particles while the cytoplasm of the bacteriocyte appeared devoid of such carbon storage (picture O. Gros). (b - symbiont, B - bacteriocytes,)

(2) Tricarboxylic acid cycle (TCA-cycle) / reductive tricarboxylic acid cycle (rTCA)

The TCA-cycle is used in heterotrophic organisms to generate energy via the respiratory chain. Both autotrophic and also heterotrophic organisms need this pathway to produce pre-cursors for anabolic processes, like amino acids or nucleotides.
All TCA-cycle proteins are encoded in the genome and were identified in the proteome of the *Codakia* symbionts. Additionally, the isocitrate-lyase (AceA), belonging to the glyoxylate pathway, was identified. However, no malate synthase (AceB) was identified in the symbiont proteome. The difficulty in the identification of the AceB was not unexpected, as it is already described in several studies of *Bacillus* (Schroeter et al., 2011; Kabisch et al., 2013;). The enzymes of the TCA-cycle are only low expressed in the symbiont proteome, but present in the endosymbiontic form of the symbiont. For the endosymbiont of *R. pachyptila* it is assumed that the heterotrophic TCA-cycle is only present in the free living form, the symbiont however uses the reductive TCA-cycle for carbon fixation (Robidart et al., 2008). The low amount of TCA-cycle proteins indicate that the TCA-cycle is not necessary for further metabolic conversion of the exchange compounds to the host. In good feeding conditions the TCA-cycle is probably only used for anabolic pre-cursors production.
Under sulfide starvation conditions the TCA-cycle was up-regulated. The symbiont may use the TCA-cycle also for energy generation via the degradation of stored glycogen.
In the *Codakia* symbiont genome only weak indications were found for the reductive TCA-cycle (table 3.6). It seems that the rTCA-cycle pathway does not exist in the *Codakia* symbiont, instead the symbiont fixes carbon dioxide only via the CBB. This is different to the endosymbiont of *R. pachyptila* which fixes CO_2 with the CBB and the rTCA-cycle (Markert et al., 2007). Proteomic data revealed the importance of both pathways in the endosymbiont of *R. pachyptila* (Markert et al., 2011). The symbiont of *C. magnifica* also uses only the CBB for carbon fixation (Newton et al., 2007).

(3) Pentose Phosphate Pathway (PPP)

The PPP is an important pathway to generate NADP(H) and for the conversion of different phosphate sugars (figure 3.21). It is divided into an oxidative branch which is irreversible and a reversible non-oxidative branch. The enzymes for the oxidative part are encoded in the genome but at the proteome level never or only less identified (table 3.22). How the symbiont avoids these NADP(H) generating steps is not known. The symbiont of *Calyptogena magnifica* encodes only the non-oxidative branch of the PPP (Newton et al., 2007).
The enzymes of the non-oxidative branch are not completely encoded in the genome, no homolog of a transaldolase was found in the *Codakia* symbiont genome. Interestingly the *R. pachyptila* endosymbiont also does not encode transaldolase in the genome (Gardebrecht et al., 2011). Further studies are necessary to clarify how the symbiont can avoid the carbohydrate transforming steps of the transaldolase. Nevertheless, the remaining enzymes of the non-oxidative way are abundantly

expressed in the symbiont's proteome. Thus, it seems that this part of the pathway plays an important role in the symbiont metabolism. A part of the non-oxidative PPP is also involved in the regeneration of ribulose-1,5-bisphosphate, the substrate for RuBisCO. In the *R. pachyptila* endosymbiont both parts of the pathway were identified in the proteome (Markert et al., 2011).

4.3.1.2 Calvin Benson Bassham Cycle (CBB)

The key enzyme of the CBB is the ribulase-1,5-bisphosphate carboxylase/oxygenase (RuBisCO). This enzyme belongs to the most abundant enzymes in plants and bacteria (Schneider et al., 1992). The enzyme is big, slow and also catalyzes the oxygenase reaction that causes a loss of efficiency by photorespiration (Schneider *et al.*, 1992).
In the genome of the *Codakia* symbiont a cluster with eight CBB gens was found (figure 3.22). In the cluster two different forms of RuBisCO are encoded, as well as its transcriptional regulator CbbR at the beginning of the cluster in the opposite transcription direction, which has never been identified in the proteome. The RuBisCO protein (ID 313) of the *Codakia* symbiont show high similarity to the RuBisCO form I of *Allochromatium vinosum* DSM 180. The second RuBisCO form II (ID 319) shows high similarity to the RuBisCO protein of the endosymbiont of *T. jerichonana*. A related genome organization was found in Proteobacteria, which also encode various forms of RuBisCO (Badger et al., 2007). It is assumed that several forms of RuBisCO in one organism could increase the flexibility to adapt to different conditions, like O_2 or CO_2 concentrations (Badger et al., 2007). Surprisingly, both RuBisCO forms (RbcL/S1, RbcL2) were indentified in the proteome under control conditions. However, form I of RuBisCO is expressed in much higher amounts and is one of the most abundant proteins in the control samples. This indicates that form I is mainly responsible for the CO_2 fixation of the symbiont in the endosymbiont lifestyle. The high amounts of RuBisCO enzymes indicate the importance of the CO_2 fixation, probably to feed the bivalve with organic compounds. In the *R. pachyptila* endosymbiont the RuBisCO protein is also among the most abundant proteins on the master-gel (Markert, 2008). In case of starvation of the *Codakia* symbionts with saturated oxygen concentration, form I RuBisCO was down-regulated, while levels of form II RuBisCO did not significantly change. Under starvation conditions the CBB proteins are down-regulated, but still present in high amounts (table 3.23), indicating the importance of this pathway. Alternatively, the symbiont may acquire the energy, needed for the CO_2 fixation, from the storage of elementrary sulfur and/or glycogen.
There are four forms of RuBisCO known. Form I, II and III catalyze the reaction of carboxylation and oxygenation of ribulose-1,5-bisphosphate, while form IV, called RuBisCO-like protein, is incapable of CO_2 fixation (Tabita, 1999). Out of the four forms of RuBisCO, form I is the most abundant enzyme (Tabita et al., 2007). It is found in bacteria and eukaryotes, predominantly in photosynthetic and aerobic chemolithoautotroph organisms (Elsaied et al., 2001). RuBisCO form I and II differ in their subunit structures, affinity to CO_2 and O_2 (Badger et al., 2007). RuBisCO form I

consists of eight large and eight small subunits (Schneider et al., 1992). The RuBisCO form I group is divided in four different clades, which were all found in Proteobacteria (Badger et al., 2007). The form I enzyme has a higher affinity to CO_2 and distinguishes better between CO_2 and O_2 compared to form II RuBisCO (Tabita, 1999; Tabita et al., 2008). The form II RuBisCO contains only a large subunit and shares approximately 30% sequence identity with the large subunit of form I (Tabita *et al.*, 2007). Both RuBisCO forms of the *Codakia* symbiont have an identity of 32%. Several bacteria produce both RuBisCO forms I and II (Hayashi et al., 1998; Shively et al., 1998; Heinhorst et al., 2002). RuBisCO II is often found in organisms that also contain RuBisCO I (Tabita et al., 2008). It is assumed that in these organisms the function of form II RuBisCO is not primarily the fixation of CO_2. Rather form II RuBisCO, together with other enzymes of the CBB pathway, makes use of CO_2 as an electron acceptor to balance the intracellular redox potential (Dubbs et al, 2004; Tabita et al., 2008). This could explain why the expression of form II RuBisCO is not coupled to the feeding conditions of the symbiont, it is not down-regulated under starvation conditions. In organisms with two different RuBisCO forms, form I is expressed when CO_2 is limiting, because of its higher affinity to CO_2 compared to form II (Tabita, 1999). In the endosymbiont of *R. pachyptila* only RuBisCO form II is encoded and expressed in very high amounts, which is not surprising, as the CO_2 concentration is extremely high in the tubeworm, due to the hemoglobin (Robinson et al., 1998; Markert et al., 2007). It can be assumed that the CO_2 concentration is lower in the *Codakia* bivalve as no hemoglobin protein is described in the gills of *Codakia* bivalve. The bivalve provides the substrates through permanently pumping surface water. Further analysis of the RuBisCO genes would improve our knowledge about the two encoded and expressed RuBisCO forms in the *Codakia* symbiont.

Down-stream of each RuBisCO form two supporting/activation proteins are encoded CbbQ (ID 315/320) and CbbO (ID 316/321). But only one set of activation proteins is expressed whereas both RuBisCO forms are expressed. Only the activator proteins CbbO and CbbQ in close proximity to the form I RuBisCO are expressed in the *Codakia* symbiont while the activation proteins CbbQ and CbbO (ID 320/321) close to form II RuBisCO were never identified. The similarity between the two CbbQ proteins is with 71% identity quite high, while the CbbO proteins share only 36% similarity. Both CbbQ proteins carry an ATP-binding motif, whereas the CbbO proteins do not contain a conserved protein domain. Functional studies of these proteins in *Pseudomonas hydrogenothermophila* showed that the RuBisCO protein is unstable and had a low activity without these genes (Hayashi et al., 1997). This indicates that both proteins, CbbQ and CbbO, posttranslationally activate RuBisCO (Hayashi et al., 1997). The presence of two sets of activator proteins could indicate that RuBisCO II plays a more important role for the CO_2 fixation in the free living form. Alternatively, form I activator proteins could also activate RuBisCO form II.

In the genome two copies of the phosphofructokinase PfkA (ID 984/1107) are encoded, with low similarity to each other (26% identity). One protein shows high similarity to the pyrophosphatase-depending PfkA (PP-PfkA) of *Candidatus Ruthia magnifica* (79% identity NCBI protein blast), which was the first sequenced chemoautotrophic sulfur oxidizing endosymbiont from the deep sea clam

Calyptogena magnifica (Newton et al., 2007)(table 3.23). The PP-PfkA is more expressed in the *Codakia* symbiont than the ATP-dependent PfkA. This hints at different physiological functions in the symbiont. The PP-PfkA might play a key role in the CO_2 fixation. The second phosphofructokinase is ATP dependent and likely needed for the glycolysis. A possible reason could be a different regulation of the two distinct pathways., Furthermore, the efficiency for a particular reaction might differ.
In the *Codakia* symbiont one enzyme, the sedoheptulose-1,7-bisphosphatase involved to the regeneration of ribulose-1,5-bisphosphate is missing in the genome. In the genome of the *Codakia* symbiont two phosphofructokinase are encoded with low similarity. One of them shows high similarity (79% identity) to PP-PfkA of the symbiont of *Calyptogena magnifica*, *Candidatus Ruthia magnifica*. It is assumed that the PP-PfkA catalyzes three reactions shown in figure 4.5 and replaces the missing enzymes, which is an unconventional CBB (Newton et al., 2007). The predicted PP-PfkA is abundantly found in the symbiont proteome, while the ATP-PfkA, related to the glycolysis is less identified. Due to the production of energy-rich pyrophosphate, the unconventional regeneration of ribulose-1,5-bisphosphate is more energy-efficient in combination with the pyrophosphate energized proton pump HppA (figure 4.6) (Kleiner et al., 2012). The *PP-pfkA* and *hppA* genes are co-localized in the genome of *R. magnifica* (Reshetnikov et al., 2008). This is also the case in the genome of the *Codakia* symbiont (figure 3.23), indicating a close metabolic relationship between these two enzymes (Kleiner et al., 2012). The author revealed the presence of the PP-PfkA in two endosymbionts of *Olavius algarvensis*, a worm without an intestine, which lives in association with three endosymbionts in coarse grained coastal sediments of the island Elba, Italy. This is a related habitat to the seagrass bed of the bivalve *Codakia.* In figure 4.4 an overview of the predicted CBB pathway is shown.

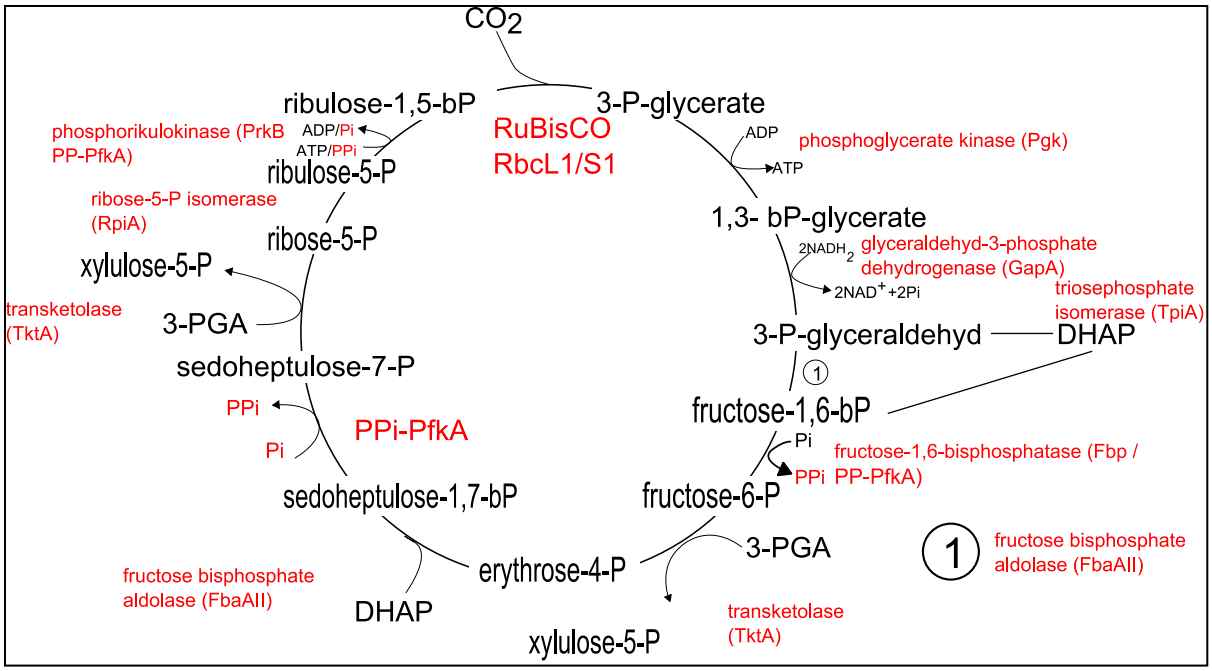

Figure 4.4. Predicted unconventional CBB in the *Codakia* symbiont.
The unconventional CBB use the PP-dep. phosphofructokinase PP-PfkA to replace missing enzymes in the symbiont (red marked enzymes are encoded in the genome and identified in the proteome). (adapted to Markert et al. 2011)

1. fructose-1,6-bisP → fructose-6-P + PPi
2. sedoheptulose -1,7-bisP → sedoheptulose-7-P + PPi
3. ribulose-5-P +PPi → ribulose-1,5-bisP

Figure 4.5. Overview about the predicted catalyzed reaction of the PP-PfkA of *R. magnifica*.
(Newton et al., 2007) (PP-PfkA - PP-dep. phosphofructokinase)

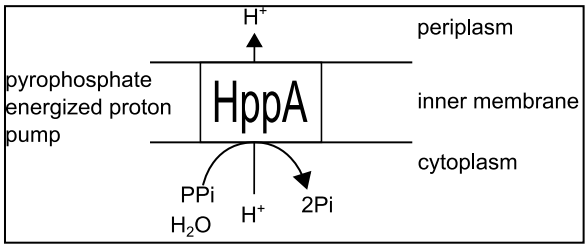

Figure 4.6. Schematic function of the HppA encoded in the symbiont genome.
(modified from Markert et al., 2011) (HppA - pyrophosphate energized proton pump)

4.3.2 Sulfur metabolism

The oxidation of reduced sulfur compounds for energy generation is ubiquitary in soil and water and only performed by prokaryotes (Friedrich et al., 2005). The sulfur oxidizing bacteria are phylogenetically diverse and can be found in a wide range of habitats (Friedrich et al., 2001). Sulfur occurs in the oxidation states from -2 (sulfide) to +6 (sulfate) (figure 1.8). The sulfur oxidizer use the electrons for energy transformation in the respiratory chain and use e. g. oxygen as terminal acceptor.

Oxidation produces the energy for the fixation of carbon dioxide via the CBB. Therefore, sulfur metabolism is an essential and characteristic pathway of the *Codakia* symbiont. Considering, that free sulfite is highly reactive and therefore a potentially toxic compound (Dorman et al., 2002), organisms have to avoid free sulfur in the cytoplasm. Stored sulfur globules are enveloped mostly with a protein layer (Prange et al., 2004). In related organisms it is described that they contain at least two sulfite oxidizing pathways to avoid free sulfur in the cytoplasm (Kappler et al., 2001).

The sulfide or hydrogen sulfide is oxidized to polysulfide or elemental sulfur. For the initial step of the sulfide oxidation in the periplasm, mainly three systems are described in literature: (i) the periplasmic sulfide dehydrogenase FccAB, (ii) the sulfide quinone reductase Sqr and (iii) the periplasmic sulfur oxidation system (Sox) (Frigaard et al., 2009).

Eight genes in the *Codakia* symbiont have a predicted function as sulfide dehydrogenase (*fccAB*) (table 6.1), but only two were identified in the proteome analysis. All *fccAB* candidates are co-localized in the genome. The annotation is difficult because the conserved domains are not specific enough for functional analysis *In silico*. FccA contains a 21 kDa cytochrome domain and FccB has a size of 46 kDa and contains a flavocytochrome c subunit. The FccA uses a cytochrome c as electron acceptor for the oxidation of sulfide to sulfur or polysulfide *in vitro*. In *A. vinosum* FccAB mutants catalyze sulfide oxidation with similar rates to the wild type (Weissgerber et al., 2011).

A putative sulfide:quinone reductase (Sqr, ID 2280) is encoded in the genome of the *Codakia* symbiont and with high similarity to the *T. jerichonana* endosymbiont (90% identity)(Gardebrecht et al., 2011). But this candidate shows no similarity to the sulfide:quinone reductase from *Rhodobacter capsulatus*, which is well described (Shahak, 1997). In the *Codakia* symbiont proteome the Sqr is only three times identified. The Sqr oxidizes sulfide and mediates the electrons to the quinone pool (Reinartz et al., 1998). As the putative Sqr is not abundantly expressed, it seems that Sqr does not play an important role in the first step of the sulfide oxidation in the periplasm, in case the function of the Sqr is predicted correctly. The oxidation of sulfide compounds generates elementary sulfur globules (Dahl et al., 2006). Elementary sulfur is unstable, highly reactive and insolube in aqueous solutions and mainly occurs in rings of mostly eight atoms (S8) (Griesbeck et al., 2000). Therefore, the intracellular sulfur globules in *A. vinosum* and in the endosymbiont of *R. pachyptila* are enveloped by three different proteins: SgpA, SgpB, SgpC (Prange et al., 2004; Markert et al., 2011). For *A. vinosum* it is described that these proteins are essential for the formation of sulfur globules (Prange et al., 2004). In the *Codakia* symbiont genome a gene with low similarity to SgpA protein was found ($4e^{-14}$/ 44% identity). Nevertheless, these proteins are small, making an exact prediction of the function of the proteins via blast difficult. On a genome level, only weak indications were found that the elementary sulfur globules are enveloped by a protein layer related to *A. vinosum*. However, transmission electron microscopy pictures of the symbiont suggest that the globules are enveloped, visible as membrane-like vesicles in the periplasmic space (personal communication of O. Gros). A SgpBC double

mutant in *A. vinosum* was unable to grow on sulfide and could not form sulfur globules. It seems that these three Sgp proteins are essential for the formation and deposition of intracellular sulfur globules (Prange et al., 2004). Further studies are necessary to improve the knowledge on how the symbiont of *Codakia* handles the elementary sulfur storage.

The Sox sulfur oxidation multi-enzyme complex is a periplasmic thiosulfate-oxidizing complex, first described in *Paracoccus versutus* (Lu et al., 1985) and later found in several chemotrophic and phototrophic bacteria of different genera (Friedrich et al., 2008). Thiosulfate is one of the most abundant forms of reduced sulfur compounds in nature and rather stable. Thiosulfate can be used by phylogenetically diverse groups, however in some cases, closely related bacteria can or cannot use thiosulfate (Ogawa et al., 2008). If the symbionts of *Codakia* can utilize thiosulfate is not known. Thiosulfate utilization is shown for the endosymbiont of *Solemya velum* (Stewart and Cavanaugh, 2006b). Two different pathways for the thiosulfate oxidation are described: (I) the thiosulfate dehydrogenase and (II) the Sox multi-enzyme complex, which is also called the thiosulfate multi-enzyme complex (TOMES) (Kelly et al., 1997). Some bacteria have both pathways (Ogawa et al., 2008). In the *Codakia* symbiont genome a candidate of a thiosulfate dehydrogenase was annotated but with only low probability (35% identity, ID 2767). The presence of the Sox-system (described below) suggests that thiosulfate utilization should not be excluded, however further studies are necessary to proof this.

The stored sulfur needs to be transported to the cytoplasm for further oxidation via a up to now unknown mechanism. In *C. tepidum* it is suggested that the polysulfide reductase (PsrABC) is responsible for this translocation (Eisen et al., 2002), which is also assumed for the endosymbiont of *R. pachyptila* (Markert et al., 2011). No homolog genes were identified in the genome of the *Codakia* symbiont genome. An putative alternative pathway for this is discussed in the Dsr-chapter.

Nevertheless, further oxidation of stored elementary sulfur occurs in the cytoplasm. In the *Codakia* symbiont genome a cluster with 15 genes of the dissimilatory sulfite reduction system (Dsr) is encoded and 11 of them were identified in the symbiont proteome (table 6.4). The Dsr-complex will be discussed below. The last step of the sulfur oxidation is the indirect oxidation via adenosine-5-phosphosulfate (APS). This step is catalyzed by the APS reductase AprAB (ID 3699/3700) and the sulfate adenylyltransferase Sat / ATP sulfurylase SopT (ID 145) with an ATP yield through substrate level phosphorylation (Parey et al., 2013) (table 6.3). The APS reductase (AprAB) catalyzes the oxidative binding of HSO_3^- to adenosinemonophosphate (AMP) to generate adenosinephosphosulfate (APS). This reaction yields two electrons, the proposed electron donor is the quinone interacting membrane bound oxidoreductase complex (QmoABC complex) (Pires et al., 2003;Ramos, 2012) (Krumholz et al., 2013). This complex shares homologies with the heterosulfide reductase. In the *Codakia* symbiont putative *qmoABC* genes (ID 3710/3711/3712) are encoded and abundantly identified in the proteome. These genes are annotated as heterodisulfide reductase. The proteins show a quite high similarity to QmoABC proteins of different sulfate reducing bacteria. In *Desulfovibrio desulfuricans* the QmoC contains six TMHs (Pires et al., 2003), whereas the homolog gene in the

Codakia symbiont does not contain TMHs (table 3.9). Pires et al. hypothesized that the electrons from the adenylsulfate reductase AprAB are transferred to the quinone-interacting membrane bound oxidoreductase (Qmo) complex and feed into the respiratory chain, which as confirmed by later studies (Krumholz et al., 2013). The *aprAB* (ID 3699/3700) genes are located closely to the putative *qmo* complex genes in the *Codakia* symbiont genome, which could hint at a coordinate expression and metabolic interaction (Pires et al., 2003). In the next step catalyzed by ATP sulfurylase (SopT ID 145), the sulfate is replaced by an inorganic pyrophosphate leading to ATP formation. The AprAB protein is one of the most abundant proteins in the *Codakia* symbiont, the SopT is less expressed, however identified in almost every sample (figure 3.26). Both enzymes are down-regulated under starvation conditions. The comparison of different oxygen concentrations revealed that both enzymes are down-regulated under low oxygen conditions, despite of an energy yield with substrate level phosphorylation (figure 3.46).

The *aprAB* genes are key genes of the dissimilatory sulfur oxidation. In the *aprAB* based phylogenetic tree, organisms can be divided into two groups, *apr* lineage I and II. *Codakia* is part of the *apr* lineage II like the endosymbiont of *R. pachyptila* (Brissac et al., 2011; Meyer et al., 2007). To remove the sulfate from the cytoplasm two high affinity sulfate transporter SulP (ID 1792, 2768) are encoded in the *Codakia* symbiont genome and one of them is identified in the proteome.

Furthermore, it is shown that the elementary sulfur works as energy storage in case of sulfide starvation (Lechaire et al., 2008). After one week sulfide starvation, sulfur globules were still visible in the light microscope (personal observations). The sulfur metabolism proteins were down-regulated after one week of sulfide starvation but still quite highly expressed (figure 3.45, 3.47). This implies that the symbiont still oxidizes sulfide after one week of sulfide starvation. The data support the hypothesis of the utilization of elementary sulfur as energy source. The lowest amount of sulfur oxidizing proteins were measured in low oxygen conditions, this indicate that the sulfur metabolism is dependent by oxygen and oxygen was limited. However no indications of nitrate use of alternative electron acceptor were observed.

The bivalve *Codakia* is living at the interface between oxic and anoxic conditions. This implies, that temporary no oxygen or only low oxygen concentrations are available. Under low oxygen conditions two proteins, SoxL and SoxB, were significantly up-regulated. For *C. orbicularis* Duplessis (2004a) showed that the symbiont utilizes the elementary sulfur storage as an electron sink to generate hydrogen sulfide during periods of anoxia. One candidate catalyzing this reaction could be the sulfide dehydrogenase (FccAB). The putative FccB is up-regulated under high oxygen conditions, this indicate that the FccAB is not involved of the hydrogen sulfide production. The anaerobic sulfide production is observed for instance in *R. pachyptila* or *Lucinoma aequizonata,* a endosymbiont bearing lucinid bivalve (Arndt et al., 2001).

A predicted overview about the sulfur metabolism in the *Codakia* symbiont is shown in figure 4.7. This overview is adapted to *A. vinosum*, which is an anaerobic phototrophic sulfur oxidizing bacterium belonging to the *Chromatiaceae.* The sulfur

metabolism of *A. vinosum* is probably one of the best characterized of all phototrophic sulfur oxidizer (Sander et al., 2009).

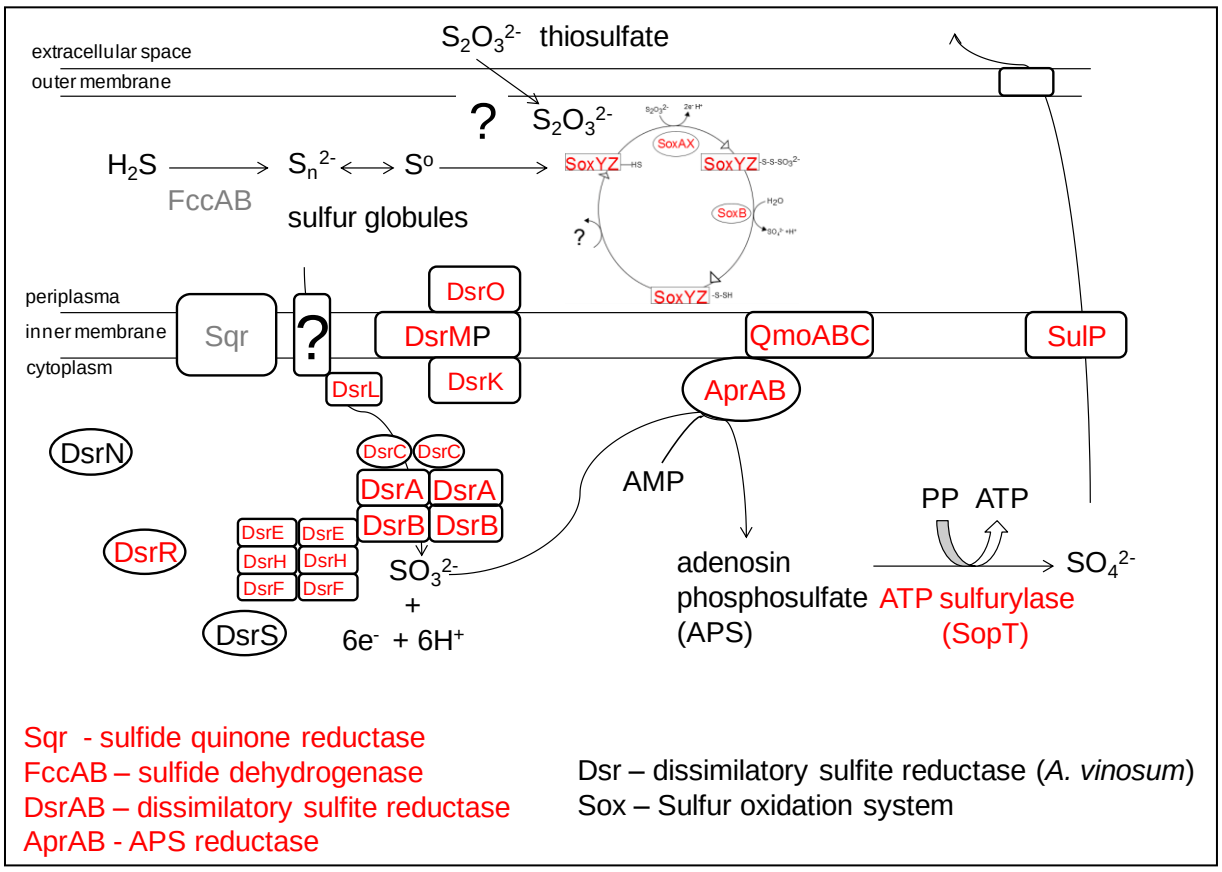

Figure 4.7. Overview about the assumed sulfur metabolism in the *Codakia* symbiont.
The grey marked FccAB and Sqr are annotated just as putative FccAB and Sqr proteins. (figure modified according to Markert *et al.* (2011) and Dahl et al. (2005)) (red marked proteins are identified in the *Codakia* proteome, the black ones are encoded in the genome but not identified in the proteome, grey - putative enzyme activity).

4.3.2.1 Sox system - Sulfur oxidation

The *Codakia* symbiont genome encodes several genes related to a Sox system, but no homologs of the *soxCD genes* were found (figure 3.25). The Sox multi-enzyme complex is located in the periplasmic space, therefore all of the Sox proteins need a signal sequence for the Sec or Tat dependent secretory pathway (Friedrich et al., 2005). The predicted localization of the Sox proteins in the *Codakia* symbiont estimated by the localization tool "Cello" is diverse between cytoplasm and periplasm.
Thiosulfate oxidizing bacteria with the Sox-system can be divided into two groups: (I) sulfur globule as intermediates can be formed, for example in case of *A. vinosum*, or (II) no sulfur globules are formed, for example *P. pantotrophus* (Hensen et al., 2006; Welte et al., 2009). Both groups can be distinguished by the presence or absence of SoxCD, which is absent in the sulfur globule forming group (Hensen et al., 2006). In the *Codakia* symbiont genome no homolog of SoxCD was found and it is further

known that the symbiont creates sulfur globules (Lechaire et al., 2008). In most cases the Sox genes are located in close proximity on the genome, indicating a close interaction of the gene products (Miyake et al., 2007). In *P. panthrophus* the Sox system compromises at least two transcriptional units with 15 genes (Friedrich et al., 2001). But only SoxZY, SoxAX, SoxB and SoxCD catalyze the oxidation of inorganic reduced sulfur compounds, like thiosulfate, sulfite, sulfur and hydrogen sulfide (Rother et al., 2001). SoxCD functions as a sulfur dehydrogenase and was described to be essential for the thiosulfate oxidation in *P. pantrotophus* (Rother et al., 2001). In *A. vinosum*, the SoxCD is also missing, however a thiosulfate utilization is described. The activity of SoxCD might be replaced by SoxL a rhodanese like protein and predicted sulfur transferase (Weissgerber et al., 2011), a homolog protein (SoxL ID 773) is encoded and abundant expressed in the Codakia symbiont.
In contrast, in *A. vinosum* the *sox* genes are separated into two independent gene clusters (Weissgerber et al., 2011). In the *Codakia* symbiont genome the *sox* genes are also separated in several gene clusters (table 3.25). Most of the annotated Sox genes show high similarity to *A. vinosum* DSM 180. In the *Codakia* symbiont genome two copies of SoxZY are encoded, both copies show low similarity to each other and only one copy was identified in the proteome. SoxYZ function as a sulfur compound binding protein, the SoxAX protein is a heme protein (Dahl and Friedrich, 2008). However, the genome organization of the *sox* genes is identical to the endosymbionts of *R. pachyptila* and *T. jerichonana* (figure 4.8). Interestingly, one copy of the *soxXY* genes in the *Codakia* symbiont genome being expressed is close to the *dsr* gene cluster. This is also the case in the endosymbiont of *T. jerichonana* (Gardebrecht et al., 2011).
In total seven Sox proteins were identified (SoxZYXA(XA)LB), four of them only in the sample after one week starvation with high and low oxygen conditions (SoxZYA(XA)). These genes were up-regulated in saturated oxygen conditions. The sulfur transferase and the SoxB protein were identified in almost every measured sample.
It seems that the Sox system under natural sulfide-rich conditions is not an important sulfur-oxidizing pathway. This appears to change in case of sulfide starvation, as the expression of these genes increased under these conditions. This data indicates an involvement of the Sox proteins in the sulfur globules utilization, as during the sulfide starvation the elementary sulfur is the only source of sulfides. Only two proteins, SoxL and Sox are up-regulated under low oxygen condition. This implies that the Sox proteins are probably not involved in hydrogen sulfide production from elementary sulfur. Further studies of the Sox system in the *Codakia* symbiont are necessary to verify the exact function and the interaction with the other sulfur oxidation systems in the *Codakia* symbiont to explain the increase of the expression of these proteins under sulfide starvation.

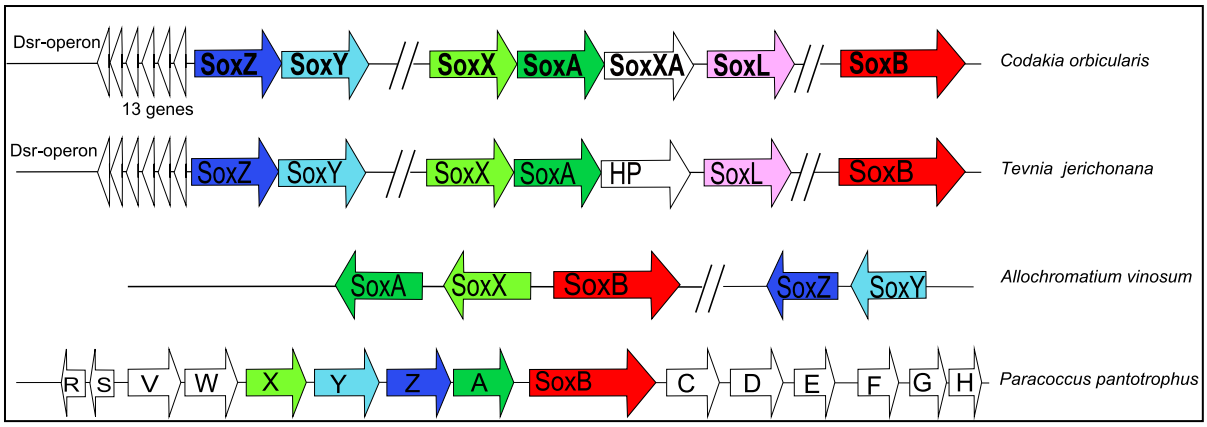

Figure 4.8. Schematic map of the *sox* gene clusters in the *Codakia* symbiont genome compared to other known *sox* gene clusters in sulfur oxidizer.Predicted homologous proteins are indicated by the same color and label. (bold marked labels are identified in the *Codakia* symbiont proteome (figure adapted from Friedrich et al., 2001)

4.3.2.2 Dsr-complex - dissimilatory sulfide reductase complex

The encoded *dsr* gene cluster in the *Codakia* genome is shown in figure 4.9, table 6.4, showing that one *dsrC* gene is encoded in the opposite direction compared to the other *dsr* genes. The *dsrC* genes in close proximity to the other *dsr* genes and in the opposite coding direction, can in this form only be found in the endosymbiont of *R. pachyptila* (Markert et al., 2011). The in *Codakia* symbiont encoded Dsr complex show a very high similarity to the endosymbiont of *T. jerichonana* and to *A. vinosum*. In *A. vinosum* the anaerobic Dsr sulfur oxidation is very well described by the pioneering work of Dahl and co. workers (Pott and Dahl, 1998; Reinartz et al., 1998). Noticeable is that the gene of *dsrJ* is missing in the *Codakia* symbiont *dsr* gene cluster, while all compared organisms show a *dsrJ* gene. In the proteome 11 Dsr proteins, DsrABCEFHLMKOR were found, only DsrP, DsrN and DsrS were not identified (table 3.26). In the control sample the highest amount of Dsr proteins was detected. Under starvation conditions with different oxygen concentrations, all identified proteins were up-regulated under high oxygen conditions.
DsrL is assumed to be involved in the electron transfer and the reductive release of sulfide (Dahl et al., 2005; Lübbe et al., 2006). The encoded DsrL gene shows similarities to the small subunit of the NADPH dependent glutamate synthase. It is possible that DsrL provides glutamine for the siroheme amidation, but it appears not to be involved in the siroamide biosynthesis. However experimental evidence for such functions of DsrL is until now lacking (Dahl et al., 2005). It is assumed that the sulfur is reductively activated and transported to the cytoplasm via a perthiolic organic carrier molecule. The predicted function is an oxidoreductase activity with an NAD(P)H:acceptor, which is essential for the sulfur oxidation in *A. vinosum* (Lübbe et al., 2006; Grimm et al., 2008). DsrS and DsrN are predicted to be cytoplasmic proteins, the function of DsrS is unknown. It does not contain similarities to proteins of known functions or co-factor binding site (Dahl et al., 2005; Grimm et al., 2008). DsrS is probably not essential for the sulfur oxidation in *A. vinosum*, a deletion mutant shows a 30% reduction of sulfur oxidation rate (Grimm, 2012). The predicted

function of DsrN is a sirohaem amidase for the co-factor biosynthesis of the sulfite reductase. Mutants in *A. vinosum* showed a reduced oxidation of stored elemental sulfur (Lübbe et al., 2006).

DsrC is a small (12.8 kDa), soluble cytoplasmic protein and is essential for the sulfur oxidation in *A. vinosum* (Cort et al., 2008). The DsrC protein appears to be a dimer $[DsrC]_2$. In the *Codakia* symbiont genome seven copies of DsrC (ID 533/ 542/ 992/ 2441/ 2727/ 3612/ 3943/ 4122) genes are encoded, whereby only one copy is close to the *dsr* gene cluster (ID 533). These genes show low similarities among each other. Five of them were identified in the proteome (ID 533/ 542/ 992/ 2727/ 4122), but only ID 533/ 542 /4122 were abundantly identified in the proteome. DsrC is part of the core set of *dsr* genes present in all sulfur oxidizing bacteria (Frigaard and Dahl, 2009) and interacts with the DsrEFH protein complex (Cort et al., 2008). In *A. vinosum* the proteins DsrEFH form a hexameric complex arranged as [a2b2c2]. The genes for the DsrEFH complex are co-located in the *dsr* gene cluster in the *Codakia* genome and all these proteins were identified in the proteome. DsrEFH and DsrC homologs were found in different non-sulfur oxidizing bacteria like *E. coli* and many other Gammaproteobacteria. A homolog of the DsrC protein is TusE, homologs for DsrEFH are TusBCD in *E. coli.* These genes are part of the sulfur relay system during 2-thiouridine biosynthesis (Ikeuchi et al., 2006; Cort et al., 2008). DsrEFH serve as an acceptor for persulfidic sulfur, imported in the cytoplasm (Stockdreher et al., 2012). In *A. vinosum* the DsrC protein interacts with DsrAB and is part of the sulfur delivery system (Cort et al., 2008).

DsrA and DsrB together form multi-subunits $[(DsrA)_2][(DsrB)_2]$ in the cytoplasm and function as a reverse acting dissimilatory sulfite reductase containing a prosthetic group identified as siroamide-$[Fe_4S_4]$. Siroamide is an amidated form of the classic sirohaem (Lübbe et al., 2006; Grimm et al., 2008). The importance of this protein for the sulfur oxidation was shown in a *A. vinosum* deletion mutant, which is unable to oxidize stored sulfur (Weissgerber et al., 2011). The dissimilatory sulfite reductase works in close interaction with other Dsr proteins (Grimm et al., 2008). It is assumed that DsrEFH and DsrC are involved in the transfer of sulfide to the key enzymes DsrAB (Grimm et al., 2010a). The *dsrEFH* genes are located next to *dsrAB* in the *Codakia* symbiont genome, as well as in *A. vinosum* (figure 4.9).

The proteins DsrMKJOP form a complex at the inner membrane in *A. vinosum* (Sander et al., 2006). In the *Codakia* symbiont genome DsrJ is missing. DsrJ is part of the DsrMPOK complex and a predicted c-type cytochrome and essential in *A. vinosum* (Dahl et al., 2005). Only the proteins DsrP and DsrM are predicted to have 5 to 10 TMHs, whereas other proteins do not encode TMHs in the *Codakia* symbiont. This is similar in *A. vinosum*, exhibiting a strong attachment of DsrK predicted in the cytoplasm and DsrO is predicted to reside in the periplasmic space (Dahl et al., 2005). It is suggested that the complex is a membrane spanning electron transporting complex that is essential for the elemental sulfur oxidation in *A. vinosum* (Dahl et al., 2005; Sander et al., 2006). DsrP is predicted to operate as a quinol oxidase and DsrM could function as a quinone reductase (Dahl et al., 2005). New studies suggest that DsrR is part of the post-transcriptional control of the *dsr* operon in *A. vinosum* (Grimm et al., 2010b), DsrR has never been identified in the proteome.

The Dsr pathway seems to be important for the sulfur oxidation in the *Codakia* symbiont, as the related proteins are abundantly expressed.

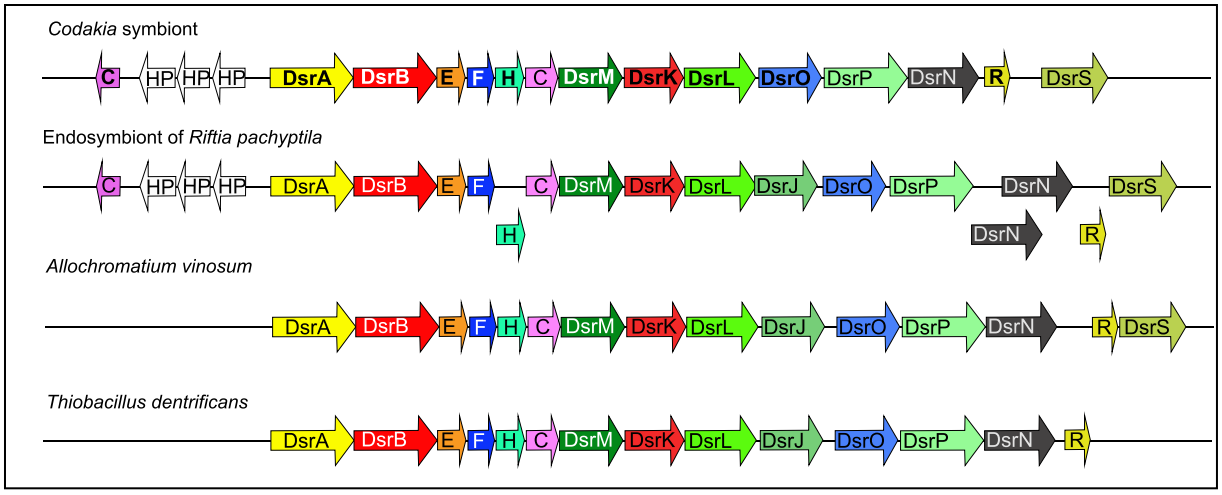

Figure 4.9. Schematic map of the of the *dsr* gene cluster in the symbiont genome of *Codakia* and *dsr* gene cluster in related organism. Predicted homolog proteins are indicated by the same color and label (bold marked labels indentified in the proteome) (figure adapted according to Friedrich et al., 2005; Markert et al., 2011).

4.3.3 Oxidative phosphorylation

The electrons from the sulfur oxidation are putatively transferred via the DsrMOPK and the QmoABC complexes to the respiratory chain and create a proton gradient between the periplasmic space and the cytoplasm. This proton gradient drives the ATP production via oxidative phosphorylation. In a previous study it was proofed that the symbiont uses oxygen as terminal electron acceptor (Duplessis et al., 2004a). The ATP synthase uses the proton gradient in the inter-membrane space and produces ATP. Six subunits of ATP synthase were identified in the proteome (table 3.24). The Pi, necessary for the ATP formation, provides the inorganic pyrophosphates PpA (ID 986), which were identified in the proteome. Several proteins involved in the membrane electron transport were identified in the proteome, like the ubiquino-cytochrome c reductase (ID 2092) and the cytochrome c oxidase FixP (ID 580), which were predicted to catalyze the last step of the electron transfer to oxygen (Markert et al., 2007).
Comparing the incubation with saturated and low oxygen conditions showed a low oxygen-dependent regulation of the ATP-synthase. This indicates that bivalve can buffer low oxygen conditions and can provide more oxygen for the symbiont through increased pumping. This is in accordance to personal observations that the bivalves shells were more open under low oxygen conditions. *Riftia* is able to buffer a temporary substrate deficiency, like oxygen and sulfide (Girguis, 2006).

4.3.4 Nitrogen metabolism

Nitrogen is an essential macroelement for life and is crucial for proteins and nucleic acids. Inorganic nitrogen has to be converted into a biologically useful form during gaseous nitrogen fixation and assimilatory nitrate reduction. Another feature of nitrate is that it can serve as an alternative electron acceptor within the respiratory pathway. The seagrass bed is described as a nitrate-poor environment (Duplessis et al., 2004a). The ammonium concentration is dependent on the depth of the sediment, and is between 8 µM at the surface and 39 µM in 10 cm sediment depth (table 4.1, figure 4.1).

4.3.4.1 Gaseous nitrogen fixation

The ability to fix gaseous nitrogen is restricted to prokaryotes and was found in most bacterial phylogenetic groups, also in the group of Gammaproteobacteria (Dixon et al., 2004). One example is *Beggiatoa*, a colorless sulfur Gammaproteobacterium, inhabited in freshwater and marine sediments, mostly in anaerobic sulfide rich zones. The ability of nitrogen fixation in *Beggiatoa* is diverse. It was shown that cultures of marine and freshwater strains fix gaseous nitrogen (Nelson et al., 1982; Polman et al., 1988; Desai et al., 2013). Nevertheless, this ability seems to be absent in *Beggiatoa* biofilms, which are abundant at hydrothermal vents (Nelson et al., 1989)(Preisler et al., 2007). *A. vinosum* also uses the nitrogenase for nitrogen fixation under microaerobic conditions (Weissgerber et al., 2011).
In the rhizophere of the eelgrass *Zostera marina*, nitrogen fixation was detected independently of the ammonium concentration in the pore water. The nitrogen fixation was typically higher under low oxygen or anaerob conditions. A symbiotic association between heterotrophic nitrogen fixing bacteria and eelgrass is suggested. Therefore, it is assumed that the diazotrophic bacteria are in direct contact with the roots of the eelgrass. A bidirectional exchange of material, ammonium to the roots and organic compounds to the heterotrophic diazotrophic bacteria, was described (McGlathery et al., 1998).
The obligate aerobe bacterium *Azotobacter vinelandii* fixes nitrogen via the nitrogenase, even under aerobic conditions (Setubal et al., 2009). For Cyanobacteria a nitrogen fixation under aerobic conditions has also been described (Bothe et al., 2010).
In the *Codakia* genome a gene cluster of at least 43 genes is encoded (figure. 4.10, table 3.28). However in *A. vinosum* the *nif* genes are spread over the genome (Weissgerber et al., 2011). The key genes for the nitrogen fixation are in all compared organisms co-localized.
The nitrogenase, catalyzing the biological reduction of dinitrogen to ammonia, contains two components NifK/D and NifH,. The "iron"-containing protein NifH is characterized by the presence of Fe^{+} ionen and functions as a ATP-dependent electron donor. NifK/D contains "molybdenum-iron" (MoFe) at the catalytic site and uses molybdenum as a co-factor (figure 4.11). The other genes in the *nif* gene cluster

are directly or indirectly supporting the nitrogenase, conducting different functions, like electron transfer, regulation or they function as co-factors for the biosynthesis. The nitrogenase is a kinetically slow enzyme (Thorneley et al., 1983) and consumes at least 16 mol ATP per mol fixed dinitrogen (figure 4.12).
Nitrogen fixation is an energy expensive pathway, therefore it is probably a growth limiting factor. Fixed nitrogen is probably provided for the host (Lechene et al., 2007). It has already been described that gaseous nitrogen fixation even occurs in the presence of ammonium (McGlathery et al., 1998; Desai et al., 2013).

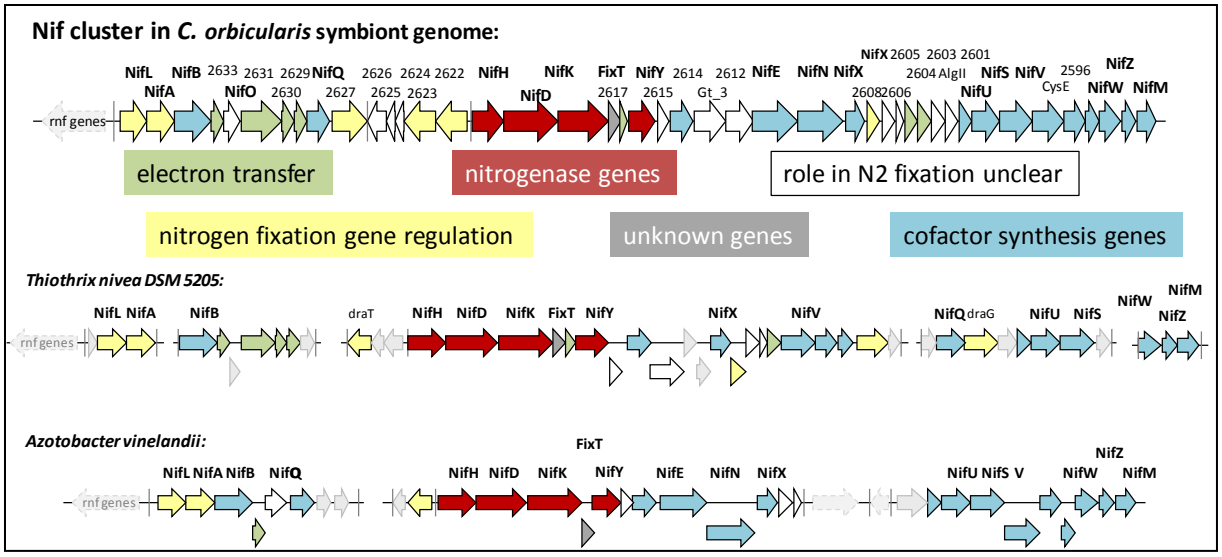

Figure 4.10. The *nif* gene cluster encoded in the *Codakia* genome and identified *nif* gene cluster in other organism. Predicted homologous proteins are indicated by the same color and label. The genes are divided in groups according to their function by the gaseous nitrogen fixation (derived from RAST SEED Viewer modified by Dr. Stephanie Markert and Sten König).

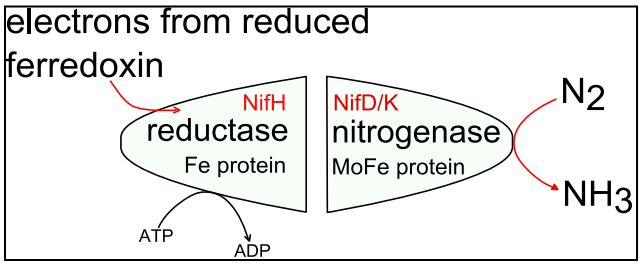

Figure 4.11.The nitrogenase is a two-protein complex.
The NifH component is an iron (Fe) containing protein that accepts electrons from ferredoxin. The NifD/K component contains an iron-molybdenum centrum and receives the electrons from NifH and fix gaseous nitrogen to ammonium.

$$N_2 + 8H^+ + 8e^- + 16ATP + 16H_2O \longrightarrow 2NH_3 + H_2 + 16ADP + 16H_3PO_4$$

Figure 4.12. Reaction equation of the nitrogen fixation with the nitrogenase.

The key proteins NifH/K/D were identified in a very high amounts in the proteome, indicating the importance of of nitrogen fixation for the symbionts of *C. orbiculata* and *C. orbicularis.* In summary, 26 proteins of the Nif cluster were identified, but sometimes only with a very low protein amounts. The proteomic data indicate that the nitrogen fixation is strictly down-regulated under sulfide starvation conditions. This can be explained by the high energy costs of the pathway.
In vitro, both components of the nitrogenase are sensitive to oxygen and irreversibly inactivated under aerobic conditions (Robson et al., 1980). Paradoxically, nitrogen

fixation also occurs under aerobic conditions, e. g. in the obligate aerobic prokaryote *Azotobacter vinelandii*, a heterotrophic Gammaproteobacterium fixing dinitrogen, (Poole and Hill, 1997) and in *Beggiatoa* spec (Nelson et al., 1982). The bacteria have to reconcile the oxygen sensitive nitrogenase and the strictly aerobic sulfur-oxidizing energy-generating metabolism. Therefore, various protection mechanisms were constructed to ensure that the nitrogenase is not damaged by oxygen. The Codakia symbiont live as endosymbionts in bacteriocytes inside the gill, where no visible barrier was observed, as compared to heterocysts of cyanobacteria with a thicker cell wall (Wolk et al., 2004). In the literature, several prospects are discussed for the nitrogenase protection, using *A. vinelandii* as a model. This organism protects the nitrogenase with a special protein called Shethna-protein FeSII, which supports a reversible conformational change of the nitrogenase. It contains two [Fe-2S] cluster for oxidation and reduction (Robson et al., 1980; Maier et al., 2000). However, in the *Codakia* symbiont no homolog protein of the Shethna-protein is encoded. In *A. vinlandii* further protection ways are described in the literature, which include an alternative way for auto-protection of the nitrogenase. The nitrogenase iron-protein (NifH) also reduces O_2 to H_2O_2, which can be further catalyzed by a catalase or peroxidase to water (Thorneley et al., 1989). However, in the proteome no indication of an increased catalase KatG (ID 399) expression was found (table 3.31). Only the superoxide dismutase SodA (ID 3897) was abundantly identified and up-regulated in the starvation sample. The expression is not directly coupled to the nitrogen fixation. Rubrerythrin, a putative peroxidase (Weinberg et al., 2004) and superoxidase (Lehmann et al., 1996) is a small protein (15.2 kDa) including a rubrerythrin ferritin-like di-iron binding domain. Rubrerythrin was found in the anaerobic sulfate reducing bacterium *Desulfovibrio vulgaris* (LeGall et al., 1988). In the *Codakia* genome three rubrerythrin genes are encoded, with low similarites to each other (ID 130/826/1096). However, only one gene is expressed (ID 1096), which is located next to a Fe-S oxidoreductase-like protein. In the symbiont of *C. orbicularis*, the rubrerythrin is the third most abundant protein, indicating a very important metabolic function for the symbiont. The highest amount of rubrerythrin was found in the control (freshly collected) sample. During starvation the amount of rubrerythrin decreased, the lowest level was found in the one of week starvation sample under low oxygen concentrations (table 3.29). That implies that the rubrerythrin shows the same expression like the nitrogenase under various conditions, indicating a coupled metabolic function. A homolog of rubrerythrin was found in the endosymbiont of *R. pachyptila*, but the symbiont does not encode nitrogen fixation genes. In air sensitive bacteria it is described that rubrerythrin plays a global role in the protection against oxygen in microaerophilic organisms (Kurtz, 2006). The main NCBI blast hits of the encoded RubY in *Codakia* symbiont are anaerobic bacteria (figure 3.33). In Cyanobacteria, Zhao et al (2007) suggests a close cooperation of the auto-protection of the nitrogenase, which produces H_2O_2, and the rubrerythrin, acting as peroxidase. Therefore, an involvment of rubrerythrin in the oxygen protection of the nitrogenase in the *Codakia* symbiont is likely.

Another possibility of nitrogenase protection against oxygen is called "respiratory protection" (Poole et al., 1997). This protects the nitrogenase by rapidly consuming

oxygen with an uncoupled respiration and thus reducing of the intracellular oxygen concentration. A special cytochrome bd type oxidase is predicted not to work as a proton pump. This uncoupling of the electron transfer allows consumption of high amounts of oxygen without being limited by ADP (Poole et al., 1997). In the *Codakia* symbiont a putative cytochrome bd-type oxidase is encoded with low similarity (ID 1282), which was not identified in the proteome. Another candidate could be the non-coupled NADH:ubiquinone oxidoreductase of *A. vinelandii* (Bertsova et al., 2001). A similar gene was found in the *Codakia* symbiont genome (ID 3792 $1e^{-141}$, 46% identity), which is however also not expressed, so probably not involved in the nitrogenase protection.

Caro et al. (2007) described different sub-populations of the symbiont within one gill, which differ in size, elemental sulfur content and quantity of genome copies. It is also possible that the sub-populations might differ in their physiology. One sub-population might create an anoxic sub-compartment inside the gill through high metabolic activity or an uncoupled respiratory chain and another sub-population might fix nitrogen. Contrary to this thesis is nevertheless the big amount of nitrogenase protein in the control sample, NifH is the ninth most abundant protein, and the absence of the putatively uncoupled cytochromes in the proteome.

In the *nif*-gene cluster, six different ferredoxin proteins are encoded. One of these encoded ferrodoxins shows similarity to ferrodoxin I FdxN (ID 2617, $6.0e^{-17}$ 65% identity) of *Rhodobacter capsulatus.* In *R. capsulatus*, it is described that ferredoxin I serves as the main electron donor to the nitrogenase (Jouanneau et al., 1995). Interestingly, the protein was also abundant identified in the *Codakia* symbiont and resembles the expression schema of the nitrogenase.

Neighboring genes of the *nif*-gene cluster in the *Codakia* symbiont genome are electron transport genes (*rnf*), indicating a close interaction. In the genome four gene clusters of *rnf* genes were found, containing 20 *rnf* genes (table 3.13). Nevertheless, only the *rnf* genes, which are encoded close to the *nif*-cluster, were identified in the proteome. The Rnf proteins build up a putative membrane bound complex to support the electron transport to the nitrogenase (Schmehl et al., 1993). An over-expression of *rnf* genes in *Rhodobacter capsulatus* showed an increased nitrogenase activity (Jeong and Jouanneau, 2000).

In the symbiont genome, two specific transporters for ammonium (ID 995/2000) and one nitrate ABC specific transporter (ID 3289) are encoded, but none is identified in the proteome. One of the ammonium transporters (ID 995) is co-localized with a nitrogen regulation protein (GlnB ID 996).

The identified Nif proteins in the *Codakia* symbiont were strictly down-regulated under starvation, probably because of the high energy costs of this pathway, making a stringent regulation necessary. In table 3.12. an overview about the predicted regulatory proteins of nitrogen fixation is shown. All the identified putative regulatory proteins are only low expressed. In *A. vinosum*, it is assumed that the regulation of the nitrogenase takes place at least on two regulatory levels (Weissgerber et al., 2011). Two nitrogen regulation proteins P-II (GlnB/K) (ID 996/ 4114/ 4148) are encoded in the *Codakia* symbiont. The proteins ID 996 and ID 4148 were identified in the proteome and both proteins show a quite high similarity (78% identity). The signal

transduction protein PII senses the intracellular nitrogen and carbon concentrations (Martinez-Argudo et al., 2004), probably using glutamine as an intracellular signal for nitrogen status (Dixon et al., 2004). The two-component system NtrB/C is co-localized, encoded in the *Codakia* genome, and identified in low amounts in the *Codakia* symbiont proteome (ID 1035/1036). It is proposed that GlnB controls the phosphorylation of NtrB and thereby also of NtrC. NtrC is a proposed key activator of the nitrogen regulation genes and is able to phosphorylate GlnK (Martinez-Argudo et al., 2004). This regulatory cascade displays the first level of nitrogenase regulation.
A second regulatory step of the nitrogenase in Proteobacteria is mostly dependent on the alternative sigma factor σ^{54}, (Martinez-Argudo et al., 2004). The *nifL and nifA* genes are co-localized on the *Codakia* genome (ID 2635/2336), but only less identified in the symbiont proteome. NifL/A is a predicted two component regulatory system for the nitrogen fixation in *Klebsiella pneumoniae* (Schmitz et al., 2002)(figure 4.13). The *NifA* protein of the *Codakia* symbiont contains several conserved domains like a sigma σ^{54} interaction domain, a DNA binding region and a Nif specific regulatory protein. This indicates an alternative sigma factor σ^{54}-dependent expression the NifA protein. The NifL protein in the *Codakia* symbiont also contains several domains, like a signal transduction histidine kinase-related domain, an ATP-binding domain and a PAS domain. The domains of NifA and NifL support the hypothesis of a sigma σ^{54}-dependent two-component system. In *Klebsiella* pneumoniae *the* NifA-NifL proteins can form a complex, which is able to measure intracellular oxygen concentrations and the presence of N-compounds (Schmitz et al., 2002). It is described that the NifA activity is directly reacting to the oxygen concentration in many diazotrophic Alpha- and Betaproteobacteria (Martinez-Argudo et al., 2004). However, for Gammaproteobacteria it was shown that the activity of NifA is totally independent of the oxygen and nitrogen concentrations (Martinez-Argudo et al., 2004). It seems that the nitrogen fixation in the *Codakia* symbiont is independent of the oxygen concentration but strictly depending on the feeding situation with sulfides. Further studies will be needed to identify how the strict regulation of the nitrogenase in the *Codakia* symbiont works.
A common method to measure the nitrogenase activity is the acetylene reduction assay. Thereby, the nitrogenase reduces acetylene gas to ethylene, both can be easily measured by gas chromatography (Dilworth, 1966). O. Gros performed first acetylene reduction experiments with gills of freshly collected *C. orbicularis* in sterile sea water. In figure 4.14, the first results are shown. A big peak of ethylene is visible, which is an evidence for an active nitrogenase in the gills of the bivalve *Codakia,* produced by the symbionts. Further experiments are necessary to improve these first results and to quantifiy the nitrogenase activity after one week starvation.

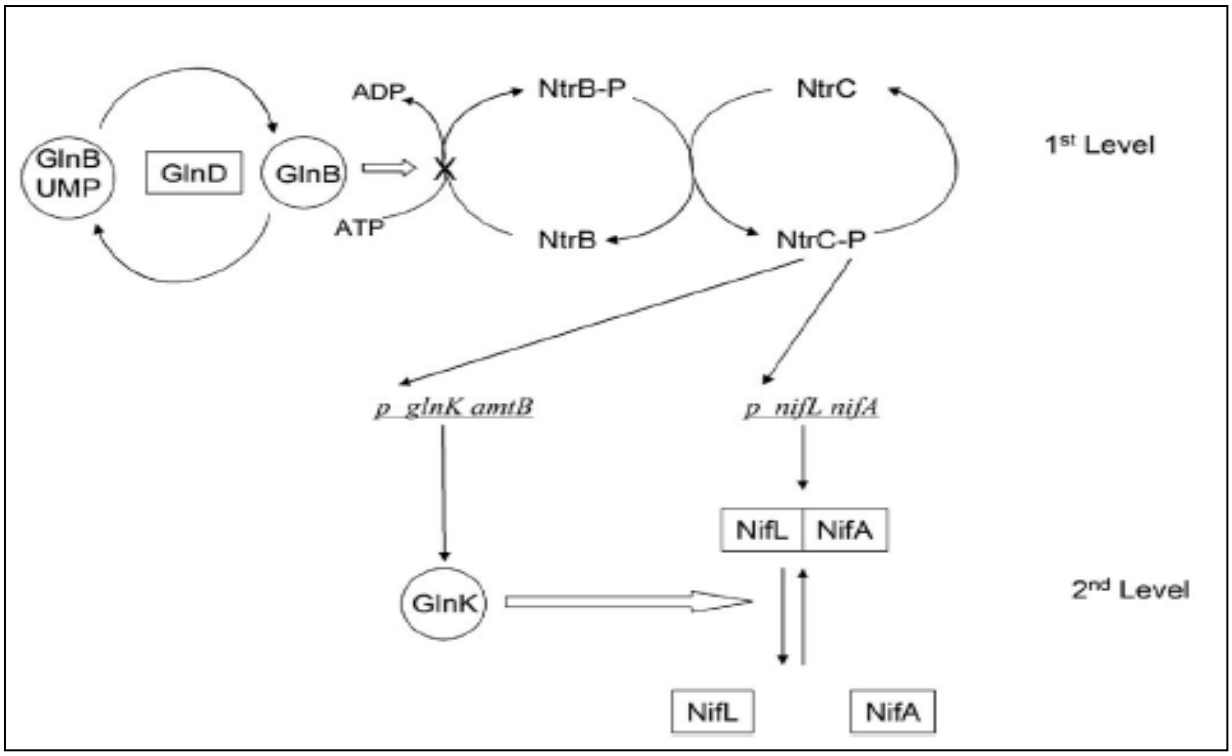

Figure 4.13. Regulation cascade of *nif* genes in *K. pneumoniae.*

The regulations proteins, GlnB, NtrB, NtrC and GlnK are encoded and expressed in the *Codakia* symbiont. (Martinez-Argudo et al., 2004).

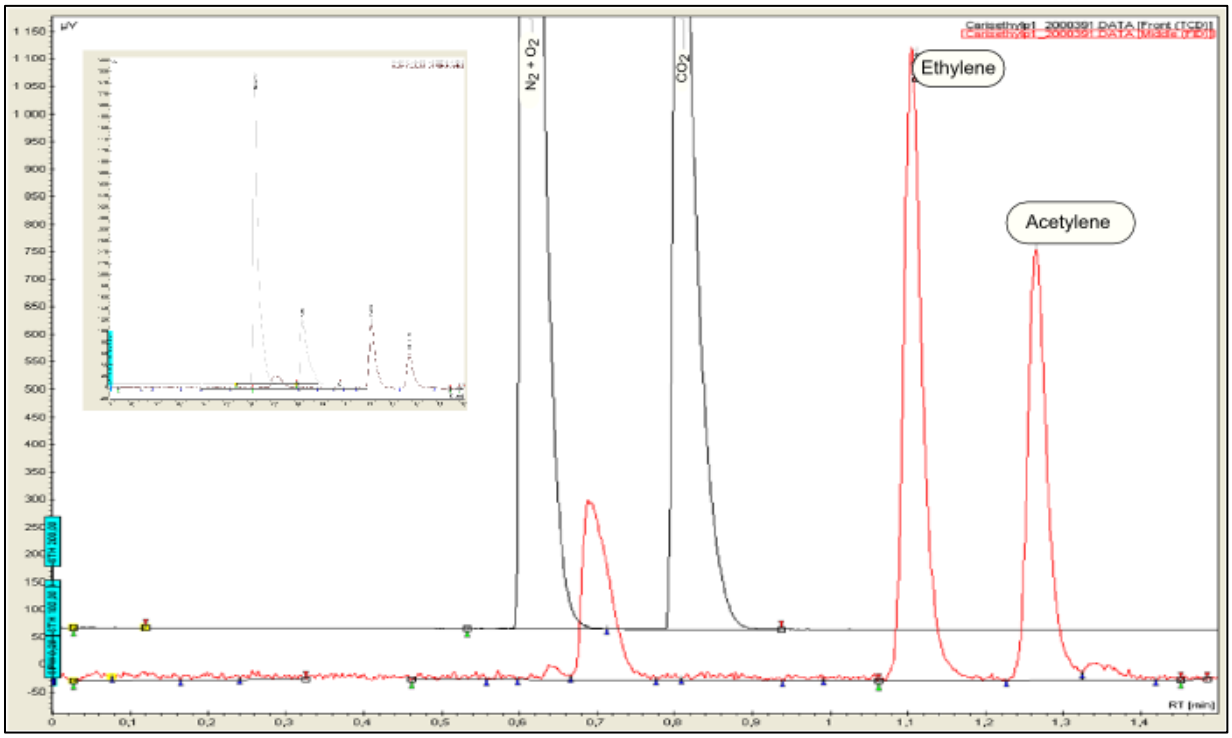

Figure 4.14. Acetylne reduction assay of *Codakia* gills.

Preliminary results of freshly gills from *C. orbicularis* performed by O. Gros measured by gas chromatography (y-axis: µV, x-axis: RT[min]). A big ethylene peak is visible an evidence for an active nitrogenase in the gills of *C. orbicularis.*

4.3.4.2 Nitrate reduction

The first evidence for nitrate reduction inside the gills was demonstrated by Berg et al. (1984), where nitrate reduction enzyme activity was detected in the *Codakia* gills. Later studies confirmed oxygen as primary electron acceptor, while the evidence of respiratory nitrate reduction was lacking (Duplessis et al., 2004a). In the *Codakia* symbiont genom two different types of nitrate reduction enzymes are encoded: (I) in the assimilatory nitrate reduction, inorganic nitrogen compounds are converted into biologically usable compounds like ammonium (table 3.11). (II) In the respiratory nitrate reduction, nitrate is used as a terminal electron acceptor (table 3.10). For the dissimilatory nitrate reduction, two systems *nar* and *nap* are encoded in the *Codakia* symbiont genome. A cluster of six Nap-genes, *napACBDGHL*, are encoded (ID 4298-4305). Additionally, two *narLX* genes, which encode a two-component system for nitrate, are co-localized with the *nap*-genes in the symbiont genome. In the literature, it is described that the Nap system is located in the periplasmic space, however, the predicted localization of most of the genes is in the cytoplasm (Moreno-Vivián et al., 1999). It is further predicted that the main function of the Nap system is not assimilatory or anaerobic respiration (Moreno-Vivián et al., 1999). The physiological role of the Nap system is not clear and may vary in different organisms. However, there is clear evidence that the Nap system is dissimilatory and involved in aerobic denitrification and/or balance the redox status (Moreno-Vivián et al., 1999). In the symbiont proteome, no indications were found for this system. The Nar system, the respiratory membrane-bound nitrate reductase, is encoded in the genome with seven *nar* genes, *narGHJLTX* (ID 3935 - 3942) (Moreno-Vivián et al., 1999). None of these proteins was identified in the proteome under the applied conditions.
Codakia lives at the interface between oxic and anoxic conditions. For the bivalve, oxygen is essential to survive longer time periods. The symbiont could handle anaerobic time periods by using respiratory nitrate reduction. Nitrate respiration is often inhibited by oxygen (Denis et al., 1990), because oxygen is the best terminal electron acceptor. In the seagrass bed the amount of nitrate or nitrite is very low (Duplessis et al., 2004b). Under low oxygen conditions, the bivalve and the symbiont are in competition for oxygen. However, in *Lucinoma aequizonata*, a marine bivalve with intracellular chemoautotrophic symbionts, the symbiont obligately uses nitrate as electron acceptor even in the presence of oxygen, which circumvents the competition for oxygen between host and symbiont. Nevertheless, in the habitat of *L. aequizonata* the oxygen concentration is low (>20 µM) and the nitrate concentration is high (30 µM) (Hentschel et al., 1995). Until now no evidence for nitrate respiration in *Codakia* has been found in the proteome, even under low oxygen conditions and an added nitrate source. In the sediment of the seagrass bed a free living form of the symbiont was found (Gros et al., 2003a). It is possible that the symbiont uses nitrate respiration in the free living form. More unlikely, these about 40 putative respiratory nitrate reduction genes are evolutionary deposited and the symbiont is only able to use oxygen in the respiratory chain. An alternative way for the symbiont to handle low oxygen conditions is the reduction of elementary sulfur to sulfide (Arndt et al., 2001; Duplessis et al., 2004a). Alternatively, glycogen usage, generating energy by

substrate chain phosphorylation, might enable symbiont and host to survive short anaerobic time periods.
The Nas system is located in the cytoplasm and responsible for the assimilatory nitrate reduction (Moreno-Vivián et al., 1999). The genes for assimilatory nitrate reduction are located in a gene cluster in the genome (ID 3284-3290) (table 3.11). The predicted assimilatory NADP(H)-dependent nitrite reductase are less identified and the assimilatory nitrate reductase were never identified under all applied conditions. However, the cluster contains two genes encoding for a nitrate transporter (*nasA*) (ID 3287/3290). Only one of them was abundantly identified (ID 3290) in the proteome. The seagrass bed is a nitrate poor environment, which is in accordance to the proteome data, where no or very weak indications were found for an assimilatory or dissimilatory nitrate respiration under the investigated conditions (Duplessis et al., 2004a).

4.4 Amino acid and lipid metabolism

Amino acids are the basic module of proteins. Prototroph organisms can synthesize all compounds and amino acids required for growth, which is in contrast to auxotroph organisms, which need supplemented compounds for their growth. As the *Codakia* symbiont can also be found as a free living from, the symbiont is a facultative symbiont (Gros et al., 2003a). Therefore, it is assumed that the symbiont is able to synthesize all amino acids. The ammonium is probably assimilated via the glutamine synthetase (GlnA, ID 1038), an enzyme that is identified in very high amounts, being one of the most abundant proteins in the symbiont proteome (figure 3.34, table 3.30). In the endosymbiont proteome of *R. pachyptila* the glutamine synthetase (GlnA) was identified in similar amounts (Markert et al., 2008).

4.5 Cytochemistry

Cytochemical methods were used to explain how lucinid bivalves survive when no organic particulate food or sulfides (needed for symbiont thiotrophic metabolism) are available. Although the digestive tract of Codakia bivalves is reduced (Reid, 1990), particle feeding is still part of the nutrient strategy of symbionts-bearing lucinids (Duplessis et al., 2004b). The symbionts occupy around 34% of the gills of freshly collected *C. orbicularis* (Caro et al., 2009). During starvation, the bacterial population decreases by one third per month (Caro et al., 2009), whereas no released bacteria are detectable using CARD-FISH (Brissac et al.,2009). The huge amount of symbionts serve as a nutrient storage for the bivalve. Lysosomes were rarely observed in freshly collected gills of *C. orbicularis* (Frenkiel et al., 1995) and are visible in figure 3.54-A/B in *C. orbiculata*. These data suggest that adult individuals can digest symbionts in their natural habitat, probably via resorption of dead bacteria and symbiont population controlling. In deep sea bivalves it was suggested that lysosomal symbiont digestion is a way of controlling symbiont population (Fiala-Medioni et al., 1990). During starvation the appearance of freshly collected lucinid gills changed from thick and lightly beige (containing large symbiont population with a big amount of elemental sulfur) to thin and dark brown in starved bivalve individuals (containing less symbionts without elemental sulfur) (personal observations) (Johnson et al., 2001; Lechaire et al., 2008; Caro et al., 2009). Activity of lysosomal enzymes was observed during starvation inside bacteriocytes close to and within the bacteria. We suggest that the host digests its bacterial endosymbionts during starvation to feed on the bacteria (i. e. farming). The lucinid *Lucina pectinata* digests symbionts in large secondary lysosomes (Liberge et al., 2001). The enzymatic degradation of the symbionts in *Codakia* appeared in another from compared to the degradation observed in *L. pectinata.* Lysosomes in *Codakia* bivalves may digest only individual symbionts, whereas in *L. pectinata* multiple bacteria are digested. *Codakia* might be less adapted to symbiont digestion than *L. pectinata.* The trophic pathway has also been described in other marine chemotrophic symbioses (Herry et al., 1989; Fiala et al., 1994; Boetius et al., 1995; Kádár et al., 2007).

Under nutrient-rich conditions, the hosts will acquire new symbiosis-competent bacteria from their environment, restoring a “normal” metabolic pathway (Gros et al., 2012).

It remains uncertain whether lucinids actively lyse bacteria to obtain nutrients during periods of host starvation, or whether the symbionts undergo autolysis, due to starvation, possibly resulting in nutritional benefits for the host. Further studies are necessary to investigate how bacterial populations are controlled by lucinids under various natural conditions, using metabolomics or stable isotope analysis.

4.6 Symbioses - interaction between host and symbiont

Several types of transport proteins were found in the genome and proteome of the symbiont. The symbiont is enclosed in vacuole membranes inside bacteriocytes. The host and the symbiont interact with each other. It is assumed that the symbionts supply the host with organic compounds. The exchange compounds are until now unknown. In *Teredinibacter turnerae*, also known as shipworm, the symbiont is also able to fix gaseous nitrogen and supplies nitrogen compounds like amino acids to the host (Lechene et al., 2007). Therefore, the symbiont needs efficient transport systems.

Protein secretion plays a central role in the interactions of bacteria with their environment, especially in case of symbiotic interactions (Tseng et al., 2009). Until now, six general classes of protein secretion systems are known in Gram-negative bacteria(Tseng et al., 2009). The general secretion pathway (Sec) and the Tat-dependent translocation are able to transport proteins in the periplasmic space. Both secretion pathways are encoded in the genome and were identified in the proteome (table 3.16).

For an effective host-symbiont interaction it is also necessary that the secreted proteins cross the plasma membrane of the vacuole, in which the symbiont is enclosed in the cytoplasm of the host. The type VI secretion (T6SS) machine is a recently identified multi-component secretion machine, which is able to translocate potential effector proteins into the cytoplasm of eukaryotic cells (Filloux et al., 2008) and thus it is probably involved in host-symbiont interaction (Records, 2011). A database screening showed that most of the bacteria with T6SS secretary machine are pathogenetic or symbiotic Proteobacteria (Filloux et al., 2008). The T6SS genes are encoded in the *Codakia* symbiont genome in a cluster of 15 genes, 11 of these genes are identified in the proteome, eventhough there are not abundantly found (table 3.15).

In a 2D gel of the soluble symbiont proteome one of the biggest spots is the "porin Gram-negative type" PorT (ID 910/4756) (table 3.18). In both symbionts of Codakia a PorT gene is encoded with 65% identity to each other, both proteins are identified in both organisms (figure 3.35). It is assumed that both PorT proteins are encoded in the symbiont genome. However, it should be noted that both symbiont genomes are still draft genomes and that only one copy was sequenced in each symbiont genome. The quantifications in the 1D gel and the gel free approach have to be handled with care, because both proteins are quite similar to each other and contain less unique peptides. Noticeably, the PorT protein contains only one TMH, which is unusual for a transmembrane protein, which normally contain at least two TMHs. Further investigations will be needed because of the unexpected high amount of PorT, to clarify the function and substrate specificity of the PorT protein.

The tripartite ATP-independent periplasmic transporter (TRAP) is a Na^+-dependent secondary transporter and is ubiquitous in prokaryotes especially in marine bacteria that live in sodium rich habitat (Mulligan et al., 2011). The TRAP type transporter contains a specific substrate-binding protein (SBP), which is responsible for the high affinity to its substrate, also known as extracytoplasmic solute receptor. The SBP is

mostly located in the periplasma or anchored in the cytoplasmic membrane (Mulligan et al., 2009). In *R. capsulatus* the TRAP transporter shows high affinity to C4-dicarboxylate, malate, succinate and fumarate (Forward et al., 1997). It is assumed that the energetic costs of transport are lower than that of an ABC transporter (Mulligan et al., 2009). The putative SBPs (DctP ID 891/ 1101/ 2444) were abundant identified in the *Codakia* symbiont proteome. In the *Codakia* symbiont genome 9 DctQ and DctM proteins are encoded but not identified in the proteome (table 3.14). DctQ and DctM encode the integral membrane-proteins containing several TMHs, which could be a reason for the difficulties in identification.

In the symbiont genome, 58 putative ABC transporter related genes are encoded (table 6.8.). Only ABC transporter related proteins were identified in the proteome, responsible for the uptake of inorganic phosphor (PstABC), molybdenum and ferric ions. The presence of transporters in the symbiont in the endosymbiont lifestyle is surprising. One hypothesis is that the host provides important growth nutrients to the symbiont, which would support the close interaction of both organisms.

Interestingly, a phosphotransferase system (PTS) is encoded and expressed, however not abundantly found (table 3.17). This bacterial system is used for sugar uptake, like glucose, mannose, fructose and the energy source is phosphoenolpyruvate translocation. The PTS does not seem to play a big role in the physiology of the symbiont under the investigated conditions, as it is only low expressed. Glycogen is a storage compound in the host and the bacteria and could be a potential interaction between host and symbiont (Sorgo et al., 2002). In the *Codakia* symbiont proteome indications for TRAP, ABC and PTS transporter were found in the endosymbiotic lifestage of the symbiont. This indicates an mixotrophically lifestyle of the symbiont even in the endosymbiont stage (Robidart et al., 2008). Additionally, It is possible that the exchange of organic compounds is bidirectional under certain conditions. This could mean that the host supports the bacteria even under freshly conditions with nutrients, like sugars or other growth factors, eventhough the main carbon transfer might be from the symbiont to the host.

In the *Codakia* symbiont genome several genes for a flagellum are encoded (table 6.7), but not identified. Inside the host, the symbiont is membrane-enclosed, where motility is not needed, so that a flagellum is supposed to be unnecessary. In the seagrass bed sediment a free living stock of bacteria was found (Gros et al., 2003a). The transmission of the symbiont occurs horizontally from an environmental stock (Gros et al., 1996). The presence of flagella genes supports the idea of a free living form of the symbiont, where flagella could be used to move inside the sediment.

The presence of two different RuBisCO forms is probably an adaption to an intracellular lifestyle of the symbionts, to reach higher CO_2 fixation rates. In the intracellular lifestyle, the RuBisCO protein form I is the most identified protein, but at the same time only a little amount of RuBisCO form II was identified. The differences between these two forms are described above.

4.7 Conclusion

The shallow water bivalves *Codakia orbicularis* and *C. orbiculata,* both belonging to the family *Lucinidae*, harbor endosymbiotic sulfur-oxidizing Gammaproteobacteria in their gills. The bivalves live in seagrass beds of *Thalassia testudinum* and harbor the same bacterial symbiont according to 16S rDNA analyses.
During starvation the symbiont population decreases while no release of symbionts were observed. We observed lysosomal enzyme activity during sulfide and nutrient starvation with cytochemical staining methods. We suggest that the host uses symbionts as a nutrient source to survive a hunger crisis. The carbon transfer from the symbiont to the host could be flexible and could consist in transfer of organic matter, "milking", under normal feeding conditions and digestion of the symbionts, "farming", under starved conditions.
Until now the symbiont alone is not cultivable. Therefore, cultivation-independent techniques, like -omics approaches were used to analyze the physiology of the symbiont (figure 4.15). Next generation sequencing (NGS) was employed to sequence the genomes of symbiont from both hosts, display the backbone for proteomics. The soluble- and membrane-associated symbiont proteomes were analyzed during different conditions.
The oxidation of sulfide is one key metabolic pathway of the *Codakia* symbiont, most probably using the periplasmic Sox system (sulfur oxidation), a cytoplasmatic sulfite reductase (DsrAB), an APS reductase (AprAB) and an ATP sulfurylase (SopT). Furthermore, indications for two additional putative sulfide oxidation systems in the periplasmic space, the sulfide quinone reductase (Sqr) and the sulfide dehydrogenase (FccAB), could be found. It was shown that the symbiont uses oxygen as terminal electron acceptor. The symbiont genome encodes several sets of dissimilatory nitrate reduction genes, however on the proteome level no evidence was found under the investigated conditions.
The CBB cycle of the symbiont is not completely encoded in the genome. The key genes, RuBisCO, are abundantly expressed. It is assumed that the regeneration of the ribulose-1,5-bisphosphate is performed unconventionally via a PPi-dependent phosphofructokinase. Another feature of the CBB is that two different forms of RuBisCO are encoded in the genome. Both are expressed at the same time, but RuBisCO form I is about 50 times more expressed. Additional to the autotrophic lifestyle all genes for the heterotrophic lifestyle are encoded in the genome. In the proteome, the enzymes related to glycolysis and TCA-cycle were low expressed, this pathway delivers important precursors for anabolic processes like for the amino acid synthesis. Interestingly, proteins for glycogen synthesis were identified in the proteome and glycogen particles were also detected using light microscope. Additionally, several types of transporters like ABC, TRAP and PTS are encoded in the genome. In the proteome several indications were found for an expression of these transporters, even in the endosymbiotic lifestyle.
Unexpectedly, in the genome a *nif* gene cluster is encoded for gaseous nitrogen fixation as ammonium source. The key genes, the nitrogenase NifH/K/D, were abundantly identified in proteome. The expression of these genes is strictly regulated

and only realized under high sulfide concentrations. The nitrogenase is irreversibly damaged by oxygen. Rubrerythrin, a strongly expressed protein, is predicted to protect the nitrogenase against oxygen stress.
The bacteria encode a specialized type 6 secretion system (T6SS) for the transport of bacterial effector molecules through the membranes to the host cytoplasm and display one possibility for a direct "communication" with the host.
In summary, genomic and proteomic analyses of the *Codakia* symbiont improved the knowledge of the major metabolic pathways of the symbiont in lucinid bivalves. The data of this study can be used as a basis for further in-depth analyses of the physiology and interaction of the symbiont with their hosts.

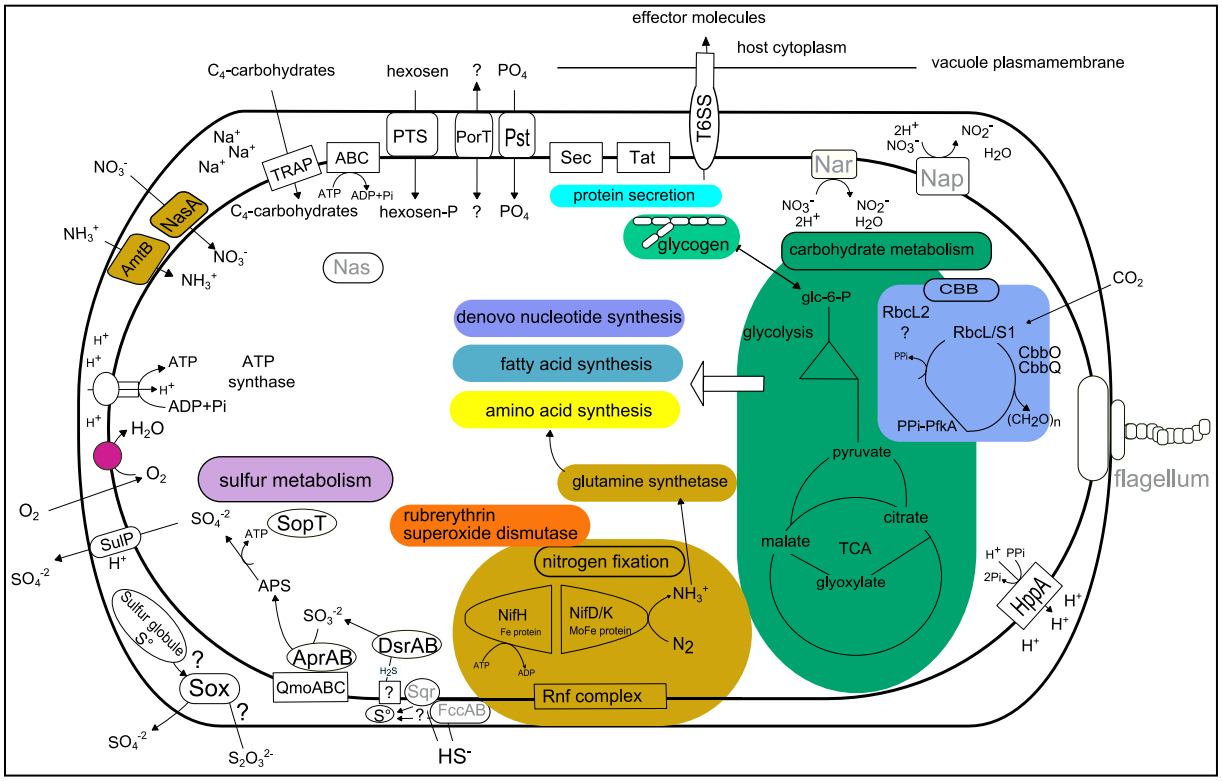

Figure 4.15. Overview about the reconstructed predicted physiology of the *Codakia* symbiont.
The dissimilatory nitrate respiration and the flagellum is encoded in the symbiont genome but in the proteome no indication was found under the investigated conditions. The other shown pathways for nitrogen fixation, CBB, glycolysis, TCA-cycle, sulfur metabolism were identified in the proteome. The first steps of sulfur oxidation in the periplasmic space were catalyzed by three systems (Sqr, FccAB and Sox-system), the Sqr and FccAB are only annotated as putative and in a low amount identified in the proteome (marked in grey). The Sox system were only identified after one week of sulfide starvation, the role of these systems in the sulfur oxidation needs further investigations. Genes for dissimilatory and assimilatory nitrate reduction are encoded in the genome but not identified unter the investigated conditions in the proteome (marked in grey). The genes for flagellum are encoded in the genome but not identified in the proteome of the endosymbiont (marked in grey). (Nap - periplasmic respiratory nitrate reductase, Nar - membrane bound respiratory nitrat reductase, Nas - assimilatory nitrate reduction, TCA - tricarboxylic acid cycle, S° - sulfur globules, Sox - Sulfur oxidation system, Dsr - dissimilatory sulfite reductase (DsrAB) and associated proteins, FccAB - putative sulfite dehydrogenase, Sqr - putative sulfide quinone reductase, HppA - proton-translocating pyrophopshatase, SopT/Sat - ATP sulfurylase/sulfate adenylyltransferase, PTS - Phosphotransferase system, TRAP - TRAP-transporter (Na^+ dep.), SulP - sulfur transporter, AmtB - ammonium transporter, NasA - nitrate transporter, T6SS - type VI secretion system)

4.8 Outlook

Symbioses of the bivalve *Codakia* are a good model system to investigate the interaction between host and bacterial symbiont. The advantages of the system are:

- the bivalve is easily accessible and can be collected from seagrass sediment
- the bivalve is easily cultivable in the laboratory for several months and it is possible to generate aposymbiotic juveniles
- re-colonization experiments are possible to study the first symbiotic interaction events
- sulfur granules in the bacteria allow easy separation of host and bacteria by density centrifugation

The bacterial symbiont is a facultative symbiont with a free-living lifestyle in the seagrass sediment. Up to date, cultivation of the bacteria alone is not possible and the identification of the necessary cultivation conditions will be a challenging task in the future (Stewart, 2012). Cultivation of only the symbiont would be a further step to being able to study symbiotic relationships with the *Codakia* model system in more detail.

Within one *Codakia* gill up to seven sub-populations of the symbiont were found, presumably all with different physiological characteristics (Caro et al., 2007). Distinguishing between these sub-populations and investigating the physiologies could elucidate the complex interplay between host and symbiont. Therefore, new extraction methods will be necessary, working independently of the density of the symbionts. A combination of different filtration steps with various pore sizes is thinkable to separate the bacteria from the host and later to differ between the sub-populations of the symbionts.

The genomic and proteomic data generated in this study can be used as a basis for further in-depth analyses of the physiology of the symbiont and its interaction with the host.

Further investigation, e. g. using microscopic techniques, will be needed to gain a more detailed knowledge of the glycogen metabolism and the metabolic interplay of host and symbiont during starvation and low oxygen conditions.

Until now, the exchange compounds between symbiont and host have not been identified. For their identification, turther metabolomics analyses are necessary. The known physiology data suggests exchange of an amino acid, due to the high amount of nitrogen fixation proteins and amino acid synthesizing proteins. To clarify the metabolic exchange compounds, the employment of a multi-isotope imaging mass spectrometry (MIMS) could be a useful approach (Lechene et al., 2007).

5. References

Amann, R.I., Ludwig, W., and Schleifer, K.-H. (1995). Phylogenetic identification and in situ detection of individual microbial cells without cultivation. Microbiol. Rev. *59*, 143–169.

An, S., and Gardner, W.S. (2002). Dissimilatory nitrate reduction to ammonium (DNRA) as a nitrogen link, versus denitrification as a sink in a shallow estuary (Laguna Madre/Baffin Bay, Texas). Mar. Ecol. Prog. Ser. *237*.

Arndt, C. (1998). Metabolic responses of the hydrothermal vent tube worm *Riftia pachyptila* to severe hypoxia. Mar. Ecol. Prog. Ser. *174*, 151–158.

Arndt, C., Gaill, F., and Felbeck, H. (2001). Anaerobic sulfur metabolism in thiotrophic symbioses. J. Exp. Biol. *204*, 741–750.

Badger, M.R., and Bek, E.J. (2007). Multiple RuBisCO forms in proteobacteria: their functional significance in relation to CO_2 acquisition by the CBB cycle. J. Exp. Bot. *59*, 1525–1541.

Barka, T., and Anderson, P.J. (1962). Histochemical methods for acid phosphatase using hexazonium pararosanalin as coupler. J. Histochem. Cytochem. *10*, 741–753.

Batstone, R.T., Laurich, J.R., Salvo, F., and Dufour, S.C. (2014). Divergent chemosymbiosis-related characters in *Thyasira cf. gouldi* (Bivalvia: Thyasiridae). PLoS ONE *9 (3): pp e92856*.

Baumann, P., and Moran, N.A. (1997). Non-cultivable microorganisms from symbiotic associations of insects and other hosts. Antonie Van Leeuwenhoek *72*, 39–48.

Baumann, P., Baumann, L., and Clark, M.A. (1996). Levels of *Buchnera aphidicola* chaperonin GroEL during growth of the aphid Schizaphis graminum. Curr. Microbiol. *32*, 279–285.

Bergemann, T.L., and Wilson, J. (2011). Proportion statistics to detect differentially expressed genes: a comparison with log-ratio statistics. BMC Bioinformatics *12*, 228.

Berg Jr., C.J., and Alatalo, P. (1984). Potential of chemosynthesis in molluscan mariculture. Aquaculture *39*, 165–179.

Bertics, V., Sohm, J., Treude, T., Chow, C., Capone, D., Fuhrman, J., and Ziebis, W. (2010). Burrowing deeper into benthic nitrogen cycling: the impact of bioturbation on nitrogen fixation coupled to sulfate reduction. Mar. Ecol. Prog. Ser. *409*, 1–15.

Bertsova, Y.V., Bogachev, A.V., and Skulachev, V.P. (2001). Noncoupled NADH:ubiquinone oxidoreductase of *Azotobacter vinelandii* is required for diazotrophic growth at high oxygen concentrations. J. Bacteriol. *183*, 6869–6874.

Betcher, M.A. (2011).Thesis: Insights into the shipworm (bivalvia: Teredinidae)-bacterial symbiosis: novel isolates and micrographic localization studies. Oregon Health & Science University.

Boetius, A., and Felbeck, H. (1995). Digestive enzymes in marine invertebrates from hydrothermal vents and other reducing environments. Mar. Biol. *122*, 105–113.

Boisvert, S., Laviolette, F., and Corbeil, J. (2010). Ray: simultaneous assembly of reads from a mix of high-throughput sequencing technologies. J. Comput. Biol. *17*, 1519–1533.

Bothe, H., Schmitz, O., Yates, M.G., and Newton, W.E. (2010). Nitrogen fixation and hydrogen metabolism in Cyanobacteria. Microbiol. Mol. Biol. Rev. *74*, 529–551.

Bradford, M.M. (1976). A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. Anal. Biochem. *72*, 248–254.

Brayant, C. (1991). Metazoan life without oxygen (London: Chapman & Hall).

Bright, M., and Bulgheresi, S. (2010). A complex journey: transmission of microbial symbionts. Nat. Rev. Microbiol. *8*, 218–230.

Brissac, T., Gros, O., and Mercot, H. (2009). Lack of endosymbiont release by two Lucinidae (Bivalvia) of the genus *Codakia*: consequences for symbiotic relationships. FEMS Microbiol. Ecol. *67*, 261–267.

Brissac, T., Merçot, H., and Gros, O. (2011). Lucinidae/sulfur-oxidizing bacteria: ancestral heritage or opportunistic association? Further insights from the Bohol Sea (the Philippines). FEMS Microbiol. Ecol. *75*, 63–76.

Brune, D.C. (1995). Isolation and characterization of sulfur globule proteins from *Chromatium vinosum* and *Thiocapsa roseopersicina*. Arch. Microbiol. *163*, 391–399.

Caro, A., Gros, O., Got, P., Wit, R.D., and Troussellier, M. (2007). Characterization of the population of the sulfur-oxidizing symbiont of *Codakia orbicularis* (Bivalvia, Lucinidae) by single-cell analyses. Appl. Environ. Microbiol. *73*, 2101–2109.

Caro, A., Got, P., Bouvy, M., Troussellier, M., and Gros, O. (2009). Effects of long-term starvation on a host bivalve *Codakia orbicularis* (Lucinidae) and its symbiont population. Appl. Environ. Microbiol. *75*, 3304–3313.

Cary, S.C. (1994). Vertical transmission of a chemoautotrophic symbiont in the protobranch bivalve, *Solemya reidi*. Mol. Mar. Biol. Biotechnol. *3*, 121–130.

Cavanaugh, C.M. (1983). Symbiotic chemoautotrophic bacteria in marine invertebrates from sulphide-rich habitats. Nature *302*, 58–61.

Cavanaugh, C.M., Gardiner, S.L., Jones, M.L., Jannasch, H.W., and Waterbury, J.B. (1981). Prokaryotic cells in the hydrothermal vent tube worm *Riftia pachyptila* Jones: possible chemoautotrophic symbionts. Science *213*, 340–342.

Cavanaugh, C.M., McKiness, Z.P., Newton, I.L.G., and Stewart, F.J. (2006). Marine Chemosynthetic Symbioses. In The Prokaryotes, M. Dworkin, S. Falkow, E. Rosenberg, K.-H. Schleifer, and E. Stackebrandt, eds. (New York, NY: Springer New York), pp. 475–507.

Childress, J.J., and Girguis, P.R. (2011). The metabolic demands of endosymbiotic chemoautotrophic metabolism on host physiological capacities. J. Exp. Biol. *214*, 312–325.

Childress, J.J., Fisher, C.R., Brooks, J.M., Kennicutt, M.C., Bidigare, R., and Anderson, A.E. (1986). A methanotrophic marine molluscan (bivalvia, mytilidae) symbiosis: mussels fueled by gas. Science *233*, 1306–1308.

Corliss, J.B., Dymond, J., Gordon, L.I., Edmond, J.M., von Herzen, R.P., Ballard, R.D., Green, K., Williams, D., Bainbridge, A., Crane, K., et al. (1979). Submarine thermal sprirngs on the galapagos rift. Science *203*, 1073–1083.

Cort, J.R., Selan, U., Schulte, A., Grimm, F., Kennedy, M.A., and Dahl, C. (2008). *Allochromatium vinosum* DsrC: solution-state NMR structure, redox properties, and interaction with DsrEFH, a protein essential for purple sulfur bacterial sulfur oxidation. J. Mol. Biol. *382*, 692–707.

Dahl, C., and Friedrich, C.G. (2008). Microbial Sulfur Metabolism (Springer Berlin Heidelberg).

Dahl, C., and Prange, A. (2006). Bacterial sulfur globules: occurrence, structure and metabolism. In Inclusions in Prokaryotes, D.J.M. Shively, ed. (Springer Berlin Heidelberg), pp. 21–51.

Dahl, C., Engels, S., Pott-Sperling, A.S., Schulte, A., Sander, J., Lubbe, Y., Deuster, O., and Brune, D.C. (2005). Novel genes of the dsr gene cluster and evidence for Close interaction of Dsr proteins during sulfur oxidation in the phototrophic sulfur bacterium *Allochromatium vinosum*. J. Bacteriol. *187*, 1392–1404.

Denis, K.S., Dias, F.M., and Rowe, J.J. (1990). Oxygen regulation of nitrate transport by diversion of electron flow in *Escherichia coli*. J. Biol. Chem. *265*, 18095–18097.

Desai, M.S., Assig, K., and Dattagupta, S. (2013). Nitrogen fixation in distinct microbial niches within a chemoautotrophy-driven cave ecosystem. ISME J. *7*, 2411–2423.

Dilworth, M.J. (1966). Acetylene reduction by nitrogen-fixing preparations from Clostridium pasteurianum. Biochim. Biophys. Acta BBA - Gen. Subj. *127*, 285–294.

Dimijian, G.G. (2000). Evolving together: the biology of symbiosis, part 1. Proc. Bayl. Univ. Med. Cent. *13*, 217–226.

Distel, D.L. (1998). Evolution of chemoautotrophic endosymbioses in bivalves. BioScience *48*, 277–286.

Distel, D.L., and Felbeck, H. (1988). Pathways of inorganic carbon fixation in the endosymbiont-bearing lucinid clam *Lucinoma aequizonata*. Part 1. Purification and characterization of the endosymbiotic bacteria. J. Exp. Zool. *247*, 1–10.

Distel, D.L., Lane, D.J., Olsen, G.J., Giovannoni, S.J., Pace, B., Pace, N.R., Stahl, D.A., and Felbeck, H. (1988). Sulfur-oxidizing bacterial endosymbionts: analysis of phylogeny and specificity by 16S rRNA sequences. J. Bacteriol. *170*, 2506–2510.

Distel, D.L., Lee, H.K., and Cavanaugh, C.M. (1995). Intracellular coexistence of methano-and thioautotrophic bacteria in a hydrothermal vent mussel. Proc. Natl. Acad. Sci. *92*, 9598–9602.

Dixon, R., and Kahn, D. (2004). Genetic regulation of biological nitrogen fixation. Nat. Rev. Microbiol. *2*, 621–631.

Dorman, D.C., Moulin, F.J.-M., McManus, B.E., Mahle, K.C., James, R.A., and Struve, M.F. (2002). Cytochrome oxidase inhibition induced by acute hydrogen sulfide inhalation: correlation with tissue sulfide concentrations in the rat brain, liver, lung, and nasal epithelium. Toxicol. Sci. *65*, 18–25.

Van Dover, C.L. (2000). The ecology of deep-sea hydrothermal vents (Princeton University Press, 2000).

Duarte, C.M., and Chiscano, C.L. (1999). Seagrass biomass and production: a reassessment. Aquat. Bot. *65*, 159–174.

Dubbs, J.M., and Robert Tabita, F. (2004). Regulators of nonsulfur purple phototrophic bacteria and the interactive control of CO2 assimilation, nitrogen fixation, hydrogen metabolism and energy generation. FEMS Microbiol. Rev. *28*, 353–376.

Dubilier, R.W. (1998). Ultrastructure and stable carbon isotope composition of the hydrothermal vent mussels *Bathymodiolus* brevior and B. sp. affinis brevior from the North Fiji Basin, western Pacific. Mar. Ecol.-Prog. Ser. *165*, 187–193.

Dubilier, N., Bergin, C., and Lott, C. (2008). Symbiotic diversity in marine animals: the art of harnessing chemosynthesis. Nat. Rev. Microbiol. *6*, 725–740.

Dufour, S.C. (2005). Gill anatomy and the evolution of symbiosis in the bivalve family Thyasiridae. Biol. Bull. *208*, 200–212.

Dufour, S.C., and Felbeck, H. (2003). Sulphide mining by the superextensile foot of symbiotic thyasirid bivalves. Nature *426*, 65–67.

Dufour, S., and Felbeck, H. (2006). Symbiont abundance in thyasirids (Bivalvia) is related to particulate food and sulphide availability. Mar. Ecol.-Prog. Ser. *320*, 185–194.

Duperron, S., Nadalig, T., Caprais, J.-C., Sibuet, M., Fiala-Medioni, A., Amann, R., and Dubilier, N. (2005). Dual symbiosis in a *Bathymodiolus sp.* mussel from a methane seep on the Gabon Continental Margin (Southeast Atlantic): 16S rRNA phylogeny and distribution of the symbionts in gills. Appl. Environ. Microbiol. *71*, 1694–1700.

Duperron, S., Gaudron, S.M., Rodrigues, C.F., Cunha, M.R., Decker, C., and Olu, K. (2013). An overview of chemosynthetic symbioses in bivalves from the North Atlantic and Mediterranean Sea. Biogeosciences *10*, 3241–3267.

Duplessis, M.R., Ziebis, W., Gros, O., Caro, A., Robidart, J., and Felbeck, H. (2004a). Respiration strategies utilized by the gill endosymbiont from the host Lucinid *Codakia orbicularis* (Bivalvia: Lucinidae). Appl. Environ. Microbiol. *70*, 4144–4150.

Duplessis, M.R., Dufour, S.C., Blankenship, L.E., Felbeck, H., and Yayanos, A.A. (2004b). Anatomical and experimental evidence for particulate feeding in *Lucinoma aequizonata* and *Parvilucina tenuisculpta* (Bivalvia: Lucinidae) from the Santa Barbara Basin. Mar. Biol. *145*, 551–561.

Durand, P., Gros, O., Frenkiel, L., and Prieur, D. (1996). Bacterial host specificity of Lucinacea endosymbionts: interspecific variation in 16S rRNA sequences. FEMS Microbiol. Lett. *140*, 193–198.

Durand P., P., Gros, O., Frenkiel, L., and Prieur, D. (1996). Phylogenetic characterization of sulfur-oxidizing bacterial endosymbionts in three tropical Lucinidae by 16S rDNA sequence analysis. Mol. Mar. Biol. Biotechnol. 5 37–42.

Eisen, J.A., Smith, S.W., and Cavanaugh, C.M. (1992). Phylogenetic relationships of chemoautotrophic bacterial symbionts of *Solemya velum* say (Mollusca: Bivalvia) determined by 16S rRNA gene sequence analysis. J. Bacteriol. *174*, 3416–3421.

Eisen, J.A., Nelson, K.E., Paulsen, I.T., Heidelberg, J.F., Wu, M., Dodson, R.J., Deboy, R., Gwinn, M.L., Nelson, W.C., Haft, D.H., et al. (2002). The complete genome sequence of *Chlorobium tepidum* TLS, a photosynthetic, anaerobic, green-sulfur bacterium. Proc. Natl. Acad. Sci. USA. *99*, 9509–9514.

Elisabeth, N. (2011), thesis: "Plasticite tissulaire et cellulaire du filament branchial des lucinidae symbiotique cotiers *Codakia orbiculata* et *Lucina pensylvanica"*. L'Université des Antilles et de la Guyane.

Elisabeth, N.H., Gustave, S.D.D., and Gros, O. (2012). Cell proliferation and apoptosis in gill filaments of the lucinid *Codakia orbiculata* (Montagu, 1808) (Mollusca: Bivalvia) during bacterial decolonization and recolonization. Microsc. Res. Tech. *75*, 1136–1146.

Elsaied, H., and Naganuma, T. (2001). Phylogenetic diversity of ribulose-1,5-bisphosphate carboxylase/oxygenase large-subunit genes from deep-sea microorganisms. Appl. Environ. Microbiol. *67*, 1751–1765.

Endow, K., and Ohta, S. (1990). Occurence of bacteria in the primary oocytes of vesicomyid clam *Calyptogena soyoae*. Mar. Ecol. Prog. Ser. *64*, 309–311.

Erftemeijer, P., and Middelburg, J. (1993). Sediment-nutrient interactions in tropical seagrass beds: a comparison between a terrigenous and a carbonate sedimentary environment in South Sulawesi (Indonesia). Mar. Ecol. Prog. Ser. *102*, 187–198.

Eymann, C., Dreisbach, A., Albrecht, D., Bernhardt, J., Becher, D., Gentner, S., Tam, L.T., Büttner, K., Buurman, G., Scharf, C., et al. (2004). A comprehensive proteome map of growing *Bacillus subtilis* cells. Proteomics *4*, 2849–2876.

Fares, M.A., Ruiz-González, M.X., Moya, A., Elena, S.F., and Barrio, E. (2002). Endosymbiotic bacteria: groEL buffers against deleterious mutations. Nature *417*, 398.

Fares, M.A., Moya, A., and Barrio, E. (2004). GroEL and the maintenance of bacterial endosymbiosis. Trends Genet. *20*, 413–416.

Felbeck, H. (1981). Chemoautotrophic potential of the hydrothermal vent tube worm, *Riftia pachyptila* Jones (Vestimentifera). Science *213*, 336–338.

Felbeck, H. (1983). Sulfide oxidation and carbon fixation by the gutless clam *Solemya reidi*: an animal-bacteria symbiosis. J. Comp. Physiol. *152*, 3–11.

Fiala, M.A., Michalski, J.C., Jolles, J., Alonso, C., and Montreuil, J. (1994). Lysosomic and lysozyme activities in the gill of bivalves from deep hydrothermal vents. C. R. Acad. Sci. III *317*, 239–244.

Fiala-Médioni, A., and Métivier, C. (1986). Ultrastructure of the gill of the hydrothermal vent bivalve *Calyptogena magnifica*, with a discussion of its nutrition. Mar. Biol. *90*, 215–222.

Fiala-Médioni, A., Métivier, C., Herry, A., and Le Pennec, M. (1986). Ultrastructure of the gill of the hydrothermal-vent mytilid *Bathymodiolus sp.* Mar. Biol. *92*, 65–72.

Fiala-Medioni, A., Felbeck, H., Childress, J.J., Fisher, C.R., and Vetter, R.D. (1990). Lysosomic resorption of bacterial symbionts in deep-sea bivalves. In Endocytobiology IV : 4th International Colloquium on Endocytobiology and Symbiosis, INSA-Villeurbanne (France), July 4-8, 1989, P. Nardon, V. Glaninazzi-Pearson, A.M. Grenier, L. Margulis, and D.C. Smith, eds. (Paris: Institut national de la recherche agronomique), pp. 335–338.

Filloux, A., Hachani, A., and Bleves, S. (2008). The bacterial type VI secretion machine: yet another player for protein transport across membranes. Microbiology *154*, 1570–1583.

Fisher, C.R., and Childress, J.J. (1986). Translocation of fixed carbon from symbiotic bacteria to host tissues in the gutless bivalve *Solemya reidi*. Mar. Biol. *93*, 59–68.

Forward, J.A., Behrendt, M.C., Wyborn, N.R., Cross, R., and Kelly, D.J. (1997). TRAP transporters: a new family of periplasmic solute transport systems encoded by the dctPQM genes of *Rhodobacter capsulatus* and by homologs in diverse gram-negative bacteria. J. Bacteriol. *179*, 5482–5493.

Frenkiel, L., and Mouëza, M. (1995). Gill ultrastructure and symbiotic bacteria in *Codakia orbicularis* (Bivalvia, Lucinidae). Zoomorphology *115*, 51–61.

Frenkiel, L., Gros, O., and Mouëza, M. (1996). Gill structure in *Lucina pectinata* (Bivalvia: Lucinidae) with reference to hemoglobin in bivalves with symbiotic sulphur-oxidizing bacteria. Mar. Biol. *125*, 511–524.

Friedrich, C.G., Rother, D., Bardischewsky, F., Quentmeier, A., and Fischer, J. (2001). Oxidation of Reduced Inorganic Sulfur Compounds by Bacteria: Emergence of a Common Mechanism? Appl. Environ. Microbiol. *67*, 2873–2882.

Friedrich, C.G., Bardischewsky, F., Rother, D., Quentmeier, A., and Fischer, J. (2005). Prokaryotic sulfur oxidation. Curr. Opin. Microbiol. *8*, 253–259.

Friedrich, D.C.G., Quentmeier, A., Bardischewsky, F., Rother, D., Orawski, G., Hellwig, P., and Fischer, J. (2008). Redox control of chemotrophic sulfur oxidation of

Paracoccus pantotrophus. In Microbial Sulfur Metabolism, D.C. Dahl, and D.C.G. Friedrich, eds. (Springer Berlin Heidelberg), pp. 139–150.

Frigaard, N.-U., and Dahl, C. (2009). Sulfur metabolism in phototrophic sulfur bacteria. Adv. Microb. Physiol. *54*, 103–200.

Gardebrecht, A., Markert, S., Sievert, S.M., Felbeck, H., Thürmer, A., Albrecht, D., Wollherr, A., Kabisch, J., Le Bris, N., Lehmann, R., et al. (2011). Physiological homogeneity among the endosymbionts of *Riftia pachyptila* and *Tevnia jerichonana* revealed by proteogenomics. ISME J. *6*, 766–776.

Girguis, P.R. (2006). Metabolite uptake, stoichiometry and chemoautotrophic function of the hydrothermal vent tubeworm *Riftia pachyptila*: responses to environmental variations in substrate concentrations and temperature. J. Exp. Biol. *209*, 3516–3528.

Glauert, A.M. (1991). Epoxy resins: an update on their selection and use in: European Microscopy and Analysis. Eur. Microsc. Anal. 15–20.

Goffredi, S., Hurtado, L., Hallam, S., and Vrijenhoek, R. (2003). Evolutionary relationships of deep-sea vent and cold seep clams (Mollusca: Vesicomyidae) of the“ pacifica/lepta” species complex. Mar. Biol. *142*, 311–320.

Goris, J., Konstantinidis, K.T., Klappenbach, J.A., Coenye, T., Vandamme, P., and Tiedje, J.M. (2007). DNA-DNA hybridization values and their relationship to whole-genome sequence similarities. Int. J. Syst. Evol. Microbiol. *57*, 81–91.

Gourdine, J., and Smithravin, E. (2007). Analysis of a cDNA-derived sequence of a novel mannose-binding lectin, codakine, from the tropical clam *Codakia orbicularis*. Fish Shellfish Immunol. *22*, 498–509.

Griesbeck, C., Hauska, G., and Schütz, M. (2000). Biological sulfide oxidation: sulfide-quinone reductase (SQR), the primary reaction. Recent Res. Dev. Microbiol. *4*, 179–203.

Grimm, F. (2012). Regulation of the dsr operon and function of the proteins DsrR and DsrS in the purple sulfur bacterium *Allochromatium vinosum*. Universitäts-und Landesbibliothek Bonn.

Grimm, F., Franz, B., and Dahl, C. (2008). Thiosulfate and sulfur oxidation in purple sulfur bacteria. In Microbial Sulfur Metabolism, (Springer), pp. 101–116.

Grimm, F., Dobler, N., and Dahl, C. (2010a). Regulation of dsr genes encoding proteins responsible for the oxidation of stored sulfur in *Allochromatium vinosum*. Microbiology *156*, 764–773.

Grimm, F., Cort, J.R., and Dahl, C. (2010b). DsrR, a novel IscA-like protein lacking iron- and Fe-S-binding functions, involved in the regulation of sulfur oxidation in *Allochromatium vinosum*. J. Bacteriol. *192*, 1652–1661.

Gros (1999). Embryonic development and endosymbiont transmission mode in the symbiotic clam *Lucinoma aequizonata* (Bivalvia:Lucinidae). Invertebr. Reprod. Amp Dev. *36*, 93–103.

Gros, O., Darrasse, A., Durand, P., Frenkiel, L., and Moueza, M. (1996). Environmental transmission of a sulfur-oxidizing bacterial gill endosymbiont in the tropical lucinid bivalve *Codakia orbicularis*. Appl. Environ. Microbiol. *62*, 2324–2330.

Gros, O., Wulf-Durand, P., Frenkiel, L., and Mouëza, M. (1998). Putative environmental transmission of sulfur-oxidizing bacterial symbionts in tropical lucinid bivalves inhabiting various environments. FEMS Microbiol. Lett. *160*, 257–262.

Gros, O., Liberge, M., Heddi, A., Khatchadourian, C., and Felbeck, H. (2003a). Detection of the free-living forms of sulfide-oxidizing gill endosymbionts in the lucinid habitat (*Thalassia testudinum* environment). Appl. Environ. Microbiol. *69*, 6264–6267.

Gros, O., Liberge, M., and Felbeck, H. (2003b). Interspecific infection of aposymbiotic juveniles of *Codakia orbicularis* by various tropical lucinid gill-endosymbionts. Mar. Biol. *142*, 57–66.

Gros, O., Elisabeth, N.H., Gustave, S.D.D., Caro, A., and Dubilier, N. (2012). Plasticity of symbiont acquisition throughout the life cycle of the shallow-water tropical lucinid *Codakia orbiculata* (Mollusca: Bivalvia): Symbiont acquisition in lucinid clams. Environ. Microbiol. *14*, 1584–1595.

Gruber-Vodicka, H.R., Dirks, U., Leisch, N., Baranyi, C., Stoecker, K., Bulgheresi, S., Heindl, N.R., Horn, M., Lott, C., Loy, A., et al. (2011). *Paracatenula*, an ancient symbiosis between thiotrophic Alphaproteobacteria and catenulid flatworms. Proc. Natl. Acad. Sci. *108*, 12078–12083.

Van Ham, R.C., Kamerbeek, J., Palacios, C., Rausell, C., Abascal, F., Bastolla, U., Fernández, J.M., Jiménez, L., Postigo, M., Silva, F.J., et al. (2003). Reductive genome evolution in *Buchnera aphidicola*. Proc. Natl. Acad. Sci. *100*, 581–586.

Hayashi, N.R., Arai, H., Kodama, T., and Igarashi, Y. (1997). The novel genes, *cbbQ* and *cbbO* located downstream from the RuBisCO genes of *Pseudomonas hydrogenothermophila,* affect the conformational states and activity of RuBisCO. Biochem. Biophys. Res. Commun. *241*, 565–569.

Hayashi, N.R., Oguni, A., Yaguchi, T., Chung, S.-Y., Nishihara, H., Kodama, T., and Igarashi, Y. (1998). Different properties of gene products of three sets ribulose 1,5-bisphosphate carboxylase/oxygenase from a marine obligately autotrophic hydrogen-oxidizing bacterium, *Hydrogenovibrio marinus* strain MH-110. J. Ferment. Bioeng. *85*, 150–155.

Van der Heide, T., Govers, L.L., de Fouw, J., Olff, H., van der Geest, M., van Katwijk, M.M., Piersma, T., van de Koppel, J., Silliman, B.R., Smolders, A.J.P., et al. (2012). A three-stage symbiosis forms the foundation of seagrass ecosystems. Science *336*, 1432–1434.

Heinhorst, S., Baker, S.H., Johnson, D.R., Davies, P.S., Cannon, G.C., and Shively, J.M. (2002). Two copies of form I RuBisCO genes in *Acidithiobacillus ferrooxidans* ATCC 23270. Curr. Microbiol. *45*, 115–117.

Hensen, D., Sperling, D., Trüper, H.G., Brune, D.C., and Dahl, C. (2006). Thiosulphate oxidation in the phototrophic sulphur bacterium *Allochromatium vinosum*. Mol. Microbiol. *62*, 794–810.

Hentschel, U., and Felbeck, H. (1995). Nitrate respiration in chemoautotrophic symbionts of the bivalve *Lucinoma aequizonata* is not regulated by oxygen. Appl. Environ. Microbiol. *61*, 1630–1633.

Herry, A., Diouris, M., and Le Pennec, M. (1989). Chemoautotrophic symbionts and translocation of fixed carbon from bacteria to host tissues in the littoral bivalve *Loripes lucinalis* (Lucinidae). Mar. Biol. *101*, 305–312.

Ikeuchi, Y., Shigi, N., Kato, J., Nishimura, A., and Suzuki, T. (2006). Mechanistic insights into sulfur relay by multiple sulfur mediators involved in thiouridine biosynthesis at tRNA wobble positions. Mol. Cell *21*, 97–108.

Imhoff, J.F., Sahling, H., Sling, J., and Kath, T. (2003). 16S rDNA-based phylogeny of sulphur-oxidising bacterial endosymbionts in marine bivalves from cold-seep habitats. Mar. Ecol. Prog. Ser. *249*, 39–51.

Jackson, J.B.C. (1972). The ecology of the molluscs of *Thalassia* communities, Jamaica, West Indies. II. Molluscan population variability along an environmental stress gradient. Mar. Biol. *14*, 304–337.

Janda, J.M., and Abbott, S.L. (2007). 16S rRNA gene sequencing for bacterial identification in the diagnostic laboratory: pluses, perils, and pitfalls. J. Clin. Microbiol. *45*, 2761–2764.

Jeong, H.-S., and Jouanneau, Y. (2000). Enhanced nitrogenase activity in strains of *Rhodobacter capsulatus* that overexpress the rnf genes. J. Bacteriol. *182*, 1208–1214.

Johnson, M.A., and Fernandez, C. (2001). Bacterial symbiosis in *Loripes lucinalis* (Mollusca: Bivalvia) with comments on reproductive strategy. J. Mar. Biol. Assoc. UK *81*, 251–257.

Johnson, K.S., Childress, J.J., Hessler, R.R., Sakamoto-Arnold, C.M., and Beehler, C.L. (1988). Chemical and biological interactions in the Rose Garden hydrothermal vent field, Galapagos spreading center. Deep Sea Res. Part Oceanogr. Res. Pap. *35*, 1723–1744.

Jørgensen, B.B. (1982). Mineralization of organic matter in the sea bed - the role of sulphate reduction. Nature *296*, 643–645.

Jørgensen, B.B., and Fenchel, T. (1974). The sulfur cycle of a marine sediment model system. Mar. Biol. *24*, 189–201.

Jouanneau, Y., Meyer, C., Naud, I., and Klipp, W. (1995). Characterization of an fdxN mutant of *Rhodobacter capsulatus* indicates that ferredoxin I serves as electron donor to nitrogenase. Biochim. Biophys. Acta *1232*, 33–42.

Kabisch, J., Pratzka, I., Meyer, H., Albrecht, D., Lalk, M., Ehrenreich, A., and Schweder, T. (2013). Metabolic engineering of *Bacillus subtilis* for growth on overflow metabolites. Micob Cell Fact *12*, 72.

Kádár, E., Davis, S.A., and Lobo-da-Cunha, A. (2007). Cytoenzymatic investigation of intracellular digestion in the symbiont-bearing hydrothermal bivalve *Bathymodiolus azoricus*. Mar. Biol. *153*, 995–1004.

Kappler, U., and Dahl, C. (2001). Enzymology and molecular biology of prokaryotic sulfite oxidation1. FEMS Microbiol. Lett. *203*, 1–9.

Kelly, D.P., Shergill, J.K., Lu, W.P., and Wood, A.P. (1997). Oxidative metabolism of inorganic sulfur compounds by bacteria. Antonie Van Leeuwenhoek *71*, 95–107.

Kinross, J.M., Darzi, A.W., and Nicholson, J.K. (2011). Gut microbiome-host interactions in health and disease. Genome Med. *3*, 14.

Kleiner, M., Wentrup, C., Lott, C., Teeling, H., Wetzel, S., Young, J., Chang, Y.-J., Shah, M., VerBerkmoes, N.C., and Zarzycki, J. (2012). Metaproteomics of a gutless marine worm and its symbiotic microbial community reveal unusual pathways for carbon and energy use. Proc. Natl. Acad. Sci. *109*, E1173–E1182.

Klose (1975). Protein mapping by combined isoelectric focusing and electrophoresis of mouse tissue. A novel approach to testing for induced point mutations in mammals. *26*, 231–243.

Kraus, D., and Doeller, J. (1995). Heme proteins in sulfide-oxidizing bacteria/mollusc symbioses. Am. Zool. *35*, 112–120.

Krueger, D.M., Dubilier, N., and Cavanaugh, C.M. (1996). Chemoautotrophic symbiosis in the tropical clam *Solemya* occidentalis (Bivalvia: Protobranchia): ultrastructural and phylogenetic analysis. Mar. Biol. *126*, 55–64.

Krumholz, L.R., Wang, L., Beck, D.A.C., Wang, T., Hackett, M., Mooney, B., Juba, T.R., McInerney, M.J., Meyer, B., Wall, J.D., et al. (2013). Membrane protein complex of APS reductase and Qmo is present in *Desulfovibrio vulgaris* and *Desulfovibrio alaskensis*. Microbiology *159*, 2162–2168.

Krylova, E.M., and Sahling, H. (2006). Recent bivalve molluscs of the genus *Calyptogena* (Vesicomyidae). J. Molluscan Stud. *72*, 359–395.

Krylova, E.M., and Sahling, H. (2010). Vesicomyidae (Bivalvia): current taxonomy and distribution. PLoS ONE *5*, e9957.

Kurtz, D.M. (2006). Avoiding high-valent iron intermediates: Superoxide reductase and rubrerythrin. J. Inorg. Biochem. *100*, 679–693.

Kurtz, S., Phillippy, A., Delcher, A.L., Smoot, M., Shumway, M., Antonescu, C., and Salzberg, S.L. (2004). Versatile and open software for comparing large genomes. Genome Biol. *5*, R12.

Laemmli (1970). Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature *227*, 680–685.

Lechaire, J.-P., Frébourg, G., Gaill, F., and Gros, O. (2008). In situ characterization of sulphur in gill-endosymbionts of the shallow water lucinid *Codakia orbicularis* (Linné,

1758) by high-pressure cryofixation and EFTEM microanalysis. Mar. Biol. *154*, 693–700.

Lechene, C.P., Luyten, Y., McMahon, G., and Distel, D.L. (2007). Quantitative imaging of nitrogen fixation by individual bacteria within animal cells. Science *317*, 1563–1566.

Lee, R.W., and Childress, J.J. (1994). Assimilation of inorganic nitrogen by marine invertebrates and their chemoautotrophic and methanotrophic symbionts. Appl. Environ. Microbiol. *60*, 1852–1858.

Lee, R.W., Thuesen, E.V., and Childress, J.J. (1992). Ammonium and free amino acids as nitrogen sources for the chemoautotrophic symbiosis *Solemya reidi* Bernard (Bivalvia: Protobranchia). J. Exp. Mar. Biol. Ecol. *158*, 75–91.

LeGall, J., Prickril, B.C., Moura, I., Xavier, A.V., Moura, J.J., and Hanh, H.B. (1988). Isolation and characterization of rubrerythrin, a non-heme iron protein from *Desulfovibrio vulgaris* that contains rubredoxin centers and a hemerythrin-like binuclear iron cluster. Biochemistry (Mosc.) *27*, 1636–1642.

Lehmann, Y., Meile, L., and Teuber, M. (1996). Rubrerythrin from Clostridium perfringens: cloning of the gene, purification of the protein, and characterization of its superoxide dismutase function. J. Bacteriol. *178*, 7152–7158.

Lewis, P., and Knight, D.P. (1992). Cytochemical staining methods for electron microscopy (Elsevier).

Liberge, M., Gros, O., and Frenkiel, L. (2001). Lysosomes and sulfide-oxidizing bodies in the bacteriocytes of *Lucina pectinata*, a cytochemical and microanalysis approach. Mar. Biol. *139*, 401–409.

Lu, W.-P., Swoboda, B.E.P., and Kelly, D.P. (1985). Properties of the thiosulphate-oxidizing multi-enzyme system from *Thiobacillus versutus*. Biochim. Biophys. Acta BBA - Protein Struct. Mol. Enzymol. *828*, 116–122.

Lübbe, Y.J., Youn, H.-S., Timkovich, R., and Dahl, C. (2006). Siro(haem)amide in *Allochromatium vinosum* and relevance of DsrL and DsrN, a homolog of cobyrinic acid *a,c* -diamide synthase, for sulphur oxidation. FEMS Microbiol. Lett. *261*, 194–202.

Lutz, R.A., Shank, T.M., Fornari, D.J., Haymon, R.M., Lilley, M.D., Von Damm, K.L., and Desbruyeres, D. (1994). Rapid growth at deep-sea vents. Nature *371*, 663–664.

Maier, R.J., and Moshiri, F. (2000). Role of the *Azotobacter vinelandii* nitrogenase-protective shethna protein in preventing oxygen-mediated cell death. J. Bacteriol. *182*, 3854–3857.

Margulis, L., and Fester, R. (1991). Symbiosis as a source of evolutionary innovation: speciation and morphogenesis (Cambridge (Massachusetts): MIT Press).

Markert, S. (2008). Functional genome analysis of the uncultured bacterial endosymbiont from the deep sea tubeworm *Riftia pachyptila*.

Markert, S., Arndt, C., Felbeck, H., Becher, D., Sievert, S.M., Hugler, M., Albrecht, D., Robidart, J., Bench, S., Feldman, R.A., et al. (2007). Physiological proteomics of the uncultured endosymbiont of *Riftia pachyptila*. Science *315*, 247–250.

Markert, S., Gardebrecht, A., Felbeck, H., Sievert, S.M., Klose, J., Becher, D., Albrecht, D., Thürmer, A., Daniel, R., Kleiner, M., et al. (2011). Status quo in physiological proteomics of the uncultured *Riftia pachyptila* endosymbiont. Proteomics *11*, 3106–3117.

Markowitz, V.M., Chen, I.-M.A., Palaniappan, K., Chu, K., Szeto, E., Pillay, M., Ratner, A., Huang, J., Woyke, T., Huntemann, M., et al. (2014). IMG 4 version of the integrated microbial genomes comparative analysis system. Nucleic Acids Res. *42*, D560–D567.

Martinez-Argudo, I., Little, R., Shearer, N., Johnson, P., and Dixon, R. (2004). The NifL-NifA system: a multidomain transcriptional regulatory complex that integrates environmental signals. J. Bacteriol. *186*, 601–610.

Maurin, L. (2009). Ecologie des nématodes marins libres et symbiotiques en milieu tropical. Développement de la microspectrométrie Raman comme outil de caractérisation des organismes thiotrophiques.

McFall-Ngai, M. (2008). Are biologists in “future shock”? Symbiosis integrates biology across domains. Nat. Rev. Microbiol. *6*, 789–792.

McFall-Ngai, M.J. (2001). Identifying “prime suspects”: symbioses and the evolution of multicellularity. Comp. Biochem. Physiol. B. Biochem. Mol. Biol. *129*, 711–723.

McFall-Ngai, M.J. (2002). Unseen forces: The influence of bacteria on animal development. Dev. Biol. *242*, 1–14.

McGlathery, K.J., RisgaardPetersen, N., and Christensen, P.B. (1998). Temporal and spatial variation in nitrogen fixation activity in the eelgrass *Zostera marina* rhizosphere. Mar. Ecol. Prog. Ser. *168*, 245–258.

Meyer, B., and Kuever, J. (2007). Molecular analysis of the distribution and phylogeny of dissimilatory adenosine-5’-phosphosulfate reductase-encoding genes (aprBA) among sulfur-oxidizing prokaryotes. Microbiology *153*, 3478–3498.

Million, M., Lagier, J.-C., Yahav, D., and Paul, M. (2013). Gut bacterial microbiota and obesity. Clin. Microbiol. Infect. *19*, 305–313.

Miyake, D., Ichiki, S., Tanabe, M., Oda, T., Kuroda, H., Nishihara, H., and Sambongi, Y. (2007). Thiosulfate oxidation by a moderately thermophilic hydrogen-oxidizing bacterium, *Hydrogenophilus thermoluteolus*. Arch. Microbiol. *188*, 199–204.

Moreno-Vivián, C., Cabello, P., Martínez-Luque, M., Blasco, R., and Castillo, F. (1999). Prokaryotic nitrate reduction: molecular properties and functional distinction among bacterial nitrate reductases. J. Bacteriol. *181*, 6573–6584.

Moya, A., Peretó, J., Gil, R., and Latorre, A. (2008). Learning how to live together: genomic insights into prokaryote–animal symbioses. Nat. Rev. Genet. *9*, 218–229.

Mulligan, C., Geertsma, E.R., Severi, E., Kelly, D.J., Poolman, B., and Thomas, G.H. (2009). The substrate-binding protein imposes directionality on an electrochemical sodium gradient-driven TRAP transporter. Proc. Natl. Acad. Sci. *106*, 1778–1783.

Mulligan, C., Fischer, M., and Thomas, G.H. (2011). Tripartite ATP-independent periplasmic (TRAP) transporters in bacteria and archaea: TRAP transporters. FEMS Microbiol. Rev. *35*, 68–86.

Nelson, D.C., Waterbury, J.B., and Jannasch, H.W. (1982). Nitrogen fixation and nitrate utilization by marine and freshwater *Beggiatoa*. Arch. Microbiol. *133*, 172–177.

Nelson, D.C., Wirsen, C.O., and Jannasch, H.W. (1989). Characterization of large, autotrophic *Beggiatoa spp.* abundant at hydrothermal vents of the Guaymas Basin. Appl. Environ. Microbiol. *55*, 2909–2917.

Newton, I.L.G., Woyke, T., Auchtung, T.A., Dilly, G.F., Dutton, R.J., Fisher, M.C., Fontanez, K.M., Lau, E., Stewart, F.J., Richardson, P.M., et al. (2007). The *Calyptogena magnifica* chemoautotrophic symbiont genome. Science *315*, 998–1000.

Nussbaumer, A.D., Bright, M., Baranyi, C., Beisser, C.J., and Ott, J.A. (2004). Attachment mechanism in a highly specific association between ectosymbiotic bacteria and marine nematodes. Aquat. Microb. Ecol. *34*, 239–246.

Nussbaumer, A.D., Fisher, C.R., and Bright, M. (2006). Horizontal endosymbiont transmission in hydrothermal vent tubeworms. Nature *441*, 345–348.

O'Farrell, P.H. (1975). High resolution two-dimensional electrophoresis of proteins. J. Biol. Chem. *250*, 4007–4021.

Ogawa, T., Furusawa, T., Nomura, R., Seo, D., Hosoya-Matsuda, N., Sakurai, H., and Inoue, K. (2008). SoxAX binding protein, a novel component of the thiosulfate-oxidizing multienzyme system in the green sulfur bacterium *Chlorobium tepidum*. J. Bacteriol. *190*, 6097–6110.

Orth, R.J., Carruthers, T.J., Dennison, W.C., Duarte, C.M., Fourqurean, J.W., Heck, K.L., Hughes, A.R., Kendrick, G.A., Kenworthy, W.J., and Olyarnik, S. (2006). A global crisis for seagrass ecosystems. Bioscience *56*, 987–996.

Ott, J., Rieger, G., Rieger, R., and Enderes, F. (1982). New mouth less interstitial worms from the sulfide system: symbiosis with prokaryotes. Mar. Ecol. *3*, 313–333.

Page, H.M., Fiala-Medioni, A., Fisher, C.R., and Childress, J.J. (1991). Experimental evidence for filter-feeding by the hydrothermal vent mussel, *Bathymodiolus thermophilus*. Deep Sea Res. Part Oceanogr. Res. Pap. *38*, 1455–1461.

Parey, K., Demmer, U., Warkentin, E., Wynen, A., Ermler, U., and Dahl, C. (2013). Structural, biochemical and genetic characterization of dissimilatory ATP sulfurylase from *Allochromatium vinosum*. PLoS ONE *8*, e74707.

Petersen, J.M., Zielinski, F.U., Pape, T., Seifert, R., Moraru, C., Amann, R., Hourdez, S., Girguis, P.R., Wankel, S.D., Barbe, V., et al. (2011). Hydrogen is an energy source for hydrothermal vent symbioses. Nature *476*, 176–180.

Pires, R.H., Lourenço, A.I., Morais, F., Teixeira, M., Xavier, A.V., Saraiva, L.M., and Pereira, I.A.. (2003). A novel membrane-bound respiratory complex from *Desulfovibrio desulfuricans* ATCC 27774. Biochim. Biophys. Acta BBA - Bioenerg. *1605*, 67–82.

Polman, J.K., and Larkin, J.M. (1988). Properties of in vivo nitrogenase activity in *Beggiatoa alba*. Arch. Microbiol. *150*, 126–130.

Poole, R., and Hill, S. (1997). Respiratory protection of nitrogenase activity in *Azotobacter vinelandii* -roles of the terminal oxidases. Biosci. Rep. *17*, 303–317.

Pott, A.S., and Dahl, C. (1998). Sirohaem sulfite reductase and other proteins encoded by genes at the dsr locus of *Chromatium vinosum* are involved in the oxidation of intracellular sulfur. Microbiology *144*, 1881–1894.

Prange, A., Engelhardt, H., Trüper, H.G., and Dahl, C. (2004). The role of the sulfur globule proteins of *Allochromatium vinosum*: mutagenesis of the sulfur globule protein genes and expression studies by real-time RT-PCR. Arch. Microbiol. *182*.

Preisler, A., De Beer, D., Lichtschlag, A., Lavik, G., Boetius, A., and Jørgensen, B.B. (2007). Biological and chemical sulfide oxidation in a *Beggiatoa* inhabited marine sediment. ISME J. *1*, 341–353.

Ramos, A.R. (2012). The membrane QmoABC complex interacts directly with the dissimilatory adenosine 5′-phosphosulfate reductase in sulfate reducing bacteria. Front. Microbiol. *3*.

Records, A.R. (2011). The type VI secretion system: a multipurpose delivery system with a phage-like machinery. Mol. Plant. Microbe Interact. *24*, 751–757.

Reid, R. (1990). Evolutionary implications of sulphide oxidizing symbioses in bivalves. In: The Bivalvia: Proceedings of a Memoriam Symposium in Honour of Sir C.M. Yonge, Edingburgh, 1986, B. Morton, ed, Hong Kong University Press 127-140.

Reinartz, M., Tschäpe, J., Brüser, T., Trüper, H.G., and Dahl, C. (1998). Sulfide oxidation in the phototrophic sulfur bacterium *Chromatium vinosum*. Arch. Microbiol. *170*, 59–68.

Reshetnikov, A.S., Rozova, O.N., Khmelenina, V.N., Mustakhimov, I.I., Beschastny, A.P., Murrell, J.C., and Trotsenko, Y.A. (2008). Characterization of the pyrophosphate-dependent 6-phosphofructokinase from *Methylococcus capsulatus* Bath. FEMS Microbiol. Lett. *288*, 202–210.

Reynolds, L.K., Berg, P., and Zieman, J.C. (2007). Lucinid clam influence on the biogeochemistry of the seagrass *Thalassia testudinum* sediments. Estuaries Coasts *30*, 482–490.

Richter, M., and Rosselló-Móra, R. (2009). Shifting the genomic gold standard for the prokaryotic species definition. Proc. Natl. Acad. Sci. *106*, 19126–19131.

Richter, M., Lombardot, T., Kostadinov, I., Kottmann, R., Duhaime, M., Peplies, J., and Glöckner, F. (2008). JCoast – A biologist-centric software tool for data mining and comparison of prokaryotic (meta)genomes. BMC Bioinformatics *9*, 177.

Robidart, J.C., Bench, S.R., Feldman, R.A., Novoradovsky, A., Podell, S.B., Gaasterland, T., Allen, E.E., and Felbeck, H. (2008). Metabolic versatility of the *Riftia pachyptila* endosymbiont revealed through metagenomics. Environ. Microbiol. *10*, 727–737.

Robinson, J.J., Stein, J.L., and Cavanaugh, C.M. (1998). Cloning and sequencing of a form II ribulose-1,5-bisphosphate carboxylase/oxygenase from the bacterial symbiont of the hydrothermal vent tubeworm *Riftia pachyptila*. J. Bacteriol. *180*, 1596–1599.

Robson, R.L., and Postgate, J.R. (1980). Oxygen and hydrogen in biological nitrogen fixation. Annu. Rev. Microbiol. *34*, 183–207.

Rossi, F., Colao, E., Martinez, M.J., Klein, J.C., Carcaillet, F., Callier, M.D., Wit, R. de, and Caro, A. (2013). Spatial distribution and nutritional requirements of the endosymbiont-bearing bivalve *Loripes lacteus* (sensu Poli, 1791) in a Mediterranean Nanozostera noltii (Hornemann) meadow. J. Exp. Mar. Biol. Ecol. *440*, 108–115.

Rother, D., Henrich, H.-J., Quentmeier, A., Bardischewsky, F., and Friedrich, C.G. (2001). Novel genes of the sox gene cluster, mutagenesis of the flavoprotein SoxF, and evidence for a general sulfur-oxidizing system in *Paracoccus pantotrophus* GB17. J. Bacteriol. *183*, 4499–4508.

Russell, J.A., Latorre, A., Sabater-Muñoz, B., Moya, A., and Moran, N.A. (2003). Side-stepping secondary symbionts: widespread horizontal transfer across and beyond the Aphidoidea. Mol. Ecol. *12*, 1061–1075.

Sander, J., and Dahl, C. (2009). Metabolism of inorganic sulfur compounds in purple bacteria. In The Purple Phototrophic Bacteria, C.N. Hunter, F. Daldal, M.C. Thurnauer, and J.T. Beatty, eds. (Springer Netherlands), pp. 595–622.

Sander, J., Engels-Schwarzlose, S., and Dahl, C. (2006). Importance of the DsrMKJOP complex for sulfur oxidation in *Allochromatium vinosum* and phylogenetic analysis of related complexes in other prokaryotes. Arch. Microbiol. *186*, 357–366.

Schink, B., and Stams, A.J.M. (2006). Syntrophism among Prokaryotes. Prokaryotes Vol 2 N. Y. Springer 2006 309–335.

Schmehl, M., Jahn, A., zu Vilsendorf, A.M., Hennecke, S., Masepohl, B., Schuppler, M., Marxer, M., Oelze, J., and Klipp, W. (1993). Identification of a new class of nitrogen fixation genes in *Rhodobacter capsalatus*: a putative membrane complex involved in electron transport to nitrogenase. Mol. Gen. Genet. MGG *241*, 602–615.

Schmitz, R.A., Klopprogge, K., and Grabbe, R. (2002). Regulation of nitrogen fixation in *Klebsiella pneumoniae* and *Azotobacter vinelandii*: NifL, transducing two environmental signals to the nif transcriptional activator NifA. J. Mol. Microbiol. Biotechnol. *4*, 235–242.

Schneider, G., Lindqvist, Y., and Branden, C.-I. (1992). RuBisCO: structure and mechanism. Annu. Rev. Biophys. Biomol. Struct. *21*, 119–143.

Schroeter, R., Voigt, B., Jürgen, B., Methling, K., Pöther, D.-C., Schäfer, H., Albrecht, D., Mostertz, J., Mäder, U., Evers, S., et al. (2011). The peroxide stress response of *Bacillus licheniformis*. Proteomics *11*, 2851–2866.

Schweimanns, M., and Felbeck, H. (1985). Significance of the occurrence of chemoautotrophic bacterial endosymbionts in lucinid clams from Bermuda. Mar. Ecol.-Prog. Ser. *24*, 113–120.

Setubal, J.C., dos Santos, P., Goldman, B.S., Ertesvag, H., Espin, G., Rubio, L.M., Valla, S., Almeida, N.F., Balasubramanian, D., Cromes, L., et al. (2009). Genome sequence of *Azotobacter vinelandii*, an obligate aerobe specialized to support diverse anaerobic metabolic processes. J. Bacteriol. *191*, 4534–4545.

Shahak, Y. (1997). Sulfide-Quinone Reductase from *Rhodobacter capsulatus*. purification, cloning and expression. J. Biol. Chem. *272*, 9890–9894.

Sharon, N., and Lis, H. (1989). Lectins as cell recognition molecules. Science *246*, 227–234.

Shively, J.M., van Keulen, G., and Meijer, W.G. (1998). Something from almost nothing: carbon dioxide fixation in chemoautotrophs. Annu. Rev. Microbiol. *52*, 191–230.

Sorgo, A., Gaill, F., Lechaire, J., Arndt, C., and Bright, M. (2002). Glycogen storage in the *Riftia pachyptila* trophosome: contribution of host and symbionts. Mar. Ecol. Prog. Ser. *231*, 115–120.

Southward, E.C. (1986). Gill symbionts in Thyasirids and other bivalve molluscs. J. Mar. Biol. Assoc. U. K. *66*, 889–914.

Stapel, J., Aarts, T.L., Van Duynhoven, B.H., De Groot, J.D., Van den Hoogen, P.H., and Hemminga, M.A. (1996). Nutrient uptake by leaves and roots of the seagrass *Thalassia hemprichii* in the Spermonde Archipelago, Indonesia. Mar. Ecol. Prog. Ser. Oldendorf *134*, 195–206.

Stewart, E.J. (2012). Growing unculturable bacteria. J. Bacteriol. *194*, 4151–4160.

Stewart, F.J., and Cavanaugh, C.M. (2006a). Bacterial endosymbioses in *Solemya* (Mollusca: Bivalvia)—Model systems for studies of symbiont–host adaptation. Antonie Van Leeuwenhoek *90*, 343–360.

Stewart, F.J., and Cavanaugh, C.M. (2006b). Symbiosis of thioautotrophic bacteria with *Riftia pachyptila*. Prog. Mol. Subcell. Biol. *41*, 197–225.

Stewart, F.J., Newton, I.L.G., and Cavanaugh, C.M. (2005). Chemosynthetic endosymbioses: adaptations to oxic–anoxic interfaces. Trends Microbiol. *13*, 439–448.

Stockdreher, Y., Venceslau, S.S., Josten, M., Sahl, H.-G., Pereira, I.A.C., and Dahl, C. (2012). Cytoplasmic sulfurtransferases in the purple sulfur bacterium *Allochromatium vinosum*: evidence for sulfur transfer from DsrEFH to DsrC. PLoS ONE *7*, e40785.

Streams, M.E., Fisher, C.R., and Fiala-Medioni, A. (1997). Methanotrophic symbiont location and fate of carbon incorporated from methane in a hydrocarbon seep mussel. Mar. Biol. *129*, 465–476.

Suzuki, I. (1999). Oxidation of inorganic sulfur compounds: chemical and enzymatic reactions. Can. J. Microbiol. *45*, 97–105.

Szafranski, K.M., Gaudron, S.M., and Duperron, S. (2014). Direct evidence for maternal inheritance of bacterial symbionts in small deep-sea clams (Bivalvia: Vesicomyidae). Naturwissenschaften *101*, 373–383.

Tabita, F.R. (1999). Microbial ribulose-1,5-bisphosphate carboxylase/oxygenase: A different perspective. Photosynth. Res. *60*, 1–28.

Tabita, F.R., Satagopan, S., Hanson, T.E., Kreel, N.E., and Scott, S.S. (2007). 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. *59*, 1515–1524.

Tabita, F.R., Hanson, T.E., Satagopan, S., Witte, B.H., and Kreel, N.E. (2008). Phylogenetic and evolutionary relationships of RuBisCO and the RuBisCO-like proteins and the functional lessons provided by diverse molecular forms. Philos. Trans. R. Soc. B Biol. Sci. *363*, 2629–2640.

Taylor, J.D., and Glover, E.A. (1997). The lucinid bivalve genus *Cardiolucina* (Mollusca, Bivalvia, Lucinidae): systematics, anatomy and relationships. Bull. Nat. Hist. Mus. Zool. Ser. *63*, 93–122.

Taylor, J.D., Glover, E.A., and Williams, S.T. (2014). Diversification of chemosymbiotic bivalves: origins and relationships of deeper water Lucinidae. Biol. J. Linn. Soc. *111*, 401–420.

Teeling, H., Meyerdierks, A., Bauer, M., Amann, R., and Glockner, F.O. (2004). Application of tetranucleotide frequencies for the assignment of genomic fragments. Environ. Microbiol. *6*, 938–947.

Thorneley, R.N., and Ashby, G.A. (1989). Oxidation of nitrogenase iron protein by dioxygen without inactivation could contribute to high respiration rates of *Azotobacter* species and facilitate nitrogen fixation in other aerobic environments. Biochem J *261*, 181–187.

Thorneley, R.N., and Lowe, D.J. (1983). Nitrogenase of *Klebsiella pneumoniae*. Kinetics of the dissociation of oxidized iron protein from molybdenum-iron protein: identification of the rate-limiting step for substrate reduction. Biochem J *215*, 393–403.

Touchette, B.W., and Burkholder, J.M. (2000). Review of nitrogen and phosphorus metabolism in seagrasses. J. Exp. Mar. Biol. Ecol. *250*, 133–167.

Tseng, T.-T., Tyler, B.M., and Setubal, J.C. (2009). Protein secretion systems in bacterial-host associations, and their description in the gene ontology. BMC Microbiol. *9*, S2.

Waycott, M., Duarte, C.M., Carruthers, T.J., Orth, R.J., Dennison, W.C., Olyarnik, S., Calladine, A., Fourqurean, J.W., Heck, K.L., Hughes, A.R., et al. (2009). Accelerating loss of seagrasses across the globe threatens coastal ecosystems. Proc. Natl. Acad. Sci. *106*, 12377–12381.

Weinberg, M.V., Jenney, F.E., Cui, X., and Adams, M.W.W. (2004). Rubrerythrin from the hyperthermophilic *Archaeon Pyrococcus* furiosus Is a rubredoxin-dependent, iron-containing peroxidase. J. Bacteriol. *186*, 7888–7895.

Wentrup, Cécilia (2012) thesis:"Acquisition and activity of bacterial symbionts in marine invertebrates", Universität Bremen.

Weissgerber, T., Zigann, R., Bruce, D., Chang, Y., Detter, J.C., Han, C., Hauser, L., Jeffries, C.D., Land, M., Munk, A.C., et al. (2011). Complete genome sequence of *Allochromatium vinosum* DSM 180T. Stand. Genomic Sci. *5*, 311–330.

Welte, C., Hafner, S., Krätzer, C., Quentmeier, A., Friedrich, C.G., and Dahl, C. (2009). Interaction between Sox proteins of two physiologically distinct bacteria and a new protein involved in thiosulfate oxidation. FEBS Lett. *583*, 1281–1286.

Wolff, S., Hahne, H., Hecker, M., and Becher, D. (2008). Complementary analysis of the vegetative membrane proteome of the human pathogen *Staphylococcus aureus*. Mol. Cell. Proteomics *7*, 1460–1468.

Wolk, C.P., Ernst, A., and Elhai, J. (2004). Heterocyst metabolism and development. In the molecular biology of Cyanobacteria, D.A. Bryant, ed. (Springer Netherlands), pp. 769–823.

Wood, A.P., Aurikko, J.P., and Kelly, D.P. (2004). A challenge for 21st century molecular biology and biochemistry: what are the causes of obligate autotrophy and methanotrophy? FEMS Microbiol. Rev. *28*, 335–352.

Woyke, T., Teeling, H., Ivanova, N.N., Huntemann, M., Richter, M., Gloeckner, F.O., Boffelli, D., Anderson, I.J., Barry, K.W., Shapiro, H.J., et al. (2006). Symbiosis insights through metagenomic analysis of a microbial consortium. Nature *443*, 950–955.

Zhang, Y., Wen, Z., Washburn, M.P., and Florens, L. (2010). Refinements to label free proteome quantitation: How to deal with peptides shared by multiple proteins. Anal. Chem. *82*, 2272–2281.

Zhao, W., Ye, Z., and Zhao, J. (2007). RbrA, a cyanobacterial rubrerythrin, functions as a FNR-dependent peroxidase in heterocysts in protection of nitrogenase from damage by hydrogen peroxide in Anabaena sp. PCC 7120. Mol. Microbiol. *66*, 1219–1230.

6. Supplementary material

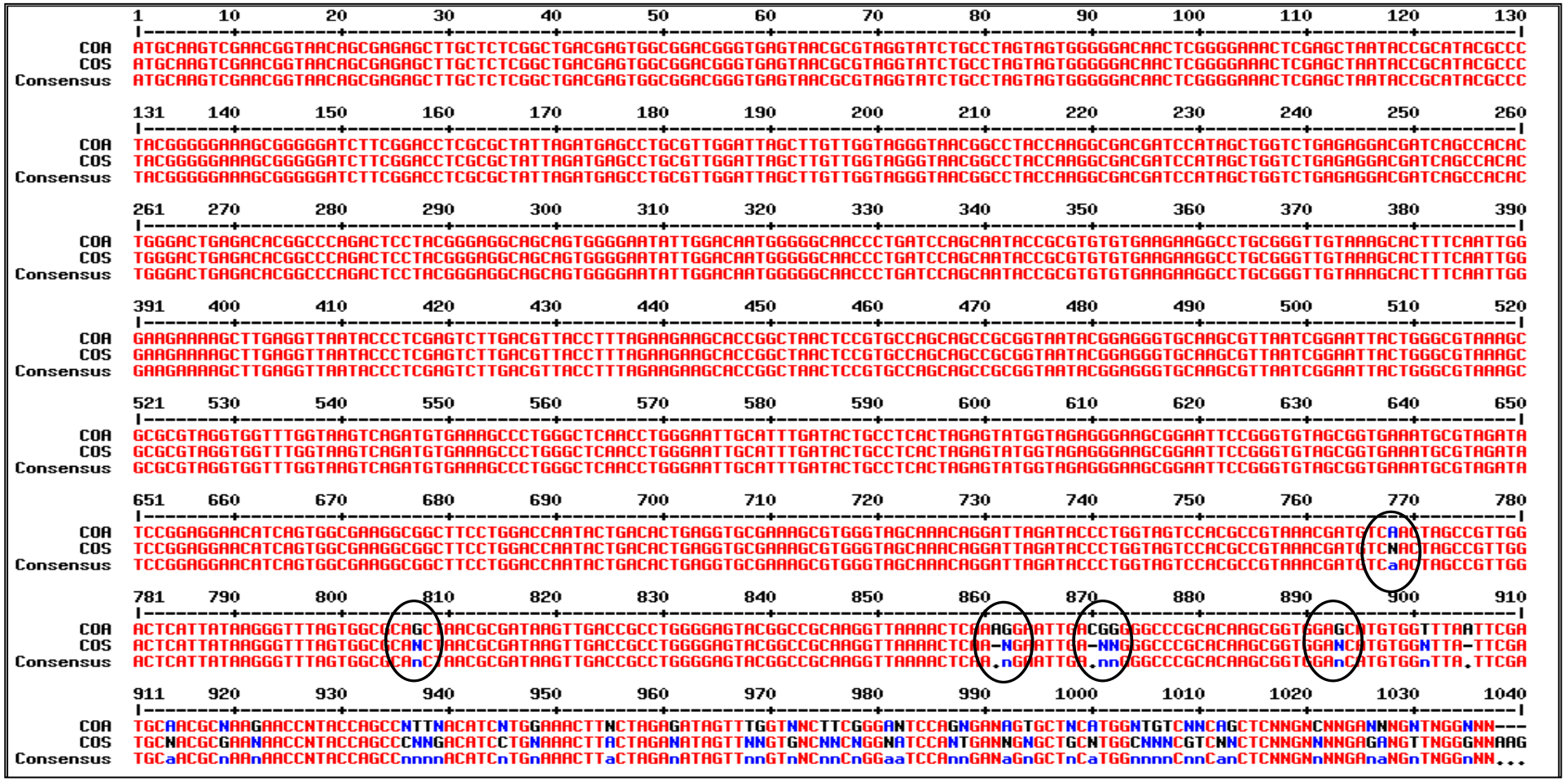

Figure 6.1. 16S rDNA comparison of the symbionts of *C. orbicularis* and *C. orbiculata*.
The differences between both symbionts are mainly caused by none determined nucleotide base (N) or a missing base (-) in the *C. orbicularis* symbiont sequence (black circles). (blue marked nucleotides are different, red marked nucleotides are equal, not sure sequenced nucleotides are marked as N)(http://multalin.toulouse.inra.fr/multalin/)(COA - *C. orbiculata*, COS - *C. orbicularis*)

Table 6.1. Putative sulfide dehydrogenase (FccAB) proteins encoded in the *C. orbicularis* symbiont genome.
Identity and eValue from NCBI protein blast. The signal peptide at the N-terminus of the protein destined the secretion with the sec-dep. or the Tat dep. secretory pathway (SigP). (Cello - subCELlular LOcalization prediction, PP - periplasm, IEP - isoelectric point, "Proteomics" - number of identifications in the symbiont proteome)

ID	Product	Label	IEP	MW (kDa)	Cello	TMH	SigP	Best NCBI blast hit	eValue	Identity %	Prote-omics
679	putative sulfide dehydrogenase (flavocytochrome), flavoprotein subunit	FccB	6.5	46.7	PP	1	yes	*Beggiatoa* sp.	$7.0e^{-101}$	54	0
680	putative cytochrome subunit of sulfide dehydrogenase	FccA	6.2	22.6	PP	1	yes	endosymbiont of *Tevnia jerichonana*	$8.0e^{-36}$	47	0
2539	putative flavocytochrome c subunit	FccA	6.2	22.6	PP	1	yes	*Thioalkalivibrio nitratreducens*	$2.0e^{-36}$	46	0
2540	putative sulfide dehydrogenase (flavocytochrome), flavoprotein subunit	FccB	6.5	46.7	PP	1	yes	*Beggiatoa* sp.	$7.0e^{-101}$	54	0
3220	putative sulfide dehydrogenase (flavocytochrome), flavoprotein subunit	FccB	9.4	45.7	PP	0	yes	*Thiocapsa marina* 5811	$9.0e^{-152}$	65	8
3221	putative cytochrome subunit of sulfide dehydrogenase	FccA	9.8	21.0	PP	0	no	endosymbiont of *Tevnia jerichonana*	$9.0e^{-33}$	49	8
3973	putative sulfide dehydrogenase (flavocytochrome), flavoprotein subunit	FccB	9.3	46.4	PP	0	yes	endosymbiont of *Tevnia jerichonana*	$9.0e^{-173}$	80	12
3974	putative cytochrome subunit of sulfide dehydrogenase	FccA	9.5	21.8	PP	0	yes	endosymbiont of *Tevnia jerichonana*	$4.0e^{-58}$	72	6

Table 6.2. Overview of the Sox-proteins encoded in the *C. orbicularis* symbiont genome.
Evalue and identity from NCBI protein Blast. (Cello - subCELlular LOcalization prediction, CP - cytoplasmic, PP - periplasma, Sec - SignalP 4.1 Server, Tat - TatP 1.0 Server)

ID	Product	Label	IEP	MW (kDa)	Tat	Sec	TMH	Cello	Best NCBI blast hit	eValue	Identity in %
563	sulfur oxidation protein SoxZ	SoxZ	10.3	11.2	no	yes	0	PP	endosymbiont *Tevnia jerichonana*	$1.0e^{-41}$	82
564	sulfur oxidation protein SoxY	SoxY	9.8	15.9	yes	yes	0	CP	endosymbiont *Tevnia jerichonana*	$3.0e^{-54}$	78
770	sulfur oxidation protein SoxX	SoxX	9.9	13.3	yes	yes	0	PP	endosymbiont of *Riftia pachyptila*	$1.0e^{-29}$	53
771	sulfur oxidation protein SoxA	SoxA	9.3	31.7	no	yes	0	PP	endosymbiont *Tevnia jerichonana*	$9.0e^{-128}$	75
772	soxAX binding protein	SoxAX	9.9	31.7	no	yes	0	PP	endosymbiont of *Riftia pachyptila*	$2.0e^{-7}$	55
773	sulfur transferase SoxL	SoxL	9.7	24.2	yes	yes	0	CP	*Allochromatium vinosum DSM 180*	$4.0e^{-70}$	65
1318	sulfur/thiosulfate oxidation protein SoxB	SoxB	5.2	63.0	no	yes	0	CP	endosymbiont *Tevnia jerichonana*	0.0	86
2485	thioredoxin SoxW	SoxW	5.7	39.1	no	yes	0	CP	*Thiorhodococcus drewsii AZ1*	$2.0e^{-98}$	56
4159	sulfur oxidation protein SoxZ	SoxZ	8.9	12.2	no	yes	0	CP	*Fulvimarina pelagi HTCCC2508*	$8.0e^{-23}$	44
4160	sulfur oxidation protein SoxY	SoxY	10.7	16.1	no	yes	0	PP	*Fulvimarina pelagi HTCCC2508*	$8.0e^{-33}$	45

Table 6.3. Sulfur metabolism proteins of the symbiont in *C. orbicularis*.
Evalue and identity from NCBI protein blast. The signal peptide at the N-terminus of the protein destined the secretion with the sec-dep. or the Tat dep. secretory pathway (SigP). (Cello - subCELlular LOcalization prediction, CP - cytoplasm, IM - inner membrane, "Proteomics" - number of identifications in the symbiont proteome, SigP - SignalP)

ID	Product	Label	IEP	MW (kDa)	Cello	TMH	SigP	Best NCBI blast hit	eValue	Identity in %	Prote-omics
3699	adenylylsulfate reductase, beta subunit	AprB	7.5	15.3	CP	0	no	endosymbiont of *Riftia pachyptila*	$3.0e^{-66}$	93	12
3700	adenylylsulfate reductase, alpha subunit	AprA	5.2	74.2	CP	0	no	endosymbiont of *Tevnia jerichonana*	0.0	92	12
145	sulfate adenylyltransferase	Sat/SopT	4.6	43.6	CP	0	no	endosymbiont of *Tevnia jerichonana*	0.0	87	11
1792	high affinity sulfate transporter	SulP	9.4	64	IM	12	no	endosymbiont of *Tevnia jerichonana*	0.0	69	6
2768	high affinity sulfate transporter	SulP	5.3	57.3	IM	11	no	*Beggiatoa sp.PS*	$2.0e^{-38}$	33	0
2280	putative sulfide-quinone reductase	Sqr	6	46.9	CP	0	no	endosymbiont of *Tevnia jerichonana*	0.0	90	3

Table 6.4. *dsr* gene cluster encoded in the *C. orbicularis* symbiont genome.
(Cello - subCELlular LOcalization prediction, CP - cytoplasm, IM - inner membrane, PP - periplasm, "Proteomics" - number of identifications in the symbiont proteome, , gene clusters are indicated by double line)

ID	Product	Label	strand	IEP	MW (KDa)	TMH	Cello	Prote-omics
533	DsrC-like protein	DsrC	-	5.5	12.8	0	CP	5
537	dissimilatory sulfite reductase, alpha subunit	DsrA	+	4.6	46.1	0	CP	12
538	dissimilatory sulfite reductase, beta subunit	DsrB	+	7.7	40.0	0	PP	12
539	sulfur oxidation protein DsrE, TusD-like sulfur relay protein	DsrE	+	5.1	14.4	0	CP	10
540	sulfur oxidation protein DsrF, TusC-like sulfur relay protein	DsrF	+	3.7	15.7	0	CP	9
541	sulfite reduction-associated protein DsrH	DsrH	+	4.7	10.9	0	PP	6
542	dissimilatory sulfite reductase (desulfoviridin), gamma subunit	DsrC	+	4.5	12.2	0	CP	12
543	sulfur oxidation protein DsrM	DsrM	+	10	27.5	5	IM	8
544	sulfur oxidation protein DsrK	DsrK	+	4.8	57.4	0	CP	8
545	sulfite reduction-associated protein DsrL	DsrL	+	4.5	71.2	0	CP	12
546	sulfite reduction-associated complex DsrMKJOP iron-sulfur protein DsrO	DsrO	+	9.4	27.0	1	PP	9
547	sulfite reduction-associated complex DsrMKJOP protein DsrP	DsrP	+	8.9	43.3	10	IM	0
548	sulfite reduction-associated protein DsrN	DsrN	+	5.1	49.2	0	CP	0
549	sulfite reduction-associated protein DsrR	DsrR	+	3.7	12.0	0	CP	4
550	sulfur oxidation-associated protein DsrS	DsrS	+	4.9	40.9	0	CP	0

Table 6.5. *nif* gene cluster in *C. orbicularis* symbiont genome.
Evalue and Identity from NCBI protein blast. (CELLO - subCELlular LOcalization prediction, CP - cytoplasm (grey color - on the plus strand)

ID	Product	Label	Stra-nd	IEP	MW (kDa)	Cello	TMH	Best NCBI blast hit	eValue	Identity in %
2593	nitrogen fixation protein NifM	NifM	-	4.9	33.0	CP	0	*Thiothrix nivea DSM 5205*	$3.0e^{-75}$	50
2594	NifZ family protein	NifZ	-	5	17.1	CP	0	*Allochromatium vinosum DSM 180*	$5.0e^{-45}$	63
2595	nitrogenase-stabilizing/ protective protein NifW	NifW	-	3.9	13.3	CP	0	*Thiocapsa marina 5811*	$5.0e^{-26}$	66
2596	hypothetical protein	HP_2596	-	4.6	19.8	CP	0	-----	-	-
2597	serine O-acetyltransferase	CysE	-	8.7	28.9	CP	0	*Thiorhodococcus drewsii AZ1*	$1.0e^{-87}$	71
2598	homocitrate synthase	NifV	-	5.5	42.5	CP	0	*Allochromatium vinosum*	$3.0e^{-135}$	67
2599	cysteine desulfurase NifS	NifS	-	5.8	44.4	CP	0	*Thiocystis violascens DSM 198*	0.0	84
2600	Fe-S cluster assembly protein NifU	NifU	-	4.3	32.0	CP	0	*Thiocystis violascens DSM 198*	$5.0e^{-116}$	76
2601	iron-sulfur cluster assembly accessory protein	2601	-	4.3	11.6	CP	0	*Thiothrix nivea DSM 5205*	$2.0e^{-35}$	65
2602	hypothetical protein, putative AlG2 protein	HP_2602	-	9.7	14.4	CP	0	*Thiorhodococcus drewsii AZ1*	$1.0e^{-33}$	61
2603	hypothetical protein	HP_2603	-	5.9	14.7	CP	0	*Beggiatoa alba B18LD*	-	-
2604	ferredoxin 2Fe-2S	2604	-	5.7	12.5	CP	0	*Methylobacter tundripaludum SV96*	$2.0e^{-31}$	58
2605	ferredoxin III 4[4Fe-4S] nif-specific	2605	-	3.7	12.2	CP	0	*Thiocystis violascens DSM 198*	$1.0e^{-24}$	58
2606	putative nitrogen fixation protein	2606	-	4.6	78.5	CP	0	-	-	-
2607	nitrogen fixation protein NifX-associated protein	NifX	-	4.9	17.5	CP	0	*Thiothrix nivea DSM 5205*	$8.0e^{-65}$	73
2608	positive regulator of sigma(E) RseC/MucC	2608	-	9.4	16.9	CP	2	*Thiocapsa marina 5811*	$1.0e^{-15}$	41
2609	dinitrogenase iron-molybdenum cofactor biosynthesis protein NifX	NifX	-	4.4	15.4	CP	0	*Thiothrix nivea DSM 5205*	$1.0e^{-42}$	66
2610	nitrogenase molybdenum-iron cofactor scaffold and assembly protein NifN	NifN	-	5.4	50.9	CP	0	*Allochromatium vinosum DSM 180*	0.0	69

ID	Product	Label	Strand	IEP	MW (kDa)	Cello	TMH	Best NCBI blast hit	eValue	Identity in %
2611	nitrogenase molybdenum-iron cofactor scaffold and assembly protein NifE	NifE	-	6.6	51.1	CP	0	*Thiorhodococcus drewsii -AZ1*	0.0	84
2612	hypothetical protein	HP_2612	-	4.8	35.9	CP	0	*Allochromatium vinosum DSM 180*	$2\text{-}0e^{-89}$	52
2613	putative glycosyl transferase family 3	Gt_3	-	4.9	36.8	CP	0	*Beggiatoa sp. PS*	$1.0e^{-81}$	47
2614	nitrogen fixing leucine rich variant repeat 4Fe-4S cluster protein	2614	-	4.9	28.4	CP	0	*Thiorhodococcus drewsii AZ1*	$7.0e^{-79}$	67
2615	hypothetical protein	HP_2615	-	4	10.1	CP	0	*Thiothrix nivea DSM 5205*	$6.0e^{-15}$	59
2616	dinitrogenase iron-molybdenum cofactor biosynthesis protein	NifY	-	5	25.0	CP	0	*Thiorhodococcus drewsii AZ1*	$2.0e^{-77}$	62
2617	4Fe-4S ferredoxin iron-sulfur binding domain-containing protein	2617	-	3.5	6.9	CP	0	*Thioflaviococcus mobilis 8321*	$6.0e^{-20}$	79
2618	putative nitrogen fixation protein FixT	FixT	-	4.5	7.9	CP	0	*Thiorhodovibrio sp. 970*	$5.0e^{-25}$	79
2619	nitrogenase molybdenum-iron protein subunit beta	NifK	-	5.2	59.5	CP	0	*Thiothrix nivea DSM 5205*	0.0	85
2620	nitrogenase molybdenum-iron protein subunit alpha	NifD	-	5.5	55.3	CP	0	*Thiocapsa marina 5811*	0.0	90
2621	nitrogenase molybdenum-iron protein	NifH	-		31.5	CP	0	*Azotobacter chroococcum mcd 1*	$5.0e^{-153}$	92
2622	NAD(+)--dinitrogen-reductase ADP-D-ribosyltransferase	2622	+	5.6	31.8	CP	0	*Thiocapsa marina 5811*		
2623	hypothetical protein	HP_2623	+	9.5	19.6	CP	0	-	-	-
2624	hypothetical protein	HP_2624	+	4.7	32	CP	0	-	-	-
2625	amidohydrolase 2	2625	+	5.5	8.9	CP	0	*Thiorhodococcus drewsii AZ 1*	$4.0e^{-17}$	59
2626	hemerythrin metal-binding domain-containing protein	2626	+	5.7	16.0	CP	0	*Teredinbacter turnerae T7901*	$3.0e^{-30}$	51
2627	ADP-ribosyl-(dinitrogen reductase) hydrolase	2627	-	6.1	33.5	CP	0	*Thiorhodococcus drewsii AZ1*	$2.0e^{-90}$	56
2628	nitrogenase FeMo cofactor synthesis molybdenum delivery protein	NifQ	-	4.8	21.6	CP	0	*Thiocapsa marina 5811*	$3.0e^{-51}$	54

ID	Product	Label	Stra-nd	IEP	MW (kDa)	Cello	TMH	Best NCBI blast hit	eValue	Identity in %
2629	ferredoxin	2629	-	4.4	13.3	CP	0	*Thiothrix nivea DSM 5205*	$1.0e^{-49}$	82
2630	ferredoxin	2630	-	5.2	10.2	CP	0	*Thiocystis violascens DSM 198*	$3.0e^{-47}$	96
2631	putative flavoprotein	2631	-	5.2	47.8	CP	0	*Thiocystis violascens DSM 198*	0.0	81
2632	nitrogenase-associated protein NifO	NifO	-	4.3	15.8	CP	0	*Thiorhodovibrio sp. 970*	$7.0e^{-35}$	53
2633	4Fe-4S ferredoxin iron-sulfur binding domain-containing protein. nitrogenase-associated	2633	-	4.1	9.9	CP	0	*Thiorhodococcus drewsii AZ1*	$1.0e^{-34}$	78
2634	nitrogenase cofactor biosynthesis protein NifB	NifB	-	5.2	54.9	CP	0	*Thiothrix nivea DSM 5205*	0.0	82
2635	transcriptional regulator. NifA subfamily Fis family	NifA	-	5.2	57.4	CP	0	*Thiorhodococcus drewsii AZ1*	0.0	68
2636	nitrogen fixation negative regulator NifL	NifL	-	4.9	59.7	CP	0	*Thiocapsa marina 5811*	$3.0e^{-145}$	51

Table 6.6. Oxidative phosphorylation related proteins in *C. orbicularis and C. orbiculata* symbiont proteome.
C. orbicularis bivalves were incubated for one week under sulfide starvation in sterile sea water with saturated oxygen (+[O_2]) and low oxygen (-[O_2]) conditions.The data represent the mean ± standard deviation of three independent experiments each performed in three technical replicates measured with the Synapt G2 mass spectrometer. The protein amount is measured in fmol/ng. Ratios more than -0,5/0,5 indicate a significant change of the protein expression. Proteins with significant positive log_2 ratio are up-regulated (green), proteins with significant negative log_2 ratio are down-regulated (red). Lower ratios are marked depending of the up- and down-regulation in lighter green and red colors. (CELLO - subCELlular LOcalization prediction, CP - cytoplasm, IM - inner membrane, EC - extracellular, PP - periplasm, OM - outer membrane, "Proteomics" - number of identifications in the symbiont proteome, gene cluster indicated by double lines) (COS - C. orbicuöaris, COA - C. orbiculata)

							Freshly (control)				COS incubation				Ratio log_2			
ID	Product	Label	IEP	MW (kDa)	Cello	TMH	COA	+/-SD	COS	+/-SD	+[O_2]	+/-SD	-[O2]	+/-SD	COS/ COA	+[O_2]/ -[O_2]	COS control/ COS +[O_2]	Prote-omics
821	NADH-quinone oxidoreductase subunit B		7.6	27.3	CP	0												6
566	cytochrome c family protein		7.4	20.3	PP	0	0.0370	0.0125	0.0650	0.0160	0.0140	0.0076	0.0210	0.0150	0.82	-0.596	2.214	8
577	cytochrome c oxidase, cbb3-type, subunit I		8.7	53.5	IM	12												6
578	cytochrome c oxidase, cbb3-type subunit II		5.8	27	PP	1	0.0160	0.0015										9
580	cytochrome c oxidase subunit CcoP	FixP	4.7	30.7	PP	1	0.0260	0.0101										9
581	type cbb3 cytochrome oxidase biogenesis protein	CcoG	9	56.7	IM	5												1
582	FixH family protein	FixH	4.5	19	PP	1												5
583	type cbb3 cytochrome oxidase biogenesis protein	CcoI	4.9	91	IM	7												2
584	cytochrome oxidase maturation protein, cbb3-type		4.2	7.4	CP	1												0
1067	ATP synthase subunit B	AtpB	5.3	17.4	CP	1	0.0140	0.0029										9
1068	ATP synthase subunit delta	AtpD	4.2	19.2	CP	0	0.0240	0.0107	0.1840	0.0301	0.0600	0.0232	0.0620	0.0692	2.934	-0.054	1.628	8
1069	F0F1 ATP synthase subunit alpha	AtpA	4.8	55.6	CP	0	0.1540	0.0261	0.1360	0.0120	0.0550	0.0188	0.0650	0.0401	-0.181	-0.237	1.292	12
1070	F0F1 ATP synthase subunit gamma	AtpG	7	31.9	CP	0	0.0360	0.0097					0.0140	0.0065				9

							Freshly (control)				COS incubation				Ratio log_2			
ID	Product	Label	IEP	MW (kDa)	Cello	TMH	COA	+/-SD	COS	+/-SD	+[O_2]	+/-SD	-[O2]	+/-SD	COS/ COA	+[O_2]/ -[O_2]	COS control/ COS +[O_2]	Prote-omics
1071	F0F1 ATP synthase subunit beta	AtpB	4.6	49.4	CP	0	0.2000	0.0192	0.2570	0.0803	0.1910	0.0330	0.1940	0.2148	0.359	-0.019	0.427	12
1072	ATPase, F1 complex, delta/epsilon subunit	AtpD	4.9	15.3	CP	0	0.0240	0.0043	0.0650	0.0087	0.0820	0.0205	0.0520	0.0202	1.440	0.653	-0.324	6
1940	V-type ATP synthase subunit E	AtpE	8.8	24.1	CP	0												7
2090	cytochrome c1		5	27.8	CP	2												8
2091	cytochrome b subunit of the bc complex		9	47.3	IM	9												7
2092	ubiquinol-cytochrome c reductase, iron-sulfur subunit		5.5	21.2	PP	1	0.0060	0.0041										8
2334	putative cytochrome c class I		4.8	16.9	PP	0	0.0200	0.0160			0.0860	0.0355	0.0620	0.0213		0.486		7
2582	cox2 cytochrome oxidase subunit 2		4.2	73.2	OM	1												8
2644	proton-translocating NADH-quinone oxidoreductase, chain N		5.2	52.3	IM	14												3
2645	proton-translocating NADH-quinone oxidoreductase, chain M		4.9	55.3	IM	14												0
2646	proton-translocating NADH-quinone oxidoreductase, chain L		6.2	72.7	IM	16												1
2647	NADH-ubiquinone oxidoreductase chain 4L/K		4.8	11.9	IM	3												1
2648	NADH-quinone oxidoreductase subunit J	NuoJ	5.2	21.9	IM	5												4
2649	NADH-quinone oxidoreductase subunit I	NuoI	8.9	18.7	CP	0												6

							Freshly (control)				COS incubation				Ratio log_2			
ID	Product	Label	IEP	MW (kDa)	Cello	TMH	COA	+/-SD	COS	+/-SD	+[O_2]	+/-SD	-[O2]	+/-SD	COS/ COA	+[O_2]/ -[O_2]	COS control/ COS +[O_2]	Prote-omics
2650	NADH-quinone oxidoreductase subunit H	NuoH	5.7	39.1	IM	8												3
2651	NADH-quinone oxidoreductase subunit G	NuoG	5	84.2	CP	0	0.0257	0.0049										7
2652	NADH-quinone oxidoreductase subunit F	NuoF	5.6	46.6	CP	0	0.0273	0.0110	0.0192	0.0036	0.0115	0.0038	0.0131	0.0085	-0.5092	-0.1877	0.7409	11
2653	NADH-quinone oxidoreductase subunit E	NuoE	4.5	18.8	CP	0							0.0104	0.0009				3
2654	NADH dehydrogenase I subunit D	NuoD	5.3	47.4	CP	0	0.0197	0.0092					0.0209	0.0009				10
2655	NADH-quinone oxidoreductase subunit C	NuoC	4.2	27.1	CP	0	0.0083	0.0030										8
2656	NADH-quinone oxidoreductase subunit B	NuoB	9.6	14.1	PP	0												5
2657	NADH-quinone oxidoreductase subunit A	NuoA	4.2	13.6	IM	3												0
2841	cytochrome c oxidase, subunit II		4.7	39.4	CP	2	0.0210	0.0088										9
3221	putative cytochrome subunit of sulfide dehydrogenase		9.8	21	PP	0												8
3377	cytochrome c class I		5.5	13.2	PP	0	0.0400	0.0423	0.0750	0.0047	0.0830	0.0305	0.0360	0.0282	0.887	1.217	-0.15	7
3607	cytochrome c4		9.1	21.5	PP	0												8

Table 6.7. Gene cluster of the motility genes in the *C. orbicularis* symbiont genome.
(CELLO - subCELlular LOcalization prediction, CP - cytoplasm, PP - periplasm, IM - inner membrane, OM - outer membrane, "Proteomics" - number of identifications in the symbiont proteome)

ID	Product	Label	IEP	MW (kDa)	Cello	TMH	Proteomics
53	flagellar hook-length control protein FliK	FliK	4.2	35.5	PP	0	0
258	flagellar motor rotation protein MotB	MotB	4.1	32.8	CP	1	0
259	flagellar motor rotation protein MotA	MotA	5.1	16.8	IM	2	0
1525	flagellar motor rotation protein MotA	MotA	4.4	8.6	IM	1	0
1526	flagellar motor rotation protein MotA	MotA	5.1	16.8	IM	2	0
1527	flagellar motor rotation protein MotB	MotB	4.1	32.8	CP	1	0
1876	flagellar basal body-associated protein FliL	FilL	5	14.2	CP	0	0
1916	flagellar export FliJ	FliJ	10.3	17.4	CP	0	0
1917	flagellum-specific ATP synthase FliI	FliI	4.8	50.9	CP	0	0
1918	flagellar assembly protein FliH	FliH	4.7	23.3	CP	0	0
1919	flagellar motor switch protein G	FliG	4.3	37.9	CP	0	0
1920	flagellar MS-ring protein	FliF	5.3	60.9	CP	2	0
1921	flagellar hook-basal body complex protein FliE	FliE	4.4	11.4	PP	0	0
1922	flagellar regulatory protein FleQ	FleQ	5.2	49.3	CP	0	0
1923	flagellar sensor histidine kinase FleS	FleS	5.2	44.2	CP	0	0
2952	flagellar protein FliS	FliS	4.9	13.8	CP	0	0
2953	flagellar hook-associated protein FliD	FliD	3.7	47.7	EC	0	0
2954	flagellar protein FlaG protein	FlaG	4.1	14.0	CP	0	0
2955	flagellin protein FlaA	FlaA	4.9	26.2	EC	0	0
2956	flagellar hook-associated protein FlgL	FlgL	4.1	33.3	CP	0	0
2957	flagellar hook-associated protein FlgK	FlgK	3.8	69.3	OM	0	0
2959	flagellar P-ring protein FlgI	FlgI	6.1	38.7	PP	0	0
2960	flagellar L-ring protein FlgH	FlgH	5.7	25.5	OM	0	0
2961	flagellar basal-body rod protein FlgG	FlgG	4.2	27.9	EC	0	0
2962	flagellar basal-body rod protein FlgF	FlgF	4	26.5	PP	0	0

ID	Product	Label	IEP	MW (kDa)	Cello	TMH	Proteomics
2963	flagellar hook protein FlgE	FlgE	3.9	44.2	EC	0	0
2965	flagellar basal-body rod protein FlgC	FlgC	4.6	14.8	EC	0	0
2966	flagellar basal-body rod protein FlgB	FlgB	7.5	14.4	OM	0	0
2967	methylase of chemotaxis methyl-accepting protein	CheR	9.6	31.0	OM	0	0
2968	response regulator receiver modulated CheW protein	CheW	4.5	33.9	CP	0	0
2969	flagella basal body P-ring formation protein FlgA	FlgA	10.6	25.7	PP	0	0
2970	conserved hypothetical protein	HP	9.2	8.7	CP	0	0
2971	putative flagella synthesis protein FlgN	FlgN	5.9	18.1	EC	0	0
4039	flagellar basal body-associated protein FliL	FliL	4.2	18.9	CP	1	0
4040	flagellar motor switch protein FliM	FliM	4.6	38.9	CP	0	0
4041	flagellar motor switch protein FliN	FliN	3.8	14.6	CP	0	0
4042	flagellar biosynthesis protein. FliO	FliO	10.3	14.4	IM	1	0
4043	flagellar biosynthetic protein FliP	FliP	3.8	26.6	IM	5	0
4044	flagellar biosynthesis protein FliQ	FliQ	4	9.6	IM	2	0
4045	flagellar biosynthetic protein FliR	FliR	5.3	28.3	IM	6	0
4046	flagellar biosynthetic protein FlhB	FlhB	8.6	42.5	IM	5	0
4047	flagellar biosynthesis protein FlhA	FlhA	4.7	75.0	IM	7	0
4048	flagellar biosynthesis protein FlhF	FlhF	4.7	26.4	CP	0	0
4049	flagellar biosynthesis protein FlhF	FlhF	6.3	21.7	CP	0	0
4050	flagellar synthesis regulator FleN	FleN	5.5	32.3	CP	0	0
4051	flagellar biosynthesis sigma factor FliA	FliA	4.7	28.1	CP	0	0
4052	chemotaxis protein CheY	CheY	4.5	14.5	CP	0	0
4053	chemotaxis phosphatase CheZ	CheZ	4.1	27.1	CP	0	0
4054	chemotaxis protein histidine kinase CheA	CheA	4.4	83.7	CP	0	0
4055	response regulator receiver modulated CheB methylesterase	CheB	9.5	37.5	CP	0	0
4056	flagellar motor rotation protein MotA	MotA	4.7	26.5	IM	4	0
4057	OmpA/MotB domain-containing protein	OmpA	5.7	31.0	OM	1	0

ID	Product	Label	IEP	MW (kDa)	Cello	TMH	Proteomics
4058	ParA family protein	ParA	5.8	29.4	CP	0	0
4059	hypothetical protein	HP	4	10.6	CP	0	0
4060	purine-binding chemotaxis protein CheW	CheW	3.8	17.8	CP	0	2

Table 6.8. Encoded ABC transporter in the *Codakia* symbiont genome.
(CELLO - subCELlular LOcalization prediction, CP - cytoplasm, PP - periplasm, "Proteomics" - number of identifications in the symbiont proteome)

ID	Product	Label	IEP	MW (kDa)	Cello	TMH	Prote-omics
2110	phosphate ABC transporter periplasmic phosphate-binding protein PstS	PstS	9	37.6	PP	0	8
3312	phosphate ABC transporter ATP-binding protein		5.2	30.4	CP	0	7
3840	putative ABC transporter ATP-binding protein YrbF	YrbF	9.5	29.6	CP	0	6
3313	phosphate ABC transporter ATPase		4.8	30.0	CP	0	5
1302	molybdenum ABC transporter ATP-binding protein		6.5	39.4	CP	0	4
3331	putative ABC transporter ATP-binding protein		6.4	27.0	CP	0	4
4313	ferric iron ABC transporter iron-binding protein		9.6	36.3	PP	0	4
3315	phosphate ABC transporter inner membrane subunit PstC	PstC	4.4	34.0	IM	6	3
4478	ABC transporter-like protein		5	28.8	CP	0	2
2109	phosphate ABC transporter permease protein PstC	PstC	7.7	49.3	IM	9	2
3314	phosphate ABC transporter inner membrane subunit PstA	PstA	10.3	31.2	IM	6	2
3539	soluble lytic transglycosylase fused to an ABC-type amino acid-binding protein		6.1	52.2	CP	0	2
3672	vitamin B12 ABC transporter B12-binding component BtuF	BtuF	9.6	30.6	PP	0	2
326	putativer ABC transporter permease protein		6.6	41.2	IM	5	1
1388	periplasmic component of amino acid ABC-type transporter/signal transduction system		4.9	32.9	CP	1	1
1666	permease protein of ABC transporter		9.8	45.0	IM	4	1
2108	phosphate ABC transporter. permease protein PstA	PstA	6.5	46.6	IM	6	1
3289	nitrate ABC transporter permease protein		10.3	38.1	IM	6	1
3631	iron-regulated ABC transporter ATPase subunit SufC	SufC	5	28.5	CP	0	1
3754	lipopolysaccharide ABC transporter ATP-binding protein LptB	LptB	6	26.6	CP	0	1
3841	putative ABC transporter permease component YrbE	YrbE	5.9	27.4	IM	5	1
4168	multidrug ABC transporter ATPase		8.7	23.7	CP	0	1
4495	ABC-type tungstate transport system. permease component		6.1	13.6	PP	0	0
4706	efflux ABC transporter permease protein		5.3	45.3	IM	5	0
4794	ABC transporter periplasmic substrate-binding protein		9.7	30.6	CP	0	0
5001	ABC-type branched-chain amino acid transport system periplasmic component		6.2	46.6	PP	1	0

ID	Product	Label	IEP	MW (kDa)	Cello	TMH	Prote-omics
5002	ABC-type branched-chain amino acid transport system ATPase component		5.5	25.2	CP	0	0
5003	ABC-type branched-chain amino acid transport system ATPase component		5.3	24.3	CP	0	0
5401	ABC transporter permease protein		9.8	37.0	IM	7	0
5465	iron(III) ABC transporter permease protein		9.8	49.6	IM	10	0
5775	ABC-type antimicrobial peptide transport system. ATPase component		6	26.5	CP	0	0
524	phosphonate ABC transporter periplasmic phosphonate-binding protein		6.3	32.4	CP	0	0
787	zinc ABC transporter periplasmic-binding protein ZnuA	ZnuA	9.9	33.4	CP	1	0
798	putative phosphonate ABC transporter periplasmic phosphonate-binding protein		6.3	34.7	CP	1	0
831	putative phosphate/phosphonate ABC transporter periplasmic protein PhnD family protein		6.3	35.0	CP	0	0
876	putative ABC transporter (ATP-binding protein)		7.7	27.2	CP	0	0
1303	molybdate ABC transporter inner membrane subunit		6.5	24.4	IM	5	0
1304	molybdenum ABC transporter periplasmic molybdate-binding protein		6.5	27.3	CP	0	0
1395	sulfonate ABC transporter permease		8.7	23.6	IM	4	0
1426	efflux ABC transporter permease protein		5.3	45.3	IM	5	0
1878	peptide ABC transporter permease		5.2	33.0	IM	6	0
2206	putative ABC transporter ATP-binding protein YheS		5.8	71.9	CP	0	0
2249	oligopeptide/dipeptide ABC transporter ATPase subunit		9.6	73.1	CP	0	0
2258	fused permease/ATPase component of ABC transporter involved in Fe-S cluster assembly		5.9	65.8	IM	6	0
2397	sulfate/tungstate uptake family ABC transporter periplasmic substrate-binding protein		6.5	31.2	CP	0	0
2897	peptide ABC transporter ATPase		6.2	26.5	CP	0	0
2905	carbohydrate ABC transporter membrane protein 2		9.9	30.8	IM	6	0
3332	permease protein. ABC-type multidrug efflux transporter		9.7	30.4	IM	6	0
3406	multidrug ABC transporter permease		5.8	37.5	IM	6	0
3438	putative ABC transporter substrate-binding protein		5.7	36.4	CP	1	0
3632	Iron-regulated ABC transporter permease protein SufD	SufD	5.5	47.8	CP	0	0
3644	efflux ABC transporter permease protein		9.2	86.6	IM	9	0
3730	putative soluble lytic transglycosylase fused to an ABC-type amino acid-binding protein		10.2	21.9	PP	0	0

ID	Product	Label	IEP	MW (kDa)	Cello	TMH	Prote-omics
3875	ABC transporter permase		10	29.2	IM	6	0
3876	ABC transporter permase		9.8	30.8	IM	6	0
4273	phosphonate ABC transporter periplasmic phosphonate-binding protein		5.3	33.3	CP	0	0
4284	phosphate/phosphonate ABC transporter periplasmic protein		4.2	27.9	CP	0	0
4314	ferric iron ABC transporter. permease protein		9.6	58.0	IM	11	0

Genome ID	Score	Genome Name
572477.4	528	Allochromatium vinosum DSM 180
637616.3	365	Methylophaga thiooxidans DMS010
396588.3	281	Thioalkalivibrio sp. HL-EbGR7
396588.4	273	Thioalkalivibrio sp. HL-EbGR7
472759.3	264	Nitrosococcus halophilus Nc4
243233.4	262	Methylococcus capsulatus str. Bath
243233.7	255	Methylococcus capsulatus str. Bath
697282.3	202	Methylobacter tundripaludum SV96
765913.3	193	Thiorhodococcus drewsii AZ1
323261.3	192	Nitrosococcus oceani ATCC 19707
765911.4	191	Thiocystis violascens DSM 198
323261.6	187	Nitrosococcus oceani ATCC 19707
349124.5	187	Halorhodospira halophila SL1
473788.3	182	Nitrosococcus oceani AFC27
349124.8	181	Halorhodospira halophila SL1
187272.10	179	Alkalilimnicola ehrlichii MLHE-1
396595.4	176	Thioalkalivibrio sp. K90mix
187272.6	173	Alkalilimnicola ehrlichii MLHE-1
105559.4	172	Nitrosococcus watsoni C-113
314278.3	150	Nitrococcus mobilis Nb-231
314278.4	144	Nitrococcus mobilis Nb-231
225937.3	144	marine bacterium HP15
768671.3	141	Thiocapsa marina 5811
500637.6	126	Providencia rustigianii DSM 4541
314276.3	125	Idiomarina baltica OS145
314276.4	122	Idiomarina baltica OS145
1051003.3	122	Pseudomonas aeruginosa 19BR
1051004.3	119	Pseudomonas aeruginosa 213BR
1051005.3	116	Pseudomonas aeruginosa 9BR
443152.3	114	Marinobacter algicola DG893

Figure 6.2. SEED Viewer closest neighbor of the *Codakia* symbiont.