Is Proteorhodopsin a General Light-driven Stress Adaptation System for Survival in Cold Environments?

By

Shi Feng

Master of Applied Science

A thesis submitted in fulfilment of the requirements for the Degree of

Doctor of Philosophy

University of Tasmania

September 2014

Declaration of Originality

This thesis contains no material which has been accepted for a degree or diploma by the University or any other institution, except by way of background information and duly acknowledged in the thesis, and to the best of my knowledge and belief, contain no copy of material previously published or written by another person, except where due reference is made in the test of the thesis. This thesis does not contain any material that infringes copyright

Statement on Access to the Thesis

The authority of access statement should reflect any agreement which exists between the University and an external organisation (such as a sponsor of the research) regarding the work. Examples of appropriate statements are:

1. This thesis may be made available for loan and limited copying and communication in accordance with the Copyright Act 1968.

2. This thesis may be made available for loan. Copying of any part of this thesis is prohibited for two years from the date this statement was signed; after that time limited copying and communication is permitted in accordance with the Copyright Act 1968.

3. This thesis is not to be made available for loan or copying for two years following the date this statement was signed. Following that time the thesis

may be made available for loan and limited copying and communication in accordance with the Copyright Act 1968.

4. (A statement of conditions applying to loan and access for copying which is consistent with any existing (intellectual property or other kind of) agreements relating to the thesis or work reported in it.)

September, 2014

Statement on Published Work

The publisher of the papers comprising Chapters 2 and 3 hold the copyright for that content, and access to the material should be sought from the respective journals. The remaining unpublished content of the thesis may be made available for loan and limited copying in accordance with the *Copyright Act 1968.*

September, 2014

Statement of Co-Authorship

The following people and institutions contributed to the publication of work undertaken as part of this thesis:

Shi Feng, School of Land and Food, University of Tasmania (Candidate)

John Bowman, School of Land and Food, University of Tasmania (Supervisor)

Shane Powell, School of Land and Food, University of Tasmania (Co-supervisor)

Richard Wilson, Central Science Laboratory, University of Tasmania.

Author details and their roles:

Paper 1: Extensive gene acquisition in the extremely psychrophilic bacterial species *Psychroflexus torquis* and the link to sea-ice ecosystem specialism

Located in Chapter 2

Shi Feng (60%) designed and conducted the experiment, analysed the data and wrote the manuscript. Shane Powell (10%), and John Bowman (25%) contributed to experimental design, analysed the data, and edited the manuscript. Richard Wilson (5%) offered 1D LC/MS proteomic service and interpretation of proteomic raw data.

Paper 2: Light stimulated growth of proteorhodopsin bearing sea-ice psychrophile *Psychroflexus torquis* is salinity-dependent

Located in Chapter 3

Shi Feng (60%) designed and conducted the experiment, analysed the data and wrote the manuscript. Shane Powell (10%), and John Bowman (25%) contributed to experimental design, analysed the data, and edited the manuscript. Richard Wilson

(5%) offered 1D LC/MS proteomic service and interpretation of proteomic raw data.

Paper 3: Life in sea ice − proteomic insights into a proteorhodopsin-containing sea-ice dwelling flavobacteria

Shi Feng (70%) designed and conducted the experiment, analysed the data and wrote the manuscript. Shane Powell (15%), and John Bowman (15%) contributed to experiment design, analysed the data, edit the manuscript.

We the undersigned agree with the above stated "proportion of work undertaken" for each of the above published or submitted peer-reviewed manuscripts contributing to this thesis

(Assoc. Prof. John Bowman)	(Prof. Holger Meinke)
Supervisor	Head of School
School of Land and Food	School of Land and Food
University of Tasmania	University of Tasmania
September, 2014	

Acknowledgements

I would like to express my gratitude and sincerest thanks to the following people for their contributions and encouragement throughout my PhD.

My academic supervisors: Assoc. Prof. John P. Bowman and Dr. Shane M. Powell for being great mentors. The guidance and assistance they provided throughout the duration of my candidature has been greatly appreciated.

University of Tasmania for generous financial support.

Prof. Tom McMeekin for his critical discussion related to all manuscript and generous advice during my candidature.

Dr. Richard Wilson for his advice and help in the proteomic work.

Prof. David Ratkowsky, Dr. Jay Kocharunchitt Mrs. Lauri Parkinson, Mr. Andrew Measham, Mr Adam Smolenski within the Tasmanian Institute of Agriculture (TIA) and Central Science Laboratory for their advice and continuous assistance during this research.

My fellow colleagues in Food Safety Centre. Ali Al-Naseri, Bianca Porteus, Tuflikha Putri, Peipei Zhang, Kamarul Zarkasi, for their friendship, insightful discussion and comments, and constant support.

Last but not least, to my family, my beloved parents Zhongyi Feng and Wei Qi for supporting me and encouraging me as they always do. Also to my beloved partner Jing Chen for emotional support and taking care of me during this time.

Table of Contents

List of Abbreviations

CCCP	Carbonylcyanide-3-chlorophenylhydrazone
CHASE	Cyclase/histidine kinase associated sensory extracellular
CRISPR	Clustered regularly interspaced short palindromic repeats
CRT	CRISPR recognition tool
EPS	Exopolysaccharides
EPA	Eicosapentaenoic acid
GAF	cGMP-specific phosphodiesterases, adenylyl cyclases, and FhlA domain
GIs	Genomic islands
HGT	Horizontal gene transfer
IVYWREL	Ile, Val, Tyr, Trp, Arg, Glu, Leu
MS	Mass spectra, mass spectrometer or mass spectrometry
NCBI	National Center for Biotechnology Information
ORFs	Open reading frames
oriC	Origin of replication
PAS	Per-Arnt-Sim domain
PEA	2-phenylethylamine
PR	Proteorhodopsin
PUFA	Polyunsaturated fatty acid
SIMCO	Sea-ice microbial communities

Abstract

This study aimed to achieve a better understanding of microbial adaptations in sea ice focusing on the physiological role of the light harvesting proton pump proteorhodopsin. To carry out these aims the research mainly focused on exploring the genome biology, physiology and life strategy of the model sea-ice bacterial species *Psychroflexus torquis*, an extremely psychrophilic member of the family *Flavobacteriaceae* (phylum *Bacteroidetes*). *P. torquis* has a bipolar distribution and is only known to occur in polar sea-ice and associated polar waters. It possesses proteorhodopsin and is believed to have a predominantly epiphytic lifestyle, mainly dwelling on sea-ice diatoms in sea-ice basal assemblages. This study extensively used gel-free label-free based proteomic approach to explore *P. torquis*' genome biology and unravel the role of proteorhodopsin in aiding the species adaptation to the extreme sea ice environment.

Sea ice has been estimated to have only become a stable feature on Earth in the last few million years ago thus it has been hypothesized that bacteria adapted to sea-ice acquired or exchange survival traits via horizontal gene transfer (HGT) between other sea ice dwelling microorganisms relatively recently. To examine the question whether sea-ice bacteria, such as *P. torquis* are endemic and display sea ice-ecosystem specialism a comparison of *P. torquis'* genome to its very closely related (99% 16S rRNA gene sequence similarity) sister species, *P. gondwanensis* ACAM 44^{T}, which is only known to dwell in Antarctic hypersaline lakes, was performed. This comparison

allowed for the determination of the level of HGT, what traits show evidence of HGT, what traits are relevant to the sea-ice ecosystem, and whether these genes are highly expressed, which would be indicative of their biological importance to *P. torquis*. The results show that in *P. torquis* ATCC 700755^{T} (genome size 4.3 Mbp) HGT has occurred much more extensively compared to *P. gondwanensis* (genome size 3.3 Mbp) and genetic features that can be linked as a sea ice specific adaptation are mainly concentrated on numerous genomic islands absent from *P. gondwanensis*. Genes encoding sea-ice ecosystem relevant traits, such as secreted exopolysaccharide, poly-unsaturated fatty acids, and ice binding proteins, form gene clusters on a number of these genomic islands. Proteomic analysis revealed that the encoded proteins for many sea-ice relevant traits are highly abundant under standard laboratory growth conditions. The genomic islands feature comparatively low gene density, a high concentration of pseudogenes, repetitive genetic elements, and addiction modules, indicative of large scale HGT either via phage or conjugation driven insertions. The overall results suggest the extensive level and nature of gene acquisition in *P. torquis* indicates its potential evolution to sea-ice ecosystem specialism. In that respect *P. torquis* seems to be an excellent model to study sea-ice functional biology. The initial screening of the *P. torquis* ATCC 700755^{T} genome revealed the presence of a proteorhodopsin gene. Previous studies have demonstrated proteorhodopsin-based phototrophy can enhance bacterial growth and survival during nutrient-stress conditions. But proteorhodopsins are widespread in natural environments and these environments may have other stress conditions for which proteorhodopsin can be

advantageous. So it can be hypothesized that proteorhodopsin may provide growth/survival advantage under stress conditions that are associated with a specific econiche. Growth studies on proteorhodopsin-containing *P. torquis* have demonstrated for the first time that light-stimulated growth can be linked to salinity stress rather than nutrient limitation. In addition, proteorhodopsin abundance and associated proton-pumping ability is also salinity dependent. The results extend the existing hypothesis that light can provide energy for marine prokaryotes through proteorhodopsin under stress conditions other than nutrient stress.

To gain a deeper insight into the physiological role of proteorhodopsin and the life strategy of *P. torquis,* a gel-free label-free quantitative proteomic approach was used. Proteome analysis revealed how *P. torquis* responded to different salinity and illumination levels by regulating its energy generation, nutrient uptake transporters, adhesion ability and gliding motility. The protein expression patterns of *P. torquis* indicates that it can use light to gain an advantage in colonizing phytoplanktonic surfaces, taking up more nutrients, and optimizing energy production. This study provided a comprehensive understanding of life style in sea ice and also partly revealed the physiological role of proteorhodopsin and its complex interrelationships.

Summary of Major Findings

Through this study, we examined the genome and the protein expression level of model sea-ice bacterial species *P. torquis* bacteria and revealed a high level of HGT seems to have driven the evolution process by which it became adapted to the adapt

dynamic sea-ice environment. We also examined the physiological role of PR within *P. torquis*, and extended the current hypothesis that PR provides light derived energy for marine prokaryotes under specific stress condition within econiches rather than only being linked to nutrient limitation. Finally, we assessed the proteome of *P. torquis* under different levels of salinity and illumination and achieved a better understanding in the proteorhodopsin-phototrophy, life strategies and physiological processes of this microorganism. Overall, we conclude that PR can be a general light-driven stress adaptation system for survival in cold environments. However, further studies are required on additional proteorhodopsin-containing strains to provide more evidence regarding this question.

Publications and conference presentations relevant to the thesis

The work presented in this thesis has so far resulted in the following peer reviewed publications and conference proceedings.

Publications

Chapter 1. Feng S, Powell SM, Wilson R, Bowman JP. (2014). Extensive gene acquisition in the extremely psychrophilic bacterial species *Psychroflexus torquis* and the link to sea-ice ecosystem specialism. *Genome Biology and Evolution* **6**: 133-148.

Chapter 2. Feng S, Powell SM, Wilson R, Bowman JP. (2013). Light-stimulated growth of proteorhodopsin-bearing sea-ice psychrophile *Psychroflexus torquis* is salinity dependent. *The ISME Journal* **7**: 2206-2213.

Chapter 3. Feng S, Powell SM, Bowman JP. (2014). Life in sea ice − proteomic insights into a proteorhodopsin-containing sea-ice dwelling flavobacteria. (Manuscript submitted to peer view journal)

Conference proceedings

Feng, S. and Bowman, JP, ‘A novel light stimulated *Flavobacterium* species from an Antarctic saline lake’, Invited presentation to the 3rd *Flavobacterium* Conference FLAVOBACTERIUM 2012, 5-7 June 2012, Turku, Finland.

Feng, S, Powell, S and Wilson, RR and Bowman, JP ‘The Light-driven proton pump proteorhodopsin promotes growth of sea-ice psychrophile during salinity stress’, Proceedings of the Australian Society for Microbiology, Annual Scientific Meeting 7-10 July 2013.

Feng, S, Powell, S and Wilson, RK and Bowman, JP, ‘Genomic and proteomic insights of enhanced growth of a sea-ice psychrophile *Psychroflexus torquis* by proteorhodopsin pohotorophy during salinity stress’, Presentation to the the First EMBO Conference on Aquatic Microbial Ecology - SAME13, 8-13 September 2013, Stresa, Italy.

Chapter 1

Literature Review

Introduction

Rhodopsins are light-absorbing pigments that are retinal binding integral membrane proteins (Béjà et al., 2000). Rhodopsins are found in the Bacteria, marine Archaea and Eukarya (Frigaard et al., 2006, Slamovits et al., 2011). Rhodopsins are currently known to belong to two distinct protein families. The visual rhodopsins, found in eyes throughout the animal kingdom, are photosensory pigments. Archeaeal rhodopsins are found in extreme halophiles and function as light-driven proton pumps (bacteriorhodopsins), chloride ion pumps (halorhodopsins), or photosensory receptors (sensory rhodopsins) (Lanyi, 2004). These two kinds of protein showed no significant sequence similarity and may have different origins (Béjà et al., 2000). However, they do share identical topologies characterized by seven transmembrane α-helical domains that form a pocket in which retinal is covalently linked via protonated Schiff base, to a lysine located in the seventh transmembrane helixices (Béjà et al., 2000). Rhodopsin in microorganisms was first found in the archaea *Halobacterium salinarum* in the 1970s and later recorded in other archaea (Oesterhe and Stoecken, 1971). Archaeal rhodopsins can be functionally divided into two classes: ion-pumps and sensory rhodopsins (Spudich et al., 2000). The ion-pumps can be further classified as proton-pumps (bacteriorhodopsin, BR) and chloride pumps, which are driven by light (halorhodopsin, HR). Both BR and HR's maximum absorption

wavelength are at about 570 nm. BR transports H^+ out of membrane and establishes a proton gradient that can be used by ATP synthase to synthesize ATP, while HR can transport Cl^- into the cell under light to sustain the balance of pH within the cell in order to acclimatize the cell to extreme environments (Spudich et al., 2000).

Rhodopsins have recently been the focus of much interest due to the discovery of a new type of rhodopsin in bacterial species (Fuhrman et al., 2008). The discovery has changed our perception and understanding of light responses in marine ecology and greatly increased the estimated abundance and diversity of this protein family.

Proteorhodopsin (PR) is a photoactive retinylidene protein first discovered in 2000 through metagenomic analyses of large genome fragments derived from a marine bacterium of the uncultivated gammaproteobacterium SAR86 clade (Béjà et al., 2000). Subsequent screening of DNA from different locations revealed a very large diversity of PR in bacteria belonging to divergent clades of the classes *Alphaproteobacteria* and *Gammaproteobacteria* (Gomez-Consarnau et al., 2007, Venter et al., 2004). When PR were functionally expressed in *Escherichia coli*, the *E. coli* cells that were expressing PR become light-powered when cellular respiration was impaired (Walter et al., 2007). Additionally light driven pumping by PR can fully replace respiration as a cellular energy source in some environmental conditions (Walter et al., 2007). PR uses light energy to translocate protons across the cell membrane and the excess extracellular protons create a proton motive force that can energetically drive flagella motility, transport processes, or ATP synthesis in cells (DeLong and Béjà, 2010). Therefore, marine bacteria which harbor PR can potentially gain a competitive advantage using

this photosynthetic pathway to create a light-driven energy source.

Since the initial discovery, PR has been identified in many marine bacteria, archaea and even marine eukaryotes (Frigaard et al., 2006, Slamovits et al., 2011). In the recent Global Ocean Sampling expedition, more than 4000 PR sequences were identified from 41 distinct surface marine environments, indicating that these PR genes are extremely abundant in the genomes of ocean bacterioplankton (Rusch et al., 2007). It is estimated that PR-containing bacteria account for 13% of the community in the Mediterranean Sea and the Red Sea, and 70% in the Sargasso Sea (Campbell et al., 2008, Venter et al., 2004, Rusch et al., 2007). The most recent PR screen from Antarctic sea ice presented a result of 52 positive and distinct PR sequences out of one thousand analyzed clone samples, with 17 positive mRNA sequences and 35 genomic DNA sequences (Koh et al., 2010). Although it is difficult to quantify the direct influence of PR-based phototrophy on the ocean ecosystems there is no doubt it is significant (DeLong and Béjà, 2010).

Structure

The PR gene encodes a polypeptide of approximately 249 amino acids, with a molecular weight of 27kD (Béjà et al., 2000). Béjà et al. (2000) demonstrated that the hydropathy plots indicated that PR has seven transmembrane domains (Fig. 1) which is similar to the rhodopsin protein family that aligned well with the corresponding helices of the archaeal rhodopsins. The amino acid residues that form a retinal binding pocket in archaeal rhodopsins are also highly conserved in PR. In particular, the

critical lysine residue in helix G, which forms the Schiff base linkage with retinal in rhodopsins, is present in PR (Béjà et al., 2000). The complete three dimensional structure of this protein is still unknown due to the difficulties in the analysis of the seven transmembrane α helical membrane protein (Shi et al., 2009). A series of conformational shifts is facilitated upon absorption of light which causes the transport of a proton across the membrane resulting in electrochemical membrane potential which drives ATP synthesis via F_0F_1-ATPase (Béjà et al., 2001).

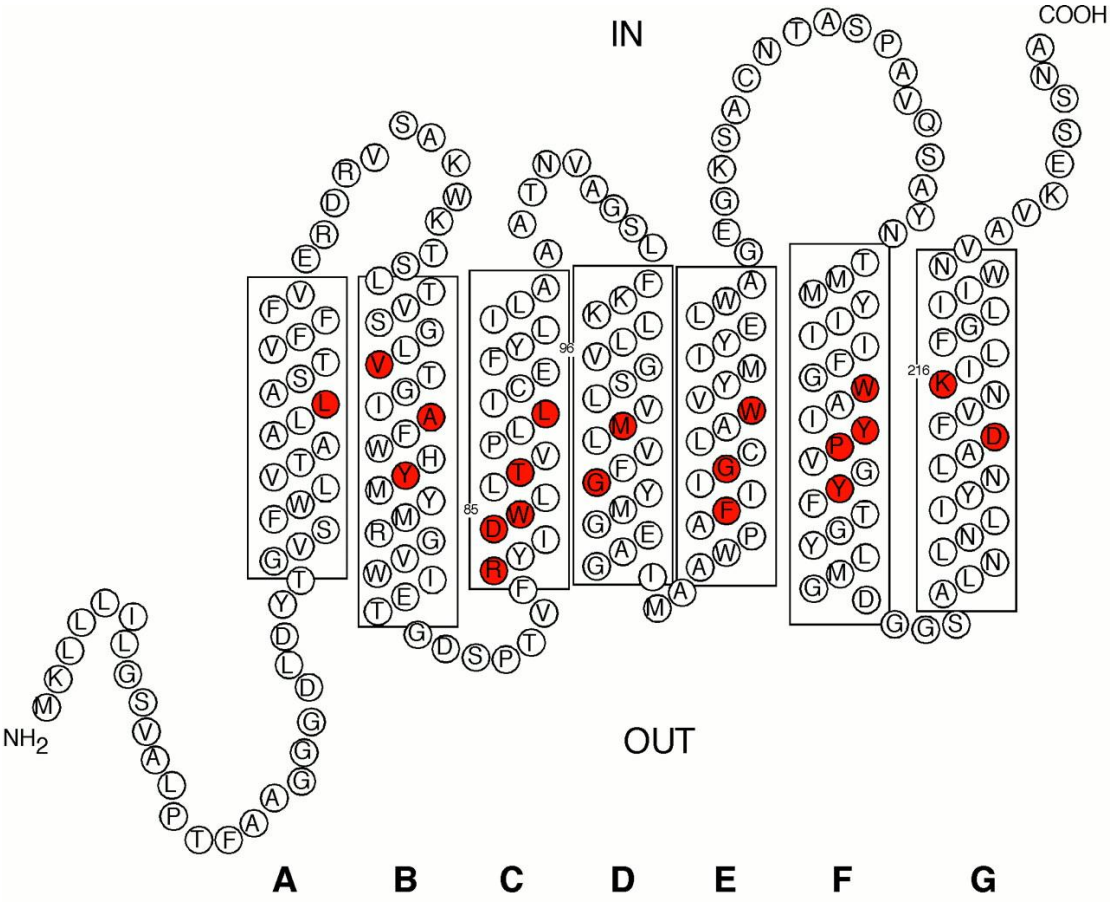


Figure.1 Secondary structure of PR. Predicted retinal binding pocket residues are marked in red. Taken from (Béjà et al., 2000)

PR has been found to share considerable functional and structural similarity with the archaeal BR. This extended protein family is distinct from the visual rhodopsins, and

differentiated from phototaxis receptors despite considerable homologies between residues forming the retinal-binding pocket (Friedrich et al., 2002). Several key amino acid positions necessary for energy generation are conserved among rhodopsins, functioning as proton or ion pumps, such as Lys 216 which binds retinal to helix G through a protonated Schiff base in BR, was conserved as Lys 230 and Asp 85 in PR. The proton acceptor from the Schiff base was conserved as Asp 97 (Gomez-Consarnau et al., 2007). This demonstrates conservation of most of the amino acid residues which are responsible for vectorial proton transfer. However,, Asp 96 in BR is replaced by Glu108 in PR, and those carboxylic acid residues involved in the proton release – although a number of Asp and Glu residues are in an appropriate position (Friedrich et al., 2002). PR and BR also showed considerable functional similarity typically of ion pumps. In PR Glu108 serves as a proton donor for re-protonation of the Schiff base rather than Asp96 in BR, and Asp97 in PR serves as proton acceptor instead of Asp 85 in BR (Friedrich et al., 2002). It has been demonstrated that PR contains two intermediates: M and O which both have a short turnover time of about 20ms (Béjà et al., 2000). PR can operate as an outward directed proton pump at alkaline pH and inwardly directed at acidic pH. Furthermore Asp97 in PR has a pK_a value of 7.68 (Friedrich et al., 2002). The protonation state of Asp97 is crucial for this bifunctional behavior by serving as the primary acceptor of the Schiff base proton under alkaline conditions.

Spectral Characteristics of PR

PR can be categorized into two major subgroups based on their visible absorption: green and blue. Their distribution was shown to be stratified with depth. Both "green-absorbing" and "blue-absorbing" PR share more than 78% of their amino acid residues, while the visible absorption peaks differ by almost 40nm between these two types (Man et al., 2003). PR present in near-surface seawater typically have leucine at position 105 and show absorption maximally around 525nm and thus represent the green light absorbing proteorhodopsins (GPR).GPRs are primarily light driven proton pumps. Spectroscopic and photochemical reactions have been characterized by a number of studies in its native membranes from Monterey Bay plankton and also expressed in *E. coli* (Béjà et al., 2000, Béjà et al., 2001, Dioumaev et al., 2002, Dioumaev et al., 2003, Lakatos et al., 2003). A variant group of PR genes differing by about 22% in predicted primary structure from the GPR has been detected in the Antarctic and deep ocean plankton (Béjà et al., 2001). Several clades isolated from deep ocean plankton exhibited a blue-shifted absorption spectrum around 490nm when PR gene expressed in *E. coli* which is identified as blue light absorbing proteorhodopsin (BPR). This type of PRs have glutamine at position 105 instead of leucine and are structurally similar to archeal sensory rhodopsin suggesting an analogous phototosensory function (Wang et al., 2003, Béjà et al., 2001, Man et al., 2003). The separation of these two subgroups is no doubt the result of selective pressure allowing each group to function optimally under different environmental conditions. GPR is generally expressed in surface water organisms where light

availability is optimal, while BPR functions further down the water column as light levels diminish (Wang et al., 2003).

A previous study has suggested that amino acid variation at just a single site (position 105) can change the spectral tuning, which made the difference between BPR and GPR (Man et al., 2003). A previous study showed a Red Sea variant (RS29) from a new family of PR that is composed of GPR type variants that possess Gln instead of Leu at position 105 (as in BPR). Its spectrum was expected to shift to a blue wavelength. However, unexpectedly, the maximum absorption of RS29 was 515 nm, which is a smaller blue shift compared to the 490 nm of GPR (Man-Aharonovich et al., 2004). They also found that there were four additional residue changes at positions 65 and 70 compared with GPR: proline instead of serine at position 65, isoleucine instead of valine at position 68, aspartic acid instead of glycine at position 70 and glutamine instead of leucine at position 105 which contribute a small red shift to the absorption spectrum of GPR (Man-Aharonovich et al., 2004) (Fig.2). Further analysis of position 105 indicated that PR absorbing green light carry other hydrophobic amino acids including methionine and alanine. (Gomez-Consarnau et al., 2007).

A

B

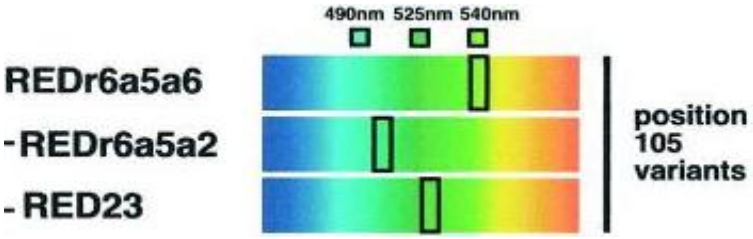


Figure. 2 Multiple protein alignment of PRs from different bacterial strain shown in A, position 105 is marked above and indicated variations. B suggests that different position 105 variants has a different spectral tuning (Man et al., 2003).

Despite the differences in spectroscopy, both GPR and BPR function as proton pumps as revealed by the laser flash-induced photolysis method. However, whether the new PR subfamily: G-B-PR is a proton-pump still remains unclear (Béjà et al., 2000, Béjà et al., 2001, Man-Aharonovich et al., 2004).

Laser flash-photolysis techniques based on the spectral analysis permit estimation of the phase of PR and cellular concentration of PR (Béjà et al., 2001). Laser flash induces the photochemical reaction which is a repeat cycle. The measurement of time-resolution absorption spectroscopy of intermediates in the cycle may characterize the intermediates and tell us what the intermediate is and when it appears. One photochemical cycle of GPR and BPR only takes 20 ms or less. Béjà et al in 2000 observed transient flash-induced absorption changes in PR. Their results showed that transient depletion occurred near the absorption maximum of the cell pigment which is around 500nm and transient absorption increase was detected at 400 nm and

590 nm, indicating that an intermediate with maximal absorption near 400 nm is produced (M intermediates) (Béjà et al., 2000).

Marine bacteria are a vital component in the ocean's ecosystems, making up a large percentage of the biomass (Whitman et al., 1998). The presence of an archaeal rhodopsin-like light driven proton pump has the potential to greatly affect the functional dynamics of marine ecosystems, as well as changing the perception of biological energy flow in the world's oceans.

Genetic Diversity and Natural Distribution of PR

Genetic Diversity

After the first record of PR in Monterey Bay, many other field investigations into environmental genome libraries and PCR techniques continue to discover new species which express PR. Although over 4000 fragments of PR sequences have been acquired by environmental genomics and PCR analysis from the global ocean metagenomic survey (Rusch et al., 2007), we are still at the beginning of the understanding of the diversity of PR. The broad distribution of these proteins has sparked debate that some photoactive rhodopsin functions may have been present in the last universal common ancestor (Sharma et al., 2006). This spatial distribution has been suggested to reflect their functional properties in relation to environmental opportunities and constraints. Selective environmental pressure or Darwinian evolution may have played a role in preserving phototrophic capabilities (Bielawski et al., 2004, Frigaard et al., 2006).

As described above in the section **Spectral Characteristics of PR** different genetic variants of PR appear to have different spectral tuning to different oceanic habitats. The two different families of PR have different absorption spectra spanning the range from blue (490 nm) to green (540 nm). A study that examined samples from North Sea showed that 97.4% of PR gene fragments were green-light absorbing PR with either a lecuine (66.7%) or a methionine (30.8%) at amino acid position 105 (Riedel et al., 2010). In 2003 Man et al. identified a third PR group in the Mediterranean Sea following comparison of DNA data (Man et al., 2003). The alignment of DNA sequences indicated that the similarities of PR from different geographical and ecological environments are divergent (Fig 3). This suggests a significant effect of natural selection in the evolution and distribution of PR. The data suggests that variants from the Mediterranean shared 88% DNA sequence identity with Monterey Bay and PRs from shallow water North Pacific, and 69% with Antarctic and deep water North Pacific varieties (Sabehi et al., 2003).

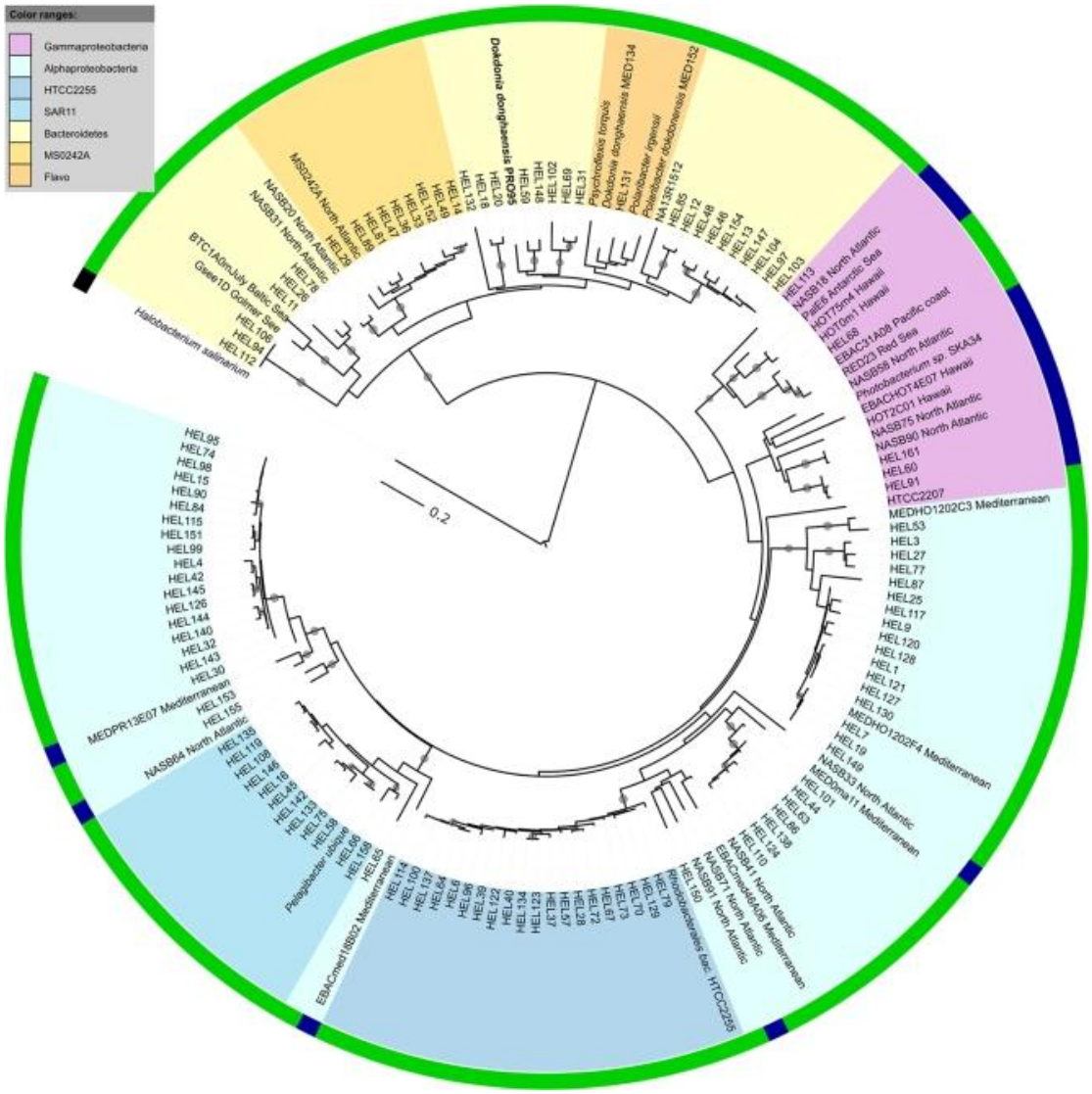


Figure. 3 Neighbour-joining tree based on PR protein sequences isolated from North Sea surface water near the island of Helgoland showing the phylogenetic distribution of PRs in the North Sea. The scale bar shows the number of amino acid substitutions per site. For comparison, PR sequences from isolates and clones from the North Atlantic and Pacific Oceans and the Baltic, Mediterranean, Red, and Arctic Seas have been included. Clone and isolate names are colored according to the relationship of the corresponding PRs to PRs from *Bacteroidetes*, *Alphaproteobacteria*, and *Gammaproteobacteria*. (Riedel et al., 2010)

Distribution of PR

PR-containing bacteria are widespread and abundant in various marine ecosystems, ranging from 13% of the community in the Mediterranean Sea and the Red Sea to 70% in the Sargasso Sea (Sabehi et al., 2005, Venter et al., 2004). PR is one of the most highly expressed proteins in marine bacterial communities (Frias-Lopez et al., 2008). At present, over 30000 PR sequences have been registered in GenBank, and some of them have been found to function as a proton-pump. So far, PCR-based gene surveys have identified a wide variety of PR genes in samples from Monterey Bay, the Hawaiian Ocean Time-series (HOT) ALOHA (A Long-term Oligotrophic Habitat Assessment) station (Central North Pacific), the Antarctic Peninsula (Southern Ocean), the Sargasso Sea, the Mediterranean Sea and the Red Sea (de la Torre et al., 2003) (Table 1).

In a recent study of around 1000 isolates from Antarctic sea ice cores, a total of 52 positive and distinct PR sequences were identified (Koh et al., 2010). It is worth noting that most of the areas described above are oligotrophic environments or under stress conditions. Theoretically since PR-bearing bacteria can use light energy to supply their heterotrophic metabolism under stress conditions, oligotrophic environments and stress condition will therefore be more suitable for PR bearing bacteria to thrive. In 2011 PRs were found in some of the marine protists. PR genes are believed been horizontally transferred from marine bacteria to protists (Sanchez, 2011).

Table 1 Examples of the widespread distribution of PR in the world's oceans

Organism	Classification	Location	Reference
Rhodobacterales sp.	*Alphaproteobacteria*	Hawaiian Fosmid 10m depth	DeLong and Béjà (2010); McCarren and DeLong (2007)
EBO_41B09	*Betaproteobacteria*	Monterey Bay Surface water	McCarren and DeLong (2007)
Porticoccus litoralis	*Gammaproteobacteria*	Yellow Sea, Korea 1m depth	Oh et al. (2010)
Polaribacter sp. Med152	*Flavobacteria*	Northwest Mediterranean Sea Surface water	Gonzàlez et al. (2007)
HOT2C01	*Gammaproteobacteria*	Central North Pacific subtropical gyre	De la Torre et al. (2003)
Vibrio sp. AND4	*Gammaproteobacteria*	Andaman Sea Surface water	Gomez-Consarnau et al. (2010)
HF10_49E08	*Planctomycetes*	Monterey Bay 10m depth	McCarren and DeLong (2007)
TNB-C-3MI	*Flavobacteria*	Antarctic Sea Ice	Koh et al. (2010)
SAR86	*Gammaproteobacteria*	Monterey Bay	Béjà et al. (2000)
HOT2C01	*Gammaproteobacteria*	Central North Pacific subtropical gyre	De la Torre et al. (2003)
Candidatus Pelagibacter ubique HTCC1062	*Alphaproteobacteria*	Coastal Oregon	Giovannoni et al. (2005)
SAR11	*Gammaproteobacteria*	Delaware coastal waters	Lami et al. (2009)
SAR11 GS06 etc.	*Gammaproteobacteria* *Flavobacteria*	Freshwater lake in Canada and Europe	Atamna-Ismaeel et al. (2008)
Flavo-NASB	*Flavobacteria*	North Atlantic Ocean	Campbell et al. (2008)
REDs3a7	*Gammaproteobacteria*	Red Sea	Sabehi et al. (2003)
MedA17r8–15	*Gammaproteobacteria*	Mediterranean	Sabehi et al. (2003)

A study carried out by the Global Ocean Sampling Expedition suggests the green-tuned PR are highly abundant in the Northern Atlantic and in non-marine environments such as fresh water. The blue-tuned PR dominated in the remaining open ocean samples (Rusch et al., 2007). Research over the last 10 years has greatly

increased the understanding of the diversity and wide distribution of PR throughout the world.

Physiological Role of Proteorhodopsin

PR function

When PR was first discovered in bacteria species in the year 2000 it challenged the assumption that chlorophyll *a* is the only important light-harvesting pigment in ocean surface waters and raised the prospect that a considerable amount of solar energy may be harvested by these alternative pigments (Moran and Miller, 2007). Before the discovery of PR, all previously characterized rhodopsins were known as light sensors which have the potential to generate ATP by using light energy via phosphorylation of ADP by the enzyme ATP synthase. PR was first postulated to elevates the proton motive force by the light-activated proton translocation, then the protons re-enter the cell through the ATP synthase complex (Béjà et al., 2001, Béjà et al., 2000). Proton-translocating ATPase – a multi-protein membrane-bound complex that can utilize the proton motive force to generate ATP (Fig 4). To evaluate the hypothesis, marine-derived PR gene were cloned in *E.coli* and showed that it codes for a protein that pumps protons when the recombinant strain is provide with retinal and light (Béjà et al., 2000). The subsequent study conducted by Martinez et al., (2007) showed there was a significant increase of ATP levels in PR positive cells exposed to light. Further experiments with ATP synthase inhibitor revealed that the expression of PR in *E. coli* can result in the phosphorylation of ADP to ATP. This discovery raised the intriguing

possibility that marine bacteria, which were previously thought to obtain their energy exclusively from the oxidation of organic matter, might generate cellular energy directly from light, which would, in turn, permit them to function more efficiently with exposure to the light compared with the dark.

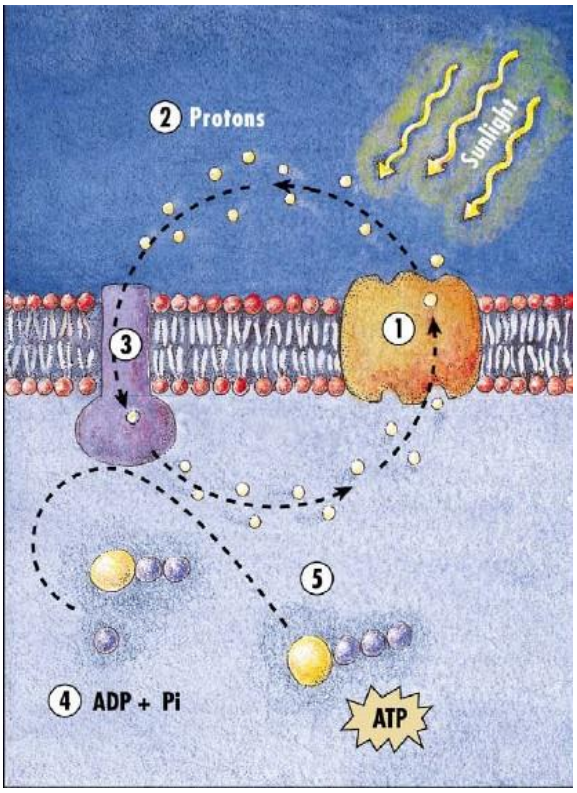


Figure.4 (1) When expose to light, PR uses light energy to translocate protons across the cell membrane. (2) The excess extracellular protons cause an imbalance in the proton motive force which can be used in ATP synthesis, chemical transport processes, or to energetically drive flagellar motility. (3-5) A multi-protein membrane-bound complex; Proton-translocating ATPase, uses the proton motive force to synthesis ATP (DeLong and Béjà, 2010).

To understand the function of PR, much effort has gone into characterizing how they work. Because *E. coli* is a well-studied bacterial model and easy to handle, some studies introduced the PR gene in *E. coli* to understand the physiological function of

PR. The first study carried out by Béjà et al in 2000 characterized the function of PR as a proton pump by detecting pH changes in a PR-containing *E. coli* suspension under illumination. Another study in 2007 suggested that when the cellular respiration is inhibited by depleting oxygen or by the respiratory poison azide, *E. coli* cells expressing PR became light-powered (Walter et al., 2007). Proton motive force can be obtained by exposure of the PR^+ cells to light coinciding with the correct PR absorption spectrum which can turn the flagella motor. Light-illuminated PR^+ cells are less sensitive to azide than PR^- cells, consistent with PR-containing cells possessing energy to maintain cellular viability. Moreover, the data obtained from measuring the Michealis-Menten constant (K_m) of PR indicate that light-driven pumping by PR can fully replace respiration as a cellular energy source in some environmental conditions (Walter et al., 2007). Although the experiments with PR expression in *E. coli* provided some insights into the function of the protein, protein-protein interactions might not occur in a foreign host cell and sometimes under specific conditions, pump protons in the absence of their cognate transducing molecules, which can potentially confound the interpretation of experiments in heterologous systems (Fuhrman et al., 2008). So it is still important to study PR in those natural PR-containing bacteria strains.

Growth Experiments

One of the current hypotheses posits that rhodopsin-based phototrophy could supply a substantial amount of the cellular bioenergetic requirements, with the relative contribution increasing as environments become more oligotrophic (Béjà et al., 2001).

Laboratory studies have examined the growth benefits of light in a few PR-containing marine bacterial cultures. In 2005 PR genes were found in *Candidatus "*Pelagibacter ubique*"* (SAR11 clade), the most abundant bacterium on Earth (Giovannoni et al., 2005). The growth experiments performed on members of the SAR11 clade demonstrated expression of PR in autoclaved seawater, and in the natural environment. However, no variation in expression could be determined between light and dark conditions, therefore the metabolic effect of PRs under natural conditions remains unclear (Giovannoni et al., 2005). Giovannoni et al. postulated that the true effect of PR is most likely more visible when organisms are exposed to limited organic matter, forcing the organisms to produce energy to continue normal cell functioning. The first case of light positively affecting the growth of a marine micro-organism *in-situ* was found in a study on *Dokdonia* sp. strain MED134. Gómez-Consarnau et al (2007) monitored the growth of MED134 in natural seawater under light and dark conditions with low carbon availability. Samples exposed to light reached a maximal abundance of $3x10^5$ cells ml^{-1} while in the samples exposed to darkness growth did not exceed $0.5x10^5$ cells ml^{-1}. When dark samples were transferred into light conditions, growth was initiated with samples reaching a maximum abundance of $2x10^5$ cells ml^{-1}. Beside a positive impact on the growth of MED134 their results suggested that PR was expressed preferentially in the light conditions compare to dark. Subsequent studies also confirmed that light stimulated the growth of other strains. In 2009 Lami et al reported light-dependent growth and PR expression by both *Flavobacteria* and SAR11. SAR11 abundance increased from 1.1×10^4 to 1.9×10^4 over 5 days incubation

under dark-light cycles, while a slight decrease from 1.1×10^4 to 0.3×10^4 was observed under continuous light. In contrast, the abundance of SAR11 decreased sharply from 1.1×10^4 to 0.03×10^4 under continuous darkness. *Flavobacteria* had a similar trend of growth under light, light dark cycle and dark. Despite this strong up-regulation of PR expression under light conditions were observed in both strains. The abundance of PR transcripts in each cell was enhanced up to 120 fold under continuous light and up to 20 fold under dark-light cycles compared to very low levels of PR mRNA while in continuous darkness (Lami et al., 2009). A more recent study in flavobacterium strain MED134 revealed that the genes encoding PR, retinal biosynthetic enzymes and several predicted light sensors were significantly up-regulated in the light when the bacteria grew in carbon-limited media (Kimura et al., 2011).

Unexpectedly some of the bacterial isolates grew as well in the light or dark, at least in the low-nutrient or carbon-limited conditions that were tested (Giovannoni et al., 2005, Riedel et al., 2010, Stingl et al., 2007). Schwalbach et al (2005) examined the growth response over 5-10 days of natural mixed communities of oceanic bacteria which were collected from low-nutrient ocean conditions and incubated in either the light or darkness and determined how light affects bacterial community structure by using Automated Ribosomal Intergenic Spacer Analysis (ARISA). Dark treatment resulted in only minor changes in ARISA profiles: only phototrophs such as cyanobacteria exhibited consistent, sharp declines in dark treatments. This indicates that the majority of non-cyanobcaterial bacteria in the ocean do not depend heavily on light to maintain themselves (Schwalbach et al., 2005).

Some short-term studies of whole seawater communities in light-dark experiments indicate an increasing protein synthesis and respiration level in bacteria cells (Michelou et al., 2007, Church et al., 2004). However it is argued that this result is usually consistent with amino acid uptake by cyanobacteria (Zubkov et al., 2004) and it might not be a direct effect from bacteria that contains PR. Shorter wavelength light and UV photolyse some dissolved organic carbon into more readily useable forms making it more available to heterotrophs (Moran and Miller, 2007). Additionally, bacteriochlorophyll-containing bacteria comprise about 20% of the total marine bacteria in some niches (Cottrell et al., 2006, Schwalbach and Fuhrman, 2005). Many heterotrophic bacteria can use photosynthate that leaks from phytoplankton. Therefore, there are still some processes, other than PR. that can also increase bacterial growth in light compared to darkness.

Light-Dependent Growth in PR-containing Bacteria under Stress Condition

In the natural environment there are a range of stressful conditions that microorganisms must overcome in order to survive and remain viable. The common stress present in the microbial environment such as nutrient availability, change of temperature, salinity and pH level affect cellular energy reserves which force the cells to make physiological changes to either continue functioning, or become dormant until optimal conditions return. Therefore it is more advantageous if physiological adaptation can be applied to cells to keep them functional during the stressful

conditions. The relatively recent discovery of PR has emerged as a mechanism in which marine bacteria may be able to grow under extreme conditions. The current concept is that PR can provide energy to cells under nutrient limiting conditions. A study confirmed that PR can enhance bacterial survival during stress (Gomez-Consarnau et al., 2010) in PR-containing gammaproteobacterial genus *Vibiro* strain AND4 (widespread marine bacteria usually referred as metabolically versatile heterotrophs). The growth experiment with AND4 in rich medium showed no differences in cell yield for both light and dark conditions. After incubating the cells in low concentration of organic or inorganic nutrients for 10 days, bacterial numbers decreased in both light and dark treatment but the light treatment remained 2.5 times higher compared to the dark treatment. In a PR deleted strain, light-dependent starvation survival was abolished, and restored when PR was complemented (Gomez-Consarnau et al., 2010). Their study has conclusively demonstrated for the first time at least one specific physiological role for PR in a native marine bacterium (DeLong and Béjà, 2010). Furthermore this study showed an important new result to the PR research, nevertheless it is not enough to solve the whole puzzle of PR photo-physiology (DeLong and Béjà, 2010). However considering the abundance of PR and possibility of lateral gene transfer (Frigaard et al., 2006), this specific physiological role can potentially influence the adaptive strategy of marine bacterioplankton (DeLong and Béjà, 2010).

The culture and field experiments that have been reported to date do not consistently support the hypothesis of PR benefitting microorganisms under oligotrophic

conditions. Although there is variance in the responses between organisms that express the same gene, it suggests the possibility of multiple PR functions in these bacteria, which could include novel physiological roles (Fuhrman et al., 2008). Therefore it is postulated that in different physiological, ecological, phylogenetic, and genomic contexts, PR activity may benefit cells in a variety of ways, some not directly related to enhanced growth rates or yields (Martinez et al., 2007).

Ecological Significance of PR

As described above, it is estimated that PR-bearing bacterial may account for up to 13%-70% of the marine microbial community. Considering that ocean covers about 71% of the Earth surface, the ocean is the biggest carbon pool on this planet. The previous understanding of photosynthesis in the ocean was mainly established on the basis of photosynthetic pigments (for example chlorophyll). The discovery of PR has brought a novel pathway of marine photosynthesis which is conducted by a single protein. However, rhodopsin-based phototrophy systems do not provide cells with reducing power (NADH etc.) directly. It is currently unclear if carbon dioxide is fixed during marine PR-based phototrophy (Fuhrman et al., 2008).

Nevertheless the discovery of rhodopsin-based phototrophy systems is challenging the conventional concept of biological utilization of light energy in the sea and the assessment of primary production based on photosynthetic pigments either through *in situ* measurements or remote sensing observations (Jiao et al., 2006). Therefore Jiao et al. (2006) proposed a model to demonstrate light utilization and carbon cycling in the

ocean (Fig 5). Pathway I is chlorophyll dependent oxygenic photosynthesis. Pathway II is bacteriochlorophyll-dependent anoxygenic photosynthesis. Pathway III is PR-depending light utilization.

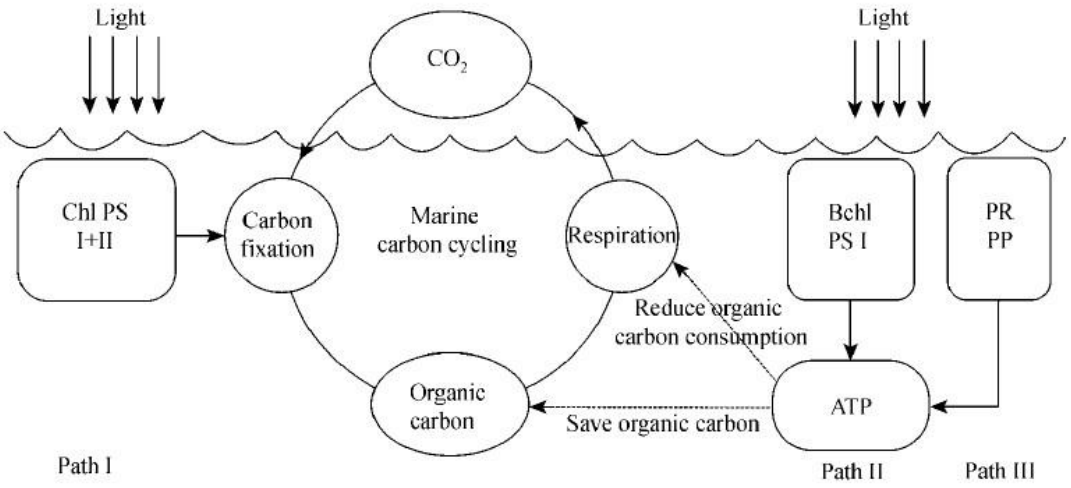


Figure 5: Path I usually used by marine algae, phytoplankton and cyanobactera captures light and transforms it into chemical energy while carbon dioxide is fixed to organic carbon. Path II is used by aerobic and anaerobic anoxygenic photosynthetic bacteria and can transform light into chemical energy but cannot fix carbon due to lack of photosystem II. Path III lacks photosystem I and II but directly captures light and uses a proton pump to produce ATP. Taken from (Jiao et al., 2006).

Currently most of the studies aiming to explain the physiological role of PR invoke arguments of growth efficiency gains during periods of carbon starvation (Gomez-Consarnau et al., 2007, Lami et al., 2009, DeLong and Béjà, 2010, Michelou et al., 2007, Gomez-Consarnau et al., 2010, Kimura et al., 2011). PR did provide a survival advantage in *Vibrio* strain AND4 during starvation (Gomez-Consarnau et al., 2010). But is that all PR can do? Consider the limited carbon in parts of the ocean,

such as the sub-Arctic and Antarctic Pacific (Church et al., 2000, Kirchman, 1990), and also consider the wide distribution of PR. Numerous ocean areas probably lack of elements other than carbon, including nitrogen, iron or have stress conditions other than nutrient stress. The role of PR may be expanded by future studies.

So far, the understanding of the role of bacteria in marine ecosystem is still limited. The current understanding of PR physiological role is still dependent on growth measurements utilizing a limited number of culturable strains and it is possible that the benefit from PR might not always be evident during typical growth conditions. Fundamentally PR is mainly used by some organisms as a survival mechanism when under extreme conditions (Fuhrman et al., 2008). It is possible that bacteria uses PR to obtain minimal level of ATP content to support crucial function or maintain the electrical gradient across the membrane under starvation (Kimura et al., 2011, Fuhrman et al., 2008, Gomez-Consarnau et al., 2010) and maintain survival.

Objectives of this Thesis

A significant proportion of PR gene expression within sea-ice microbial communities corresponds to the presence of metabolically active cells (Koh et al., 2010), which suggests PR may be highly functional within sea ice. However, the complete physiological role of PR-mediated proton transport within bacterial systems remains unclear. Furthermore, it is unknown how PR contributes to survival in sea-ice. Fundamental details on the life strategies, physiological aspects, and genome biology of sea ice bacteria have rarely been addressed. Up till now studies conducted on PR

include growth experiments under controlled conditions, transfer of the PR gene into other well known species and transcriptomics-based analysis. This thesis attempts to combine the studies of genome biology and proteomics research to reach a new level of understanding of the function of this protein. To do this a "whole proteome approach is undertake, examing the abundance of as many proteins as possible within the subject species studied. In addition PR function is investigated within the context of adaptation to a distinct and inherently stressful environment. For this study this environment is the sea-ice ecosystem of Antarctica, which combines salinity and temperature stress. The research relies on the model psychrophilic species *Psychroflexus torquis*, which has many interesting traits relevant to sea-ice inhabitation and also possesses a PR gene.

The following objectives are presented to achieve the proposed aim:

1. to examine and compare the genome of *P. torquis* and *P. gondwanensis*, and measure the expression of predicted genes at protein level in *P. torquis* to achieve a better understanding in how this species may have evolutionarily adapted to the sea-ice environment.
2. to test the hypothesis that PR can provide a growth/survival advantage to bacteria under specific stress conditions within their own econiche.
3. to assess the proteomic profile of *P. torquis* under different salinity and illumination levels to determine how PR links into the species photophysiology and to predict how this would be advantageous within its sea-ice ecosystem.

Chapter 2

Extensive gene acquisition in the extremely psychrophilic bacterial species *Psychroflexus torquis* and the link to sea-ice ecosystem specialism

Abstract

Sea-ice is a highly dynamic but productive environment that includes an array of diverse prokaryotic and eukaryotic psychrophilic taxa distinct from the underlying water column. Because se ice has only been extensive on Earth since the since the early-Eocene it was hypothesized that bacteria highly adapted to the sea-ice inhabitation have traits that have been acquired through horizontal gene transfer (HGT). Here we compare the genomes of the bipolar distributed sea-ice associated bacterial species *Psychroflexus torquis* ATCC 700755^{T}, and its closely related hypersaline lake dwelling sister species *P. gondwanensis* ACAM 44^{T}. We show that HGT-driven take up of genes has occurred to a much greater scale, yielding more functionally diverse outcomes in *P. torquis* ATCC 700755^{T}. Genetic features that can be linked to the psychrophilic and sea-ice specific lifestyle of *P. torquis* include genes for exopolysaccharide (EPS) and polyunsaturated fatty acid biosynthesis (PUFA), numerous specific modes of nutrient acquisition, and proteins putatively associated with ice-binding, light sensing (bacteriophytochromes), and programmed cell death (metacaspases). Proteomic analysis shows that several genes associated with these traits are highly translated, especially those involved with EPS and PUFA production.

Since most of the genes relating to the ability of *P. torquis*' to dwell in sea-ice ecosystems occur on genomic islands absent in closely related *P. gondwanensis* the adaptation to the sea-ice environment seems mainly driven by HGT. The genomic islands are also pseudogene, IS element and addiction module-rich suggesting gene acquisition is being followed by a process of reduction potentially indicative of evolving ecosystem specialism.

Introduction

Sea ice is a major feature of the surface of high-latitude oceans. It is relatively biologically productive due to extensive blooms of sea-ice algae, embedded in the ice floes as a band of growth or associated with the basal section of the ice floe that contacts the underlying seawater. Sea-ice algae and their epiphytic bacteria form the foundation of an active microbial loop comprised of taxa distinct from the underlying seawater (Bowman et al. 1997; Brown and Bowman 2001; Brinkmeyer et al. 2003; Bowman et al. 2012). Tightly coupled to algal-driven primary production, sympagic bacterial populations increase one to two weeks after the phytoplankton bloom peaks in late summer and become increasingly dominant when solar irradiance levels decline and algae subsequently senesce and/or become dormant (Kottmeier et al. 1987; Kottmeier and Sullivan 1987; Grossman and Dieckmann 1994; McMinn and Martin 2013). As sea ice forms, brine is ejected from the ice crystal matrix and collects in cracks and channels (referred to as brine channels) making up 5–20% of the sea ice volume. Though a very cold and saline environment sea-ice microbial communities

(SIMCO) thrive in sea ice brines at temperatures of -10 °C and at salinities three or more times the concentration of seawater (Thomas and Dieckmann 2002).

The extent of most sea ice changes by more than an order of magnitude between summer and winter. Long-term, stable ice tends to only occur connected to land at high latitudes. Recent trends in Arctic Ocean sea ice decline and a simultaneous increase in Antarctic sea ice suggests that climate change may have an impact on sea ice associated taxa, though the extent of this impact is so far difficult to predict (Berge et al. 2012). Based on detection of sea-ice diatom fossils and geological signatures indicating iceberg rafting, ice formation in high latitude oceans has been extensive since the mid-late Eocene about 35–47 million years ago (Mya). However, multi-year ice that would act as a more stable sympagic platform than seasonal sea ice may not have appeared until as late as the Pliocene or Pleistocene (2.5–3 Mya) with the advent of sustained polar glaciation (Polyak 2010). In that time, microbial life associated with sea ice may have had the opportunity to specialise. Currently, our knowledge of the diversity of microbial sea-ice communities and their obligate sympagy remains limited (Bluhm et al. 2011; Poulin et al. 2011).

The genome sequence of the psychrophilic marine species *Colwellia psychrerythraea* (strain 34H) provided the first genome-based perspective on the traits that allow not only for psychrophilic growth, but also the possible means to grow and persist in sympagic ecosystems (Methé et al. 2005). The main traits examined included amino acid composition of proteins and their relation to tertiary structure, secreted and non-secreted cold-active enzymes, omega-3 polyunsaturated fatty acids (PUFA),

compatible solute synthesis, and secreted exopolysaccharides. Ice-active proteins that act to modify ice crystal structure have also been studied (Raymond et al. 2007; Bayer-Giraldi et al. 2010). The important sea ice-dwelling diatom *Fragilariopsis cylindrus* has several ice-active proteins orthologous to generally uncharacterised proteins in cold-adapted bacteria, suggesting that inter-domain horizontal gene transfer (HGT) of these proteins may have occurred (Bayer-Giraldi et al. 2010). This raises the question of whether other traits allowing inhabitation and successful competition in sea ice have also been acquired by HGT processes.

In this study we investigated the genomic properties of the extremely psychrophilic bacterial species *Psychroflexus torquis*, an unusual member of the family *Flavobacteriaceae* (phylum *Bacteroidetes*) that has several traits linked to sea-ice inhabitation and dependence on algae via epiphytism. *P. torquis* was originally isolated from algal assemblages in Antarctic multi-year sea ice. It differs from all other related species, including its closest relative *P. gondwanensis*, in being filamentous at an early stage of growth, extremely psychrophilic, able to synthesize omega-3 and omega-6 polyunsaturated fatty acids, and prolifically secreting soluble extracellular polysaccharide substances (EPS) (Bowman et al. 1998). The species, though chemoorganotrophic, can also harness energy from light via PR-driven proton pumping, a feature enhanced under osmotic stress (Feng et al. 2013). Interestingly proteorhodopsin (PR) gene is located adjacent to a genomic island which is absent in *P. gondwanensis*. The genus *Psychroflexus* is found within moderately hypersaline ecosystems across the world; however, the combined traits of psychrophily and PUFA

synthesis in *P. torquis* make this species stand out among other members of phylum *Bacteroidetes*. To explore these and other ecologically relevant genomic aspects of *P. torquis* that may provide insight into the relatively recent evolution of psychrophily, we compared the genome of the type strain ATCC 700755^T to that of its closest relative *P. gondwanensis* ACAM 44^T. To better discern important genes we also performed proteomics on ATCC 700755^T grown under a range of conditions. In particular, we searched for mobile genetic elements and pseudogenes as evidence for HGT and its possible role in sea-ice ecosystem specialism.

Materials and Methods

Genome Sequence Determination

P. torquis ATCC 700755^T (T = type strain) and *P. gondwanensis* ACAM 44^T (ATCC 51278^T) were cultivated on modified marine agar (0.5% w/v proteose peptone, 0.2% w/v yeast extract, 1.5% w/v agar, 3.5% w/v sea salts) at 4 °C and 25 °C, respectively. High molecular weight DNA was extracted and purified from biomass using the Marmur method. DNA was sequenced using the 454 GS-FLX / Plus (454 Life Sciences, Branford, Conn., USA) platform following the manufacturer's *de novo* sequencing protocol. For ATCC 700755^T and ACAM 44^T 146.5 and 149.1 Mbp of sequence data (430-440 bp average length) was assembled using Newbler v. 2.6 (454 Life Sciences). The Sanger sequence draft already available for ATCC 700755^T (generated through the Gordon and Betty Moore Foundation Genome Sequencing Project at the J. Craig Venter Institute) was compared to the pyrosequenced contigs

using Artemis (Carver et al. 2012) with the number of contigs reduced from 39 to 9. Gaps between contigs were closed using PCR analysis. Apparent mis-assemblies and regions with sequence inconsistencies were corrected in Artemis after PCR and sequencing confirmation of the regions.

Post-Sequence Analysis

Gene annotation for the complete ATCC 700755^T sequence was carried out in Artemis and also compared with annotations generated via the Prodigal server (Hyatt et al. 2010) and Glimmer v. 3.02 (Delcher et al. 1999). Transfer RNAs were predicted using tRNAscan-SE (Lowe and Eddy 1997). Predicted CDSs were compared against the NCBI database. Annotation utilised the NCBI Prokaryotic Genomes Automatic Annotation Pipeline. Pseudogenes in both genomes were defined as described by Lerat and Ochman (2005) based on comparisons with highly similar sequences in ATCC 700755^T, ACAM 44^T and with highly similar orthologs in related taxa. The presence of protein signal peptides and transmembrane helices were predicted using the SignalP 4.1 (Petersen et al. 2011) and THMM v. 2.0 servers (Centre for Biological Sequence Analysis, Technical University of Denmark), respectively. The genome of ATCC 700755^T was visualised using DNAPlotter (Carver et al. 2009). CRISPR palindromic repeats were detected using CRT (Bland et al. 2007).

Protein Extraction and Post-Treatment

P. torquis ATCC 700755^{T} was grown on modified marine agar at different sea salt salinities and light intensities at 4 °C in order to examine the broadest possible set of proteins produced by *P. torquis*. The different salinities were achieved by adding 17.5, 35, 52.5 and 70 g.L^{-1} of sea salt (Red Sea) to the marine agar and three levels of illumination (0, 3-4, 20-30 μmol photon s^{-1} m^{-2}) were used. Cells were lysed in 1 ml 50 mM Tris-HCl buffer (pH 7.0) by sonication in an ice bath, with 10 s of sonication with a 10 s wait period, cycled 15 times until the opaque cell suspension became translucent. The suspensions were then centrifuged at 16,000 × *g* for 25 min at 4 °C. The supernatant protein was precipitated using trichloroacetic acid and the protein pellets were then treated with 0.2 M NaOH to improve subsequent solubilisation (Nandakumar et al. 2003). The pellets were then solubilised using 100 μl 50 mM Tris-HCl buffer (pH 7.0) and protein concentration determined using the Bradford assay (Bio-Rad). Volumes of samples containing 50 μg of soluble protein extract were then reduced in a solution of 50 mM dithiothreitol, 100 mM ammonium bicarbonate for 1 h at room temperature. The samples were then alkylated with 200 mM iodoacetamide in 100 mM ammonium bicarbonate for 1h at room temperature. After alkylation, the reduction of proteins was repeated and they were digested in a buffer (50 mM ammonium bicarbonate, 1 mM calcium chloride) that contained sequencing grade modified trypsin (Promega), at a sample protein to trypsin ratio of 25:1, at 37 °C with gentle shaking overnight. Digestion was stopped by acidification with 10 μl of 10% (vol/vol) formic acid. The samples were then centrifuged for 5 min at

14,000 × *g* to remove any insoluble material, and an aliquot (100-200 µl) of peptides was transferred to HPLC vials for mass spectrometric analysis. Samples were prepared with three biological replicates and for each biological replicate, two technical replicates were performed.

NanoLC-Orbitrap Tandem Mass Spectrometry

The separation of peptides utilized a Surveyor Plus HPLC system fitted in line with an LTQ-Orbitrap XLmass spectrometer (ThermoFisher Scientific). Aliquots of peptide samples were loaded at 0.05 mL/min onto a C_{18} capillary trapping column (Peptide CapTrap, Michrom BioResources) controlled by an Alliance 2690 separations module (Waters). Peptides were then separated on an analytical HPLC column packed with 5 µm C18 media (PicoFrit Column, 15 µm i.d. pulled tip, 10 cm, New Objective) using four linear gradient segments controlled using a Surveyor MS Pump Plus (ThermoFisher Scientific) at 200 nl/min The solvent series included initially 5% acetonitrile in 0.2% formic acid (solvent A) shifting up to 90% acetonitrile in 0.2% formic acid (solvent B). The 4 stage process where solvent B gradually replaced solvent A comprised: 0–10% solvent B over 7.5 min; 10–25% solvent B over 50 min, 25–55% B over 20 min and then 55–100% solvent B over 5 min. This process was followed by re-equilibration of the column with solvent A for 15 min. The LTQ-Orbitrap was fitted with a dynamic nanoelectrospray ion source (Proxeon) containing a 30 µm inner diameter uncoated silica emitter (New Objective). The LTQ-Orbitrap XL was controlled using Xcalibur 2.0 software (ThermoFisher

Scientific) and operated in data-dependent acquisition mode whereby the survey scan was acquired in the Orbitrap with a resolving power set to 60,000 (at 400 m/z). MS/MS spectra were concurrently acquired in the LTQ mass analyser on the seven most intense ions from the FT survey scan. Charge state filtering, where unassigned and singly charged precursor ions were not selected for fragmentation, and dynamic exclusion (repeat count, 1; repeat duration, 30 s; exclusion list size, 500), was used. Fragmentation conditions in the LTQ were 35% normalized collision energy, activation *q* of 0.25, 30 ms activation time, and minimum ion selection intensity of 500 counts.

Protein Identification and Bioinformatics

The acquired MS/MS data were converted to .mzXML pleak list files from .RAW files using the msConvert command in Proteowizard. The MS/MS data was searched against the proteome of *P. torquis* ATCC 700755^{T} (NCBI accession code CP003879) employing X!Tandem running in the open source Computational Proteomics Analysis System environment (Rauch et al. 2006). For searches parent ion tolerance of 10 ppm and fragment ion mass tolerances of 0.5 Da were used, and enzyme cleavage was set to trypsin, allowing for a maximum of two missed cleavages. Amino acid residue alterations were also accounted for including: S-carboxamidomethylation of cysteine residues specified as a fixed modification; and cyclisation of N-terminal glutamine to pyroglutamic acid, deamidation of asparagine, hydroxylation of proline and oxidation of methionine specified as variable modifications. Proteins identifications were

assessed by the Peptide Prophet and Protein Prophet algorithms (Nesvizhskii et al. 2003). Protein identifications were filtered by assigning a Protein Prophet probability > 0.95 and by accepting only peptides with charge states of +1, +2, and +3 because charge states ≥ 4 were not a good fit to Peptide Prophet models. This filtration constrains the protein false discovery rate to < 1%. Spectral counting was used to determine relative protein abundance taking into account the number of amino acid residues (Liu et al 2004). In order to acquire the maximum coverage of proteins, spectra obtained from all samples were pooled together in this study. Filtration removed 6.4% peptides with a total of 743361 peptide spectra matched to 2598 protein identifications with 1936 proteins possessing two or more unique peptides.

Results and Discussion

A Highly Conserved Core Genome is Shared by P. torquis and P. gondwanensis

Based on a meta-analysis of the latest metagenome and NCBI database records, *P. torquis* is restricted to sea ice or seawater around ice floes and has a bipolar distribution, occurring in both the Arctic and the Antarctic (Figure 1). *P. gondwanensis*, the closest cultivated relative of *P. torquis*, so far is only known to reside in Antarctic hypersaline lakes, where it can be a dominant member of the microbial community (Yau et al. 2013). An equidistant lineage detected in the ice cover of Lake Vida, Antarctica (Mosier et al. 2007), and in salinated Yellow River delta soil from China, is so far uncultured but suggests a broader distribution of

closely related *Psychroflexus* genotypes. The 4.32 Mbp genome of the *P. torquis* type strain ATCC 700755^{T} was obtained in two stages, first as an 8 × coverage Sanger-sequenced draft and then closed in this study by a combination of pyrosequencing, gap filling and subsequent PCR-based checks of potentially misassembled regions. The 3.32 Mbp draft genome of *P. gondwanensis* type strain ACAM 44^{T} (NCBI accession code APLF00000000) was obtained at a coverage level of 44-fold. Details for the genomes are summarized in Table 1.

The 16S rRNA gene sequence similarity between *P. torquis* and *P. gondwanensis* strains and clones averages 99.0%, higher than the empirical 98.5% similarity cut-off that has often been used as preliminary evidence for defining distinct prokaryotic species (Stackebrandt et al. 2006). However, a full genome comparison between the two genomes yielded an average nucleotide similarity (ANI) (Goris et al. 2007) of only 68%, consistent with an overall low DNA:DNA hybridization level of < 20% (Bowman et al. 1998). Despite the differences in size and ANI score, both strains share 2308 genes out of an effective pan-genome of 4225 protein-coding genes. The genes that are in common have a high mean nucleotide similarity (92 ± 4%) and extensive stretches of synteny (see Appendix B, Table B.1).

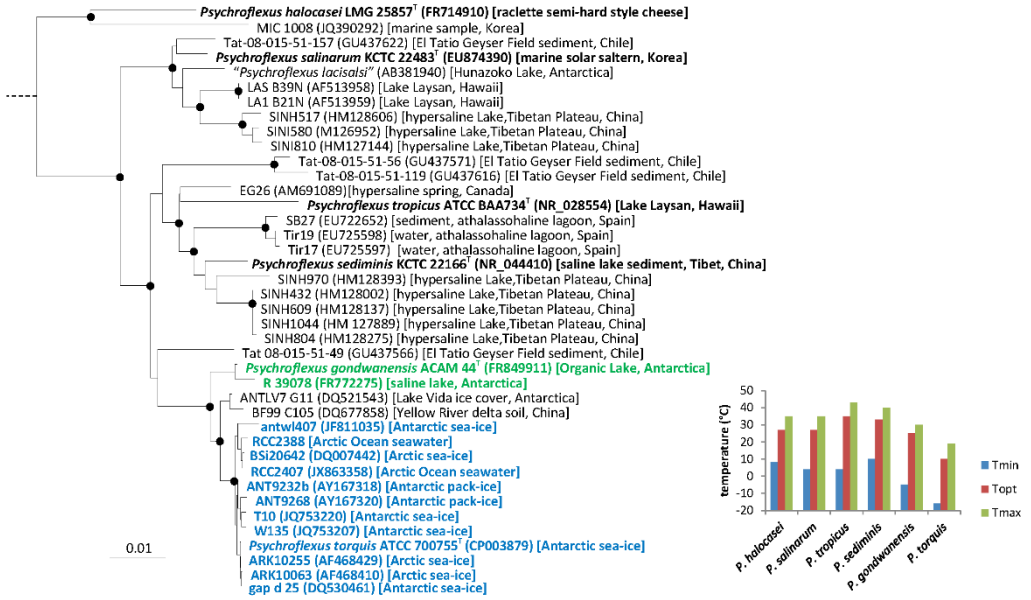


Figure 1. 16S rRNA gene-based phylogenetic tree of the genus *Psychroflexus* (family *Flavobacteriaceae*, phylum *Bacteroidetes*). The tree was constructed with complete or near complete sequences, aligned with CLUSTALW, clustered with the Maximum Likelihood algorithm, and created using Neighbour-Joining. Black circles indicate that bootstrapping support for nodes is >80%. The outgroup used was *Capnocytophaga ochracea*. The bar indicates relative sequence distance. NCBI accession codes are given in parentheses followed by location of isolation. For type strains of described species cardinal temperatures for growth rate are indicated in the inset bar graph including values for temperature: minimum (T_{min}), optimal (T_{opt}) and maximum (T_{max}) temperature values. Values were determined in liquid media, with the minimum temperature a theoretical value calculated from the square root model of Ratkowsky and colleagues (1983).

Table 1. Genome data for *Psychroflexus torquis* ATCC 700755^T and *P. gondwanensis* ACAM 44^T

Species	***Psychroflexus torquis***	***Psychroflexus gondwanensis***[a]
Strain	ATCC 700755^T (= ACAM 623^T)	ACAM 44^T (=ATCC 52178^T, DSM 5423^T)
Taxonomic hierarchy	*Flavobacteriaceae*, *Flavobacteriales*, *Flavobacteria*, *Bacteroidetes*	
Genome status	Finished	Non-contiguous draft
Platform	Sanger (JCVI), 454 GX FLX	454 GS FLX
Coverage	8 ×, 33 ×	44 × (62 contigs)
Number of replicons	1	1
Extrachromosomal elements	0	0
GenBank ID	CP003879, NC_018721	APLF00000000
Genome size (bp)	4,321,832	3,325,075
DNA coding region (bp)	3,503,343 (81.06%)	2,852,146 (85.77%)
DNA G+C content (bp)	34.51% (34.9/33.5)	35.72%[b] (36.0/34.4)
Total genes	3951	3007
RNA genes	45 (1.14%)	47 (1.61%)
rRNA operons	3	3
Protein coding genes	3526	2912
Pseudogenes	379 (9.57%)	48 (1.60%)
Genes with a predicted function	2278 (64.58%)[a]	2006 (68.32%)
Proteins with signal peptides or POR secretion system sorting domains	380 (10.78%)	299 (10.18%)
Proteins with transmembrane domains	761 (21.58%)	658 (22.41%)

[a]Estimates based on available non-contiguous sequence data.

A dissection of genes by functional classification demonstrates that the conserved gene overlap includes virtually all RNA-coding genes and genes involved in fundamental cellular processes, including DNA-related processes, protein translation, ribosome structure and biogenesis, tRNA processing, protein secretion, cytokinesis, nucleic acid transport/metabolism, cofactor transport/metabolism, and gliding motility (Figure 2). Many of these genes have similarly co-located and syntenic arrangements amongst related taxa within the family *Flavobacteriaceae*, demonstrating an apparent ancestral nature. Other functional classes more moderately conserved between the two

species include metabolism, protein modification, folding and turnover, and defence/detoxification systems. Overall this pattern of gene sharing is consistent with both *P. torquis* and *P. gondwanensis* being inhabitants of cold saline ecosystems and having broadly similar though not identical morphological and metabolic phenotypes (Bowman et al. 1998).

Evidence of Extensive HGT in the Species-Dependent Section of the P. torquis ATCC 700755^T genome

Large tracts of the ATCC 700755^T and ACAM 44^T genomes share no significant nucleotide similarity with 1265 and 653 strain distinct genes respectively. The structural and functional categorisation of the species-dependent genome regions reveal that in general, *P. torquis* ATCC 700755^T has a wide variety of functional genes in this set compared to *P. gondwanensis* ACAM 44^T (Figure 2, Table B.1, Table B.2). When directly compared against the ACAM 44^T genome (Figure 3) these regions occur in distinct genomic islands (GI) of which 44 could be defined (size range 3.7 to 153.2 kb, Table B.1). Fourteen of these GIs are flanked by tRNA genes and the majority are characterised by low gene density (average 74.5%) and are relatively rich in pseudogene, insertional element and/or addiction modules. The presence of tRNA genes flanking GIs is consistent with tRNA being a well known hot spot for site-specific recombination (Ou et al. 2006). One tRNA gene (tRNA-Met between P700755_00961/00963) has been disrupted in ATCC 700755^T by an integrase suggesting other tRNAs may have also been lost by GI insertions. An example of such

may include a tRNA-Ala-GGC present as a pair in ACAM 44^{T} but only singly in ATCC 700755^{T} (between P700755_01366/01368), located directly adjacent to a 153.2 kb GI (GI no. 17; Table B.1).

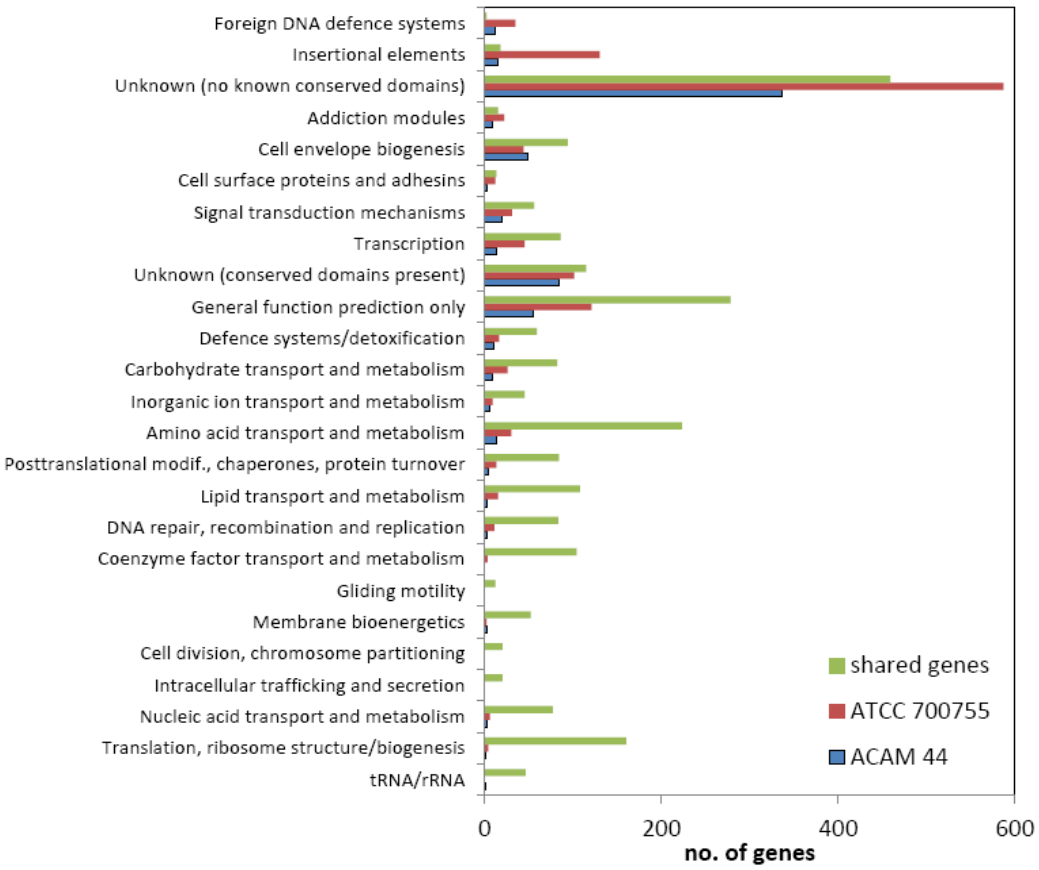


Figure 2. Numbers of genes shared or strain dependent in *P. torquis* ATCC 700755^{T} and *P. gondwanensis* ACAM 44^{T} organised by functional class.

It is suggested that the variation in G+C content can be important evidence for HGT in regions of a genome (Ochman et al. 2000). Thus we plotted the overall GC% of *P. torquis* using DNAMAN. The overall GC% of *P. torquis* is 34.5%, and we were able to locate several regions that deviated from this mean ranging between 26 to 50% with 15 GIs are located within these areas (see Appendix B, Figure B.1) suggestive of an external origin and ancestry of these genomic regions. Furthermore, comparisons

of GI-associated genes with the NCBI database were made and the highest match noted on the basis of amino acid similarity. A high proportion of genes (27%) did not match anything on the NCBI database, while 45% of matches occurred with proteins from members of the family *Flavobacteriaceae* that mainly occur in marine, especially polar, ecosystems (i.e. *Polaribacter* spp., *Cellulophaga algicola*, *Gillisia* spp.). The remainder of matches were with a wide variety of bacteria within the phylum *Bacteroidetes* and in other phyla (Table B.1). A similar pattern was observed for ACAM 44^{T}, though more muted in terms of the sheer scale of gene acquisition. Sea-ice has been proposed as a hotspot for genetic recombination due to its high density of bacteriophage, a result of the concentration of brines during sea ice formation (Wells and Deming 2006). Despite no genes of definitive phage origin being found on the ATCC 700755^{T} genome, extensive phage defence systems were detected including four large regions containing 7 to 28 CRISPR repeats with the main concentration located immediately after a cluster of CRISPR genes (P700755_00291 to 00295). In addition, six restriction-modification systems and a variety of proteins in the phage infection resistance family were detected (Table B.1). Approximately 4% of genes were IS elements belonging to 40 families of retron-associated reverse transcriptase, integrase and tranposase proteins. Ten families of addiction (toxin/antitoxin) modules were also present. These gene types and their distribution suggest that mobile genetic elements and potentially phages were involved in building GI-associated genomic content. Overall, the data suggests that there has been a high level of gene acquisition in *P. torquis* ATCC 700755^{T}. Based on

the collective nature of the GIs as explained above, we hypothesize that HGT processes drove this acquisition. Since direct evidence such as prophage and conjugative systems are no longer evident on the genome it would be ideal to collect data from a wider range of strains as well as other sea-ice-derived bacteria to assess the degree of HGT and whether there is a pool of shared genetic homologs within SIMCO.

Modern Evolution of Psychrophily in P. torquis ATCC 700755^T and its Relation to Genome Sequence-Derived Criteria.

Psychrophilic prokaryotes unlike thermophilic taxa are neither phylogenetically deep-branching nor tend to cluster together. In terms of phylogenetic relationships, psychrophilic strains tend to occur almost without exception in genera with higher temperature adapted relatives and are usually located at the tips of branches. This aspect of psychrophilic prokaryote phylogeny was first noted by Franzmann (Fanzmann 1996). Though it is conceivable that various cold-adapted microbes evolved during earlier cold times on Earth, for example during the purported Cryogenian period (MacDonald et al. 2010) and Snowball Earth periods (Hoffman and Scharg 2002), psychrophily amongst prokaryotes is a modern phenomenon based on the temperature history of Earth and underlying prokaryotic evolution (Schwartzman and Lineweaver 2005). As seen in Figure 1 the deeper branching *Psychroflexus* species are classic mesophiles and derive from widespread locations, such as geyser field soils, marine sites, saline lakes, and even cheese. The tip position

of *P. torquis* in the *Psychroflexus* phylogenetic tree (Figure 1) and its apparent restriction to polar marine locations, specifically multi-year sea ice, fits the concept of a modern evolution of psychrophily. The cardinal growth temperatures of *P. gondwanensis* and *P. torquis* differ by more than 10 °C and this value is even greater for other *Psychroflexus* spp (Figure 1).

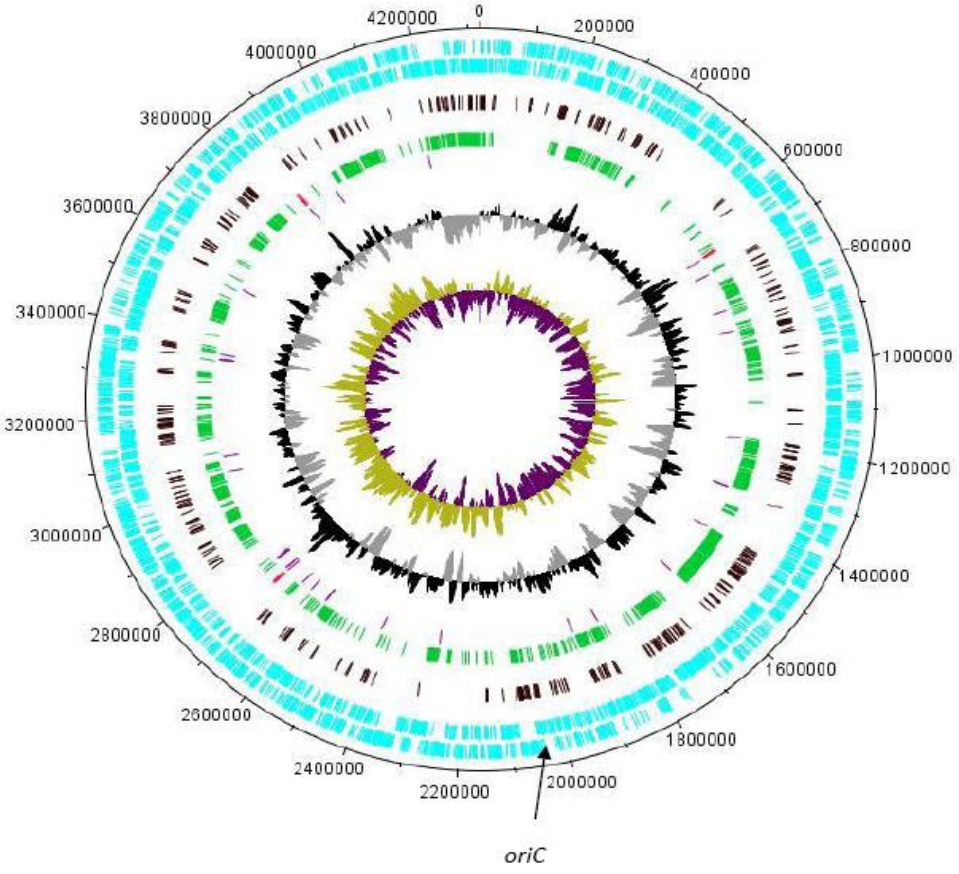


Figure 3. Genome map of *P. torquis* ATCC 7000755^T drawn using DNAPLOTTER (Carver et al. 2009). The first and second rings (blue) show gene annotations for the sense and antisense strands, respectively. The third ring (brown) shows the position of annotated pseudogenes. The fourth ring (green) shows genes that occur in *P. torquis* ATCC 700755^T but not *P. gondwanensis* ACAM 44^T. On this ring (in red) are the positions of the three rRNA operons, which are located in syntenic regions of the ACAM 44^T genome. The fifth ring (purple) shows the position of tRNA coding regions. The sixth ring demonstrates GC-bias (black – positive, gray – negative) across the genome calculated from 1000 bp segments. The inner ring shows GC-skew with the leading strand commencing at the predicted *oriC*.

Typically amino acid composition in general terms is correlated to growth temperature preference, but in the case of the strains studied here overall amino acid content did not vary significantly. Specific amino acids (abbreviated as IVYWREL) have been found to be strong determinants of thermostability (Zeldovich et al. 2007); however, predictions based on the levels of these amino acids for ATCC 700755^{T} and ACAM 44^{T} overestimated their optimal growth temperatures as 29 °C and 34 °C respectively. This result is not surprising in that predictions of thermoadaptation are relatively insensitive at the mesophilic/psychrophilic end of the biokinetic spectrum, suggesting an inherent limit to thermostability of proteins and thus growth (Corkrey et al. 2012). Codon usage analysis was also performed to determine whether any significant trends occurred between the strains based on highly translated gene products. This analysis was done by comparing the top 10 percentile of the most abundant proteins of ATCC 700755^{T} that had highly similar orthologs in ACAM 44^{T} (n = 306) determined from the protein spectral count dataset (Table B.1). Analysis of the expected codon adaptation index (Puigbò et al. 2007) revealed significant bias mainly due to differences in synonymous 3rd-base positions being more AT-rich (average % GC3: 27 vs 35). Overall, this suggests that amino acid content and codon usage criteria can discriminate between the two species examined; however, the trends may not necessarily relate only to psychrophily. A similar situation was found with the extreme psychrophile *Psychromonas ingrahamii*, which did not exhibit unusual codon usage trends (Riley et al. 2008) nor was its level of IVYWREL amino acid content informative. Another explanation for temperature sensitivity is that *P.*

torquis is inherently unstable at mesophilic temperatures and that, because proteins do not seem obviously involved, it is possible another part of the cell is the temperature "Achilles Heel". A logical candidate is the cell membrane of ATCC 700755^T which is quite different from that of ACAM 44^T as it is rich in PUFA and anteiso-branched fatty acids (Bowman et al. 1998). This combination creates membrane fluidity compatible with very low temperature even though it has the disadvantage in being potentially more thermolabile (D'Aoust and Kushner 1971).

Evidence of Significant Gene Decay in the P. torquis ATCC 700755^T Genome.

The genome of ATCC 700755^T has an unusually large number of pseudogenes (n = 379) making up 9.5% of the total number of protein coding genes. This percentage represents a conservative 8-fold increase over that of the ACAM 44^T genome for which pseudogene numbers have likely been overestimated due to its draft status. The appearance of pseudogenes is believed to be associated with recent mutational processes because they seem to be rapidly deleted from genomes (Kuo and Ochman 2010). In ATCC 700755^T the pseudogenes primarily occurred as truncated fragments or segmented ORFs due to one or more nonsense mutations and/or indels, while more rarely pseudogenes were derived from direct transposon or retron disruptions. In the case of the more overtly truncated ORFs most have been affected by subsequent frame shifts and partial deletion because the coding region remnants were usually less than 40% of the full length version with the remaining degenerate region sometimes

still adjacent to the pseudogene. Indeed numerous examples of fragmentary relics lacking both stop and start codons were detectable in intergenic regions. Insertional (IN) element types are very diverse in *P. torquis* ATCC 700755^{T} with a high proportion also being pseudogenes (Table B.3). Given that many IS elements and other mobile genetic elements are concentrated in GIs, insertion and recombination appears to have shaped the genome of ATCC 700755^{T} extensively. Such high proportions of pseudogenes essentially indicate a process of both gene decay and adaptation, as has been observed in bacteria transiting to a lifestyle of obligate parasitism or symbiotism (Burke and Moran 2011).

The large number of pseudogenes in ATCC 700755^{T} relative to ACAM 44^{T} suggest that HGT gene acquisition may have been both advantageous and deleterious. Unnecessary genes, copies of genes involved in the HGT processes themselves, as well as those accidentally disrupted via integration events have been and are being progressively deleted. This process may be due to selection against pseudogenes themselves or selection of processes that actually remove them from genomes (Kuo and Ochman 2010). Other strains will need to be examined to determine if this pseudogene decay is consistent at a species level and if the “burden” of pseudogenes correlates to fitness. The predicted location of the point of origin of replication (*oriC*), detected using DoriC v. 5.0 (Gao and Zhang 2012) is of interest in this regard. Surprisingly, two putative *oriC* sites were found, both located near each other within a 31.2 kb GI (no. 22) immediately adjacent to retron-type reverse transcriptases (between P700755_01930/01931; P700755_01949/ 01953), the latter of which is a

pseudogene.

Pseudogenes are generally assumed not to be expressed or translated although exceptions have been detected (Rusk 2011). Based on our proteomic spectral count data the vast majority of pseudogenes were not detected after filtration (Table B.3). However, some pseudogenes have substantial number of collated spectral counts that had high confidence of identification. In all cases these peptides occurred on IS elements that occurred multiple times in the genome either as full length genes or truncated pseudogene versions. It is possible that truncated protein products are still translated but in general appear to have low cellular abundance based on the spectral count data.

Functional Prioritization Suggested by Abundance of Protein Products in P. torquis ATCC 700755^T Genome

We assessed to what degree the genes of ATCC 700755^T are translated using proteomic analysis. It is assumed that the more abundant proteins are inherently fundamental to the system biology of ATCC 700755^T. Above an arbitrary threshold set at 0.005% of the filtered and normalized total spectral count, proteins were regarded as being abundant. At this threshold, spectral counts for multiple peptides were observed in most sample replicates, thus suggesting sustained production under the range of growth conditions tested. The proteomic dataset is however logistically limited due to huge dynamic ranges with a natural bias against low abundance and transmembrane proteins (Borg et al. 2011). Many proteins belonging to the

overlapping proteomes of ATCC 700755^T and ACAM 44^T were readily detected, as expected, as were high proportions of functionally conserved, mainly cytosolic proteins (Fig. 4). Amongst these the least detected proteins were associated with DNA-related processes, including repair and recombination (only 20% of proteins observed), while the highest proportion (63%) was associated with electron transport. This difference may indicate the prioritization of proteins in terms of cellular processes where certain functional proteins such as DNA repair are only up-regulated to high levels when needed. Other proteins, such as those involved in central metabolism and energy production, are constantly required by the organism. At the other extreme, some functional groups of proteins from the ATCC 700755^T genome were rarely detected, including addiction modules, foreign defence proteins, insertional elements, and a high proportion of proteins with seemingly no function at all based on their lack of conserved domains. The weak translation of these latter proteins suggests that their presence could be largely strain dependent (Ou et al. 2005).

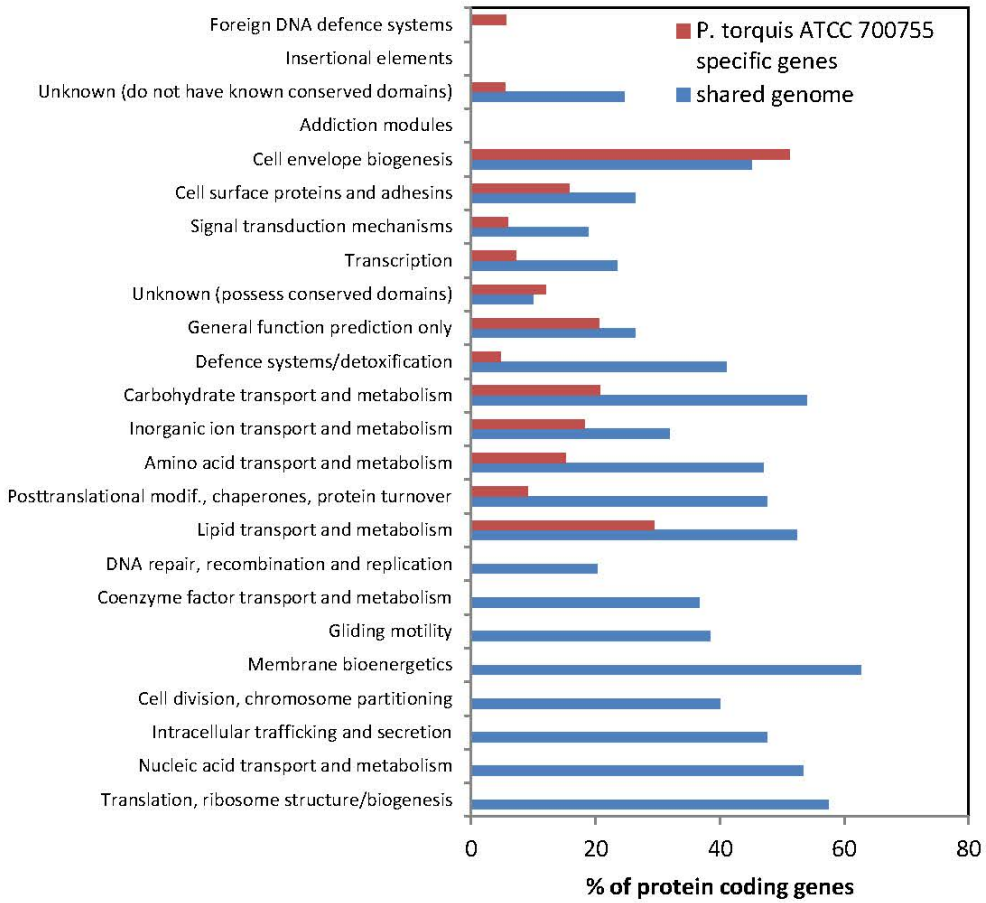


Figure 4. Distribution of most abundant proteins of *P. torquis* ATCC 700755^{T} and whether they are strain-specific or shared with the genomeof *P. gondwanensis* ACAM 44^{T}, grouped by functional classes. Proteomic data was pooled from all treatment samples, an abundant protein was defined as 0.005% of the total spectral counts; each was detected in most replicates and represented by multiple peptide sequences.

Genes that Code Abundant Proteins in P. torquis ATCC 700755 are Potentially Important for Sea-Ice Inhabitation

We assume that species and/or strain-specific features (as summarised in figure 2 in terms of gene distribution) greatly contribute to the inherent uniqueness of bacterial strains at a phenotypic level and subsequently determine their ecological nature.

Features that are strongly expressed likely have important roles in defining this identity. Based on protein abundance analysis, the most prominent products of the unique genome regions of *P. torquis* ATCC 700755^T (Figure 4) are proteins associated with cell envelope biogenesis, cell surface proteins/adhesions, proteins involved in the transport and metabolism of carbohydrates, lipid and inorganic ion transport and metabolism, and those with so far only generalised functions. Many strain-dependent features are likely not well observed due to the inability of laboratory-based growth conditions to adequately capture this information. Given that sea-ice ecosystem conditions are highly complex in terms of resource availability, physicochemical pressures and biological interactions, many other more transiently expressed proteins potentially play important roles. Nevertheless, these experiments provide a preliminary view of the functionality of ATCC 700755^T in the context of the ecosystem to which it is specialised.

A list of proteins found to be relatively abundant (as defined above) and produced by genes only on GIs (Table B.1), summarizes as far as it is possible, the potential unique biology of ATCC 700755^T and provides a suite of relevant functions for its sea-ice ecosystem associations. In particular, two GIs (no. 18 and 43) have a large number of moderately to highly abundant proteins while several GIs include no abundant proteins at all. The selective pressure enforced in a sea-ice ecosystem may lead to the retention of some GIs as stable sections of the genome, while others could be eventually lost.

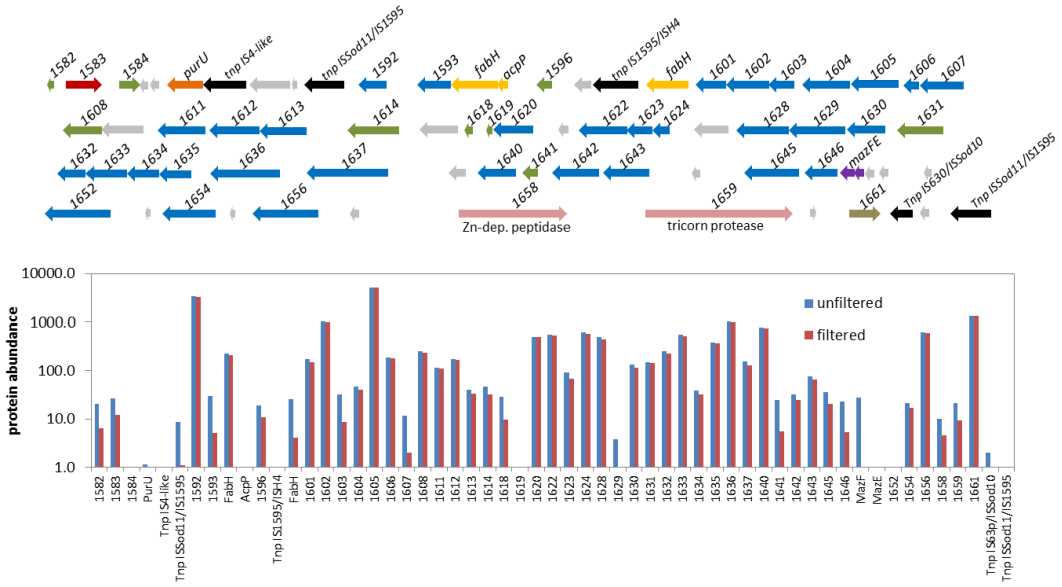


Figure 5. The genomic island region 18 of *P. torquis* ATCC 700755^{T} containing a large EPS biosynthesis cluster that is highly translated and likely involved in the organism's manufacture of complex EPS. Relative abundance of gene products for this region is shown in the lower graph. Gray genes denote pseudogenes. Black genes denote intact transposases. Green genes are hypothetical proteins with conserved domain regions. Blue genes are enzymes involved in polysaccharide biosynthesis including synthesis of modified sugars, glycosylation, polymerases, and flippases; P700755_001654 and 1656 are near identical copies of putative UDP-glucose 6-dehydrogenase genes. Yellow genes are associated with lipid metabolism and include two FabH homologs. A dark red gene denotes a NusG-like transcriptional elongation/antitermination factor. Pink genes separated by large intergenic regions include putative exported proteases. Purple genes include a MazEF family addiction module. Further details are shown in Appendix B data files Table B.1 and B.3.

The EPS production by ATCC 700755^{T}, which includes multiple sulphated, uronic acid-containing, and N-acetylated polysaccharides, is prodigious substantially increasing medium viscosity, with production levels and viscosity increasing with decreasing growth temperature (Mancuso Nichols et al. 2005; Bowman 2008). GI no. 18 includes a 60-kb EPS biosynthesis locus (Figure 5). Though the exact structure of

the EPS and its specific functional benefits remain to be elucidated, EPS production has been found to be a common feature of sea-ice bacteria and a crucial factor for sea-ice inhabitation. It provides cryoprotection (Marx et al. 2009), encourages ice crystal modification, retains liquid brine in brine channels, and enhances recruitment into forming sea ice (Ewert and Deming 2011; Krembs et al. 2011). It also has a role in nutrient acquisition, especially trace elements such as iron and cobalt since anionic EPS can act as powerful ligands (Hassler et al. 2011). The highly translated level of EPS biosynthesis gene products (Figure 5) suggests that EPS may be crucial for the low temperature growth and activity of ATCC 700755^{T} even outside of the sea ice environment. The EPS cluster contains genes that match those of bacterial relatives within and outside of the phylum *Bacteroidetes*. Intermingled with these genes are intact and remnant transposase genes, as well as a MazEF family addiction module and duplicated genes coding putative UDP-glucose 6-dehydrogenases, suggesting recent acquisition via HGT (Figure 5).

The second region of highly translated proteins is located on GI no. 43 and includes a 42-kb cluster of transporter-like proteins mostly of the protein family referred to as acidobacterial duplicated orphan permeases (ADOP) (P700755_03736 to 03759) first observed in the genomes of acidobacteria (Ward et al. 2009). The ADOP proteins include a cluster of 10 paralogs associated with other transporter proteins (Figure 6). The ADOPs all contain a MacB-periplasmic domain, suggesting they could be involved in efflux. The relatively high protein abundance and clustered nature of numerous ADOP paralogs and surrounding transporters (Figure 6) is intriguing and

suggests that they have an important role in the biology of *P. torquis* ATCC 700755^{T}. Whether this role is for export of toxic products of metabolism or deliberate release of products that can influence its interactions with other bacteria or algae is unknown.

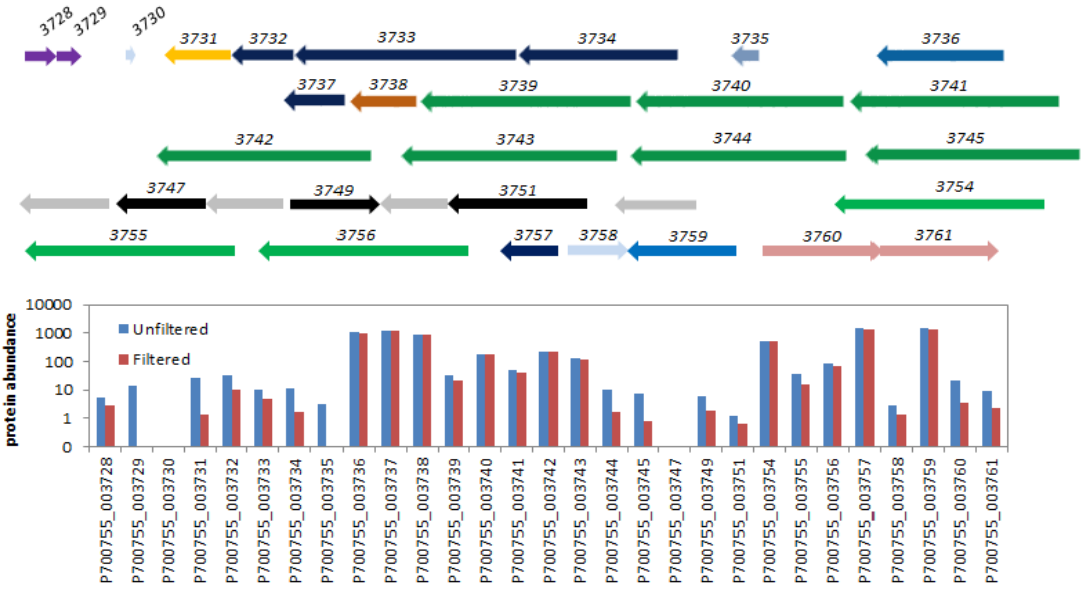


Figure 6. The genomic island region 43 of *P. torquis* ATCC 700755^{T} containing a highly translated cluster of transporter proteins including 8 ADOP family permease-like proteins. Relative protein abundance of gene products for this region is shown in the lower graph. Gray genes denote pseudogenes. Black genes denote intact transposases. Purple genes comprise an addiction module. Green genes denote the ADOP family permease-coding genes. Dark blue and indigo genes show different families of other transporters including those with ATP-binding regions. The orange gene codes a putative multifunctional acyl-CoA thioesterase. Other genes code hypothetical proteins. Further details are shown in Appendix B data files Table B.1 and B.3.

The genes coding the PUFA biosynthesis (*pfa*) cluster (P700755_01456 to _01462) (Figure 7) are located on a third GI (no.17). The ability to synthesise omega-3 and omega-6-type PUFA is a rare trait amongst bacteria; it is restricted to class *Gammaproteobacteria* and phylum *Bacteroidetes* and, within those groups, is largely restricted to marine psychrophiles (Shulse and Allen 2011). The *pfa* cluster of ATCC

700755^T is similar to previously described clusters but has an altered structural arrangement of conserved domains compared to those found in *Gammaproteobacteria* suggesting a different evolutionary process of acquisition. ATCC 700755^T, which can synthesise eicosapentaenoic acid (EPA) via this cluster, has higher levels of EPA in its cytoplasmic membrane at low temperatures (Nichols et al. 1997). Protein abundances were found to be substantial for *pfa* cluster gene products (Figure 7). Unlike *Gammaproteobacteria*, which form either EPA or docosahexaenoic acid, ATCC 700755^T also forms arachidonic acid, though its levels do not increase with temperature (Nichols et al. 1997) so it may have another role. We assume arachidonic acid is another by-product from the *pfa* cluster. PUFA is capable of maintaining homeoviscosity of membranes at very low temperatures (Russel and Nichols 1999; Usui et al. 2012) and, due to increasing cell hydrophobicity, also potentially shields cells against hydrophilic toxic substances such as peroxides (Nishida et al 2010). Two genome sequenced relatives of *P. torquis*, strain SCB49 (related to the genus *Ulvibacter*) isolated from the Arctic Ocean and a strain of the genus *Dokdonia* isolated from Arctic Ocean marine sediment (classified as *Krokinobacter* sp. 4H-3-7-5) possess *pfa* gene clusters very similar to that of *P. torquis*. This suggests that this type of *pfa* gene cluster could be prevalent in other psychrophilic members of the family *Flavobacteriaceae*. The *pfa* cluster in ATCC 700755^T is flanked by a number of intact and remnant transposases and a DNA-binding excisionase is located immediately upstream, suggesting the *pfa* cluster may have been mobilised into the ATCC 700755^T genome via a phage insertion or conjugative transposon. PUFA production is also

very sensitive to lipid oxidation (Imlay 2003). The sea-ice environment is generally saturated with oxygen due to photosynthetic activity and low temperature (D'Amico et al. 2006), which may partly explain why ATCC 700755^{T} possesses a wide array of enzymes that provide immediate protection against reactive oxygen species (ROS) and organic peroxides. The array of defences includes genes for two catalases (P700755_00288, _02059), seven peroxidases (P700755_00120, _0196, _01102, _01308, _03056, _03338, _03478) and three superoxide dismutases (P700755_00728, _00729, 01787). Some of these have homologs in ACAM 44^{T} which likely also experiences photooxidative stress in its lake environment. However, several are located on GIs in ATCC 700755^{T} including genes coding a diheme cytochrome peroxidase (GI no. 41 - P700755_03478), nickel- and iron-based superoxide dismutases (GI no. 8 - P700755_00728, _00729, GI no. 19 - _01787) and a putatively secreted catalase (P700755_00288).

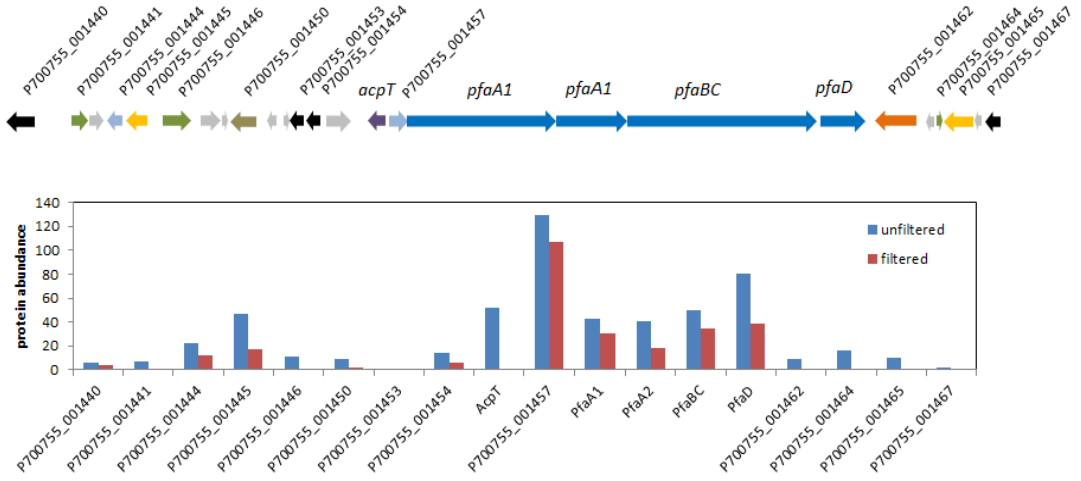


Figure 7. The gene content of the genomic island region 17 of *P.torquis* ATCC 700755^{T} (upper graph) containing the omega-3 polyunsaturated fatty acid (*pfa*) gene cluster and the relative protein abundance of gene products for this region (lower graph). The *acpT*/P700755_1457 (TetR family protein)/*pfaA1A2BCD* region is conserved in two other genome sequenced bacteria: strain SCB49 and *Krokinobacter* sp. 4H-3-7-5. Gray genes denote pseudogenes. Black genes denote transposases. Green genes are hypothetical proteins. Putative TetR family transcriptional regulators are shown in pale blue. Other genes shown are likely enzymes but function is only generally predicted from conserved domain data.

Occurrence of the Proteorhodopsin Gene of Psychroflexus torquis ATCC 700755^{T} at the Edge of a Genomic Island that Contains Putative Ice-Binding Proteins

Some *P. torquis* genes are suspected of having a role in sea ice inhabitation but the functions remain tentative and the coded proteins were generally only weakly abundant in the proteomic survey. The suppositions are based on the likelihood they could be advantageous in a sea-ice ecosystem setting with translation requiring specific conditions. One such trait is the ability to bind and/or interact with ice crystalline surfaces, aiding recruitment and persistence within sea ice (Raymond et al.

2007). Genes that could have this role in ATCC 700755^{T} includes a cluster of secreted proteins that may act as adhesins and that have a C-terminal domain homologous to other ice binding proteins, including those of the sea-ice diatom *Fragilariopsis cylindrus* (Bayer-Giraldi et al. 2010). Interestingly, the putative ice-binding/adhesin proteins are located in a GI (no. 2), which has at its edge a proteorhodopsin (PR) gene and its cognate carotenoid monooxygenase and, immediately upstream, ABC-type transporters for complexed iron or cobalamin. The genetic arrangement of PR is largely conserved in ACAM 44^{T}; however, the putative ice-binding protein cluster is mostly absent (Figure 8). The adjacent ice-binding proteins can be readily detected but are far-less abundant, perhaps pointing to a more generalised and important role for PR. The putative ice-active proteins found in *P. torquis* are related to other proteins found in bacteria, algae, and yeast (Figure 8), several with confirmed ice-binding and anti-freeze functions; nevertheless, substantial work is required to substantiate their functionality. Located on GI no. 17, adjacent to yet another putative ice-binding protein, lies a series of two component sensor systems clustered over a 16kb region, which are weakly orthologous to bacteriophytochromes and contain multiple PAS and GAF domains that have been suggested to be light sensors in the genome of the PR-possessing strain MED152 (Gonzalez et al. 2008). These proteins are only weakly to moderately abundant and a photobiological function remains unconfirmed. *P. torquis*, however, responded significantly to light, increasing its growth yield by 2-3 times depending on the salinity and light level (Feng et al. 2013), which suggests the presence of sensing and regulatory systems that must involve an

ability to respond to changing light and salinity conditions.

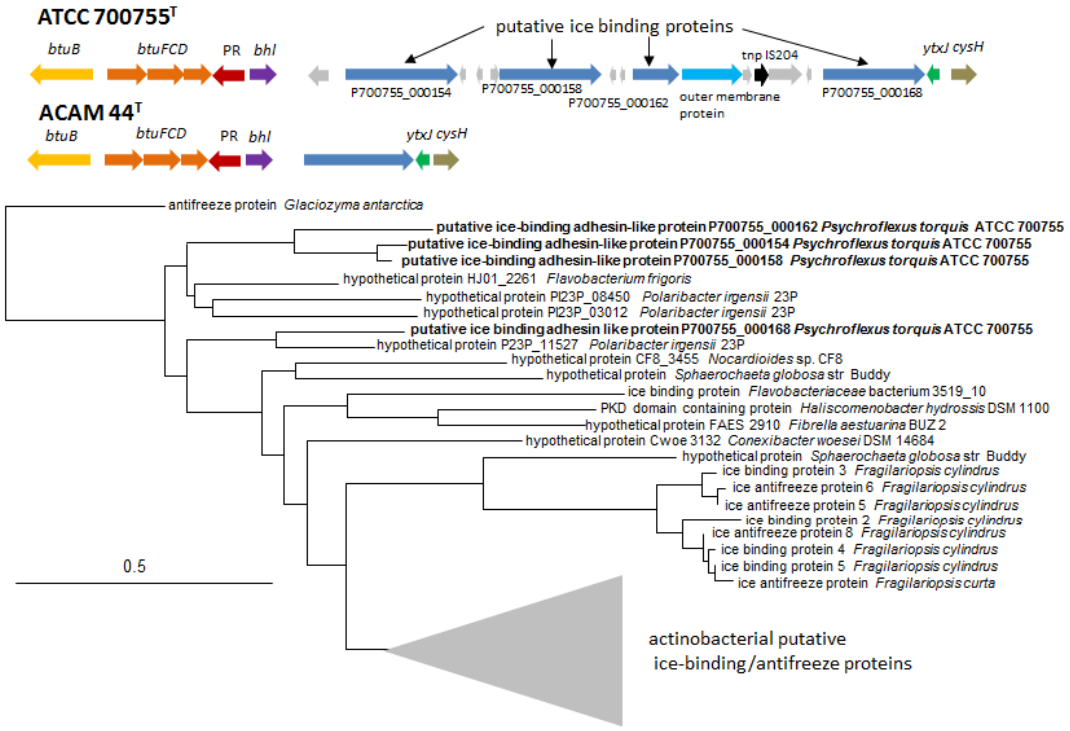


Figure 8. The PR and putative ice-binding protein gene cluster associated with the genomic island region 2 of *P. torquis* ATCC 700755^{T} genome compared with equivalent region from *P. gondwanensis* ACAM 44^{T}. Genes shown in gray are pseudogenes (see Table B.1, B.2 and B.3 for more details). Known and putative ice binding proteins within GI no. 2 are compared to other equivalent proteins from bacteria, diatoms (*Fragilariopsis* spp.) and yeast (*Glaciozyma antarctica*) in a protein sequence-based tree, where distances were calculated with the Grishin algorithm and clustering was via Neighbour-Joining, calculated using the constraint-based multiple alignment tool (www.ncbi.nlm.nih.gov) with default parameters.

Aspects of the P. torquis ATCC 700755 Genome Linked to Nutrient Acquisition and Algal Colonisation in Sea-Ice

A critical aspect for sea ice adaptation is nutrient acquisition. The nature of this ability can be partly surmised by the already known phenotype of strains of *P. torquis*. The species uses an eclectic range of carbon and energy sources including limited ranges

of carbohydrates, amino acids, organic acids, and odd chain length lipid oxidation products (Bowman et al 1998). In sea ice, levels of dissolved organic compounds are at much higher levels than in the water column; they include large amounts of exopolymeric material, carbohydrates, free amino acids, and lipids (Norman et al 2011). ATCC 700755^{T} was thus expected to have a complement of genes specifically geared to access abundant substrates formed during algal primary production, which it may share with ACAM 44^{T}. Both strains are relatively fastidious and require growth factors, including vitamin B6 and cobalamin consistent with their gene content for cofactor metabolism. Several nutrient acquisition-related proteins coded by GI-associated genes were found to be comparatively abundant based on the proteomic analysis (Table B.1). These included: a number of secreted and non-secreted peptidases (GI no. 4 – glutamate carboxpeptidase II, L-carnosine dipeptidase; GI no. 13 two subtilisin-like proteases); enzymes for catabolism of DL-hydroxyproline (GI no. 34, HCP deaminase, 4-hydroxyproline epimerase); a putative SGLT1-like family glucose/Na^{+} ion symporter; and several TonB-dependent outer membrane receptor and binding proteins, mainly of the RagA/SusCD families (GI no. 5, 13, 16), which typically take up carbohydrates, polypeptides, and/or chelated metallic cations. Since sea ice is a sub-zero temperature environment these proteins are almost certainly cold-active (Huston et al. 2000). The genome of *Colwellia psychrerythraea* 34H was noted to possess a large number of extracellular enzyme coding genes (Methé et al. 2005) and dwelling at extreme cold seems to require the production of large amounts of extracellular enzymes to overcome the mass transfer limitations caused by low

temperature impairment of enzyme function and transport (Struvay and Feller 2012).

Another aspect of sea ice inhabitation is the link *P. torquis* has to algae as an epiphyte. Given that *P. torquis* has substantial growth factor requirements and is slow growing (fastest doubling time ~ 1 day), persisting in a dynamic system may require close interaction with sources of nutrients. Details of bacterial/algal associations derivable from the genome data are so far effectively limited to inferences; however, they are compelling and provide several possible lines of research on competitive and mutualistic interactions in sea ice. Both ATCC 700755^{T} and ACAM 44^{T} possess a conserved set of surface-gliding motility-associated genes (McBride and Zhu 2013) and the associated Por secretion system (Sato et al 2010); however they strongly differ in terms of cell-surface proteins and in the presence of several adhesin-like proteins (Fig. 2). ATCC 700755^{T} is able to perform a form of slow gliding motility (Bowman et al. 1998) which has not been demonstrated in ACAM 44^{T}.

Putative adhesins present in ATCC 700755^{T} include several types such as the aforementioned putative ice-binding adhesin-like proteins, VCBS and fasciclin repeat domain-containing proteins and autotransporter adhesins. ATCC 700755^{T} also possesses a large surface protein of 468 kDA (P700755_00663) homologous to colossin A protein from the slime mold *Dictyostelium discoideum* (Whitney et al. 2010) that could also be involved with surface interactions. The diverse range of these proteins along with ice-active proteins seems to suggest a flexible attachment ability potentially necessary in the highly dynamic sea-ice ecosystem.

The possibility emerges that the algal interactions of ATCC 700755 run deeper than

simple commensalism. Both ATCC 700755^{T} and ACAM 44^{T} strains possess several signalling proteins that putatively respond to the presence of plant hormones, including GH3 auxin promoter proteins that may allow them to coordinate growth and activity with that of algal hosts (Lambrecht et al. 2000). Unusually, ATCC 700755^{T} secretes 2-phenylethylamine (PEA) in substantial levels (Hamana and Niitsu 2001). The physiological function of PEA in relation to algae is unclear, but PEA has been found to induce production of oxidative bursts in tobacco that has relatively high levels of PEA in its leaves and could be linked to triggering defence systems (Kawano et al. 2000). An intriguing 35-kb cluster of proteins (P700755_01235 to _01257) on GI no. 15, neighbouring a glycogen synthesis cluster (P700755_01229 to _01232) could also be involved in algal interactions (Figure 9). The cluster includes signalling proteins with CHASE domains as well as several large WD40 repeat domains containing 4 metacaspase family proteins. Metacaspases have been linked to programmed cell death functions in lower eukaryotes including phytoplankton (Madeo et al. 2012, Choi and Berges 2013). At this stage few details on the functionality of bacterial metacaspases are available (Vercammen et al. 2007). Cooperative and non-cooperative interactive mechanisms between algae and bacteria, whose populations are functionally tightly coupled, are a critical aspect of sea-ice ecology, yet remain largely unknown.

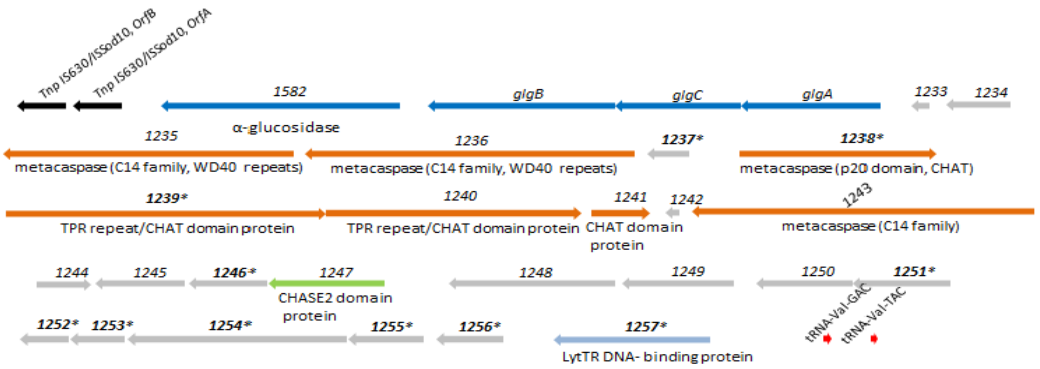


Figure 9. The genomic island region 15 of *P. torquis* ATCC 700755^{T} that contains a cluster of putative metacaspase coding genes (orange) following a glycogen-synthesis/utilisation cluster (blue) with a possible role linked to programmed cell death. The domain structure of metacaspase genes and associated genes are indicated. Genes flagged with asterisks are those with signal peptide regions and thus putatively secreted. A putative LytTR family regulatory protein is coded by gene P700755_001257, while a CHASE2 family protein (usually associated with extracellular sensory proteins) is shown in green. The region is flanked by tranposases (black genes) and two tRNA-Val genes.

Conclusions

Staley and Gosink (1999) indicated that a number of bacterial genera exist in both Arctic and Antarctic sea ice，but whether this finding can be extended to the species level was unknown (Staley and Gosink 1999). They argued that the global distribution of psychrophiles was essentially blocked by a tropical marine barrier, providing the possibility that ecosystem-linked endemism could emerge. Recently evidence based on high throughput sequencing of marine microbial communities has been presented that, like macroscopic organisms, marine bacteria seem to be subject to biogeographic limitations affecting their current and presumably long term distribution. Not only do these proposed limitations suggest the potential for localised speciation in certain ecosystems (Whitaker 2006) but it also suggests that some communities of bacteria

are vulnerable to extinction brought about by external disruption such as climate change, habitat destruction, or invasion by other organisms (Sul et al. 2013). Here we demonstrate the comparative genomic features of the bipolar bacterial species *P. torquis* that could be an excellent example of evolving endemism in a bacterial species. A next step in genome-level analysis would be to compare Arctic and Antarctic strains to determine genetic similarity and degree of change, especially in the number and content of GIs, relative state of gene decay and overall occurrence of HGT gene regions, which made up approximately 35% of the genome of strain ATCC 700755^{T}. This major finding of our study is consistent with the suggestion, based abundance of phage and extracellular DNA in sea-ice brines, that sea ice is a hotspot for HGT (Collins and Deming 2013). High levels of HGT in other sea ice-associated bacteria can also be expected. Sea ice or polar seawater associated *Octadecabacter* species, which possess the light-driven proton pump xanthorhodopsin, also show evidence of high levels of HGT and have genomes rich in IS elements, pseudogenes and possess plasmids (Vollmer et al. 2013). Sea ice specialism apparent in *P. torquis* appears to be linked to its GI-associated genes including those for EPS and PUFA synthesis, modes of nutrient acquisition and potential ice-binding and algal interactions. Proteomic data provided evidence that many of the associated genes are being actively translated and thus important to the biology of *P. torquis*. Overall this study suggests *P. torquis* could be an excellent model to study sea ice functional biology and evolutionary processes linked with psychrophily, endemism and algal interactions.

Chapter 3

Light stimulated growth of proteorhodopsin bearing sea-ice psychrophile *Psychroflexus torquis* is salinity-dependent

Abstract

Proteorhodopsins (PRs) are commonly found in marine prokaryotes and allow microbes to use light as an energy source. In recent studies it was reported that PR stimulates growth and survival under nutrient-limited conditions. In this study we tested the effect of nutrient and salinity stress on the extremely psychrophilic sea-ice bacterial species *Psychroflexus torquis*, which possesses PR. We demonstrated for the first time that light stimulated growth occurs under conditions of salinity stress rather than nutrient limitation and that elevated salinity is related to increased growth yields, PR levels and associated proton pumping activity. PR abundance in *P. torquis* also is post-transcriptionally regulated by both light and salinity and thus could represent an adaptation to its sea-ice habitat. Our findings extend the existing paradigm that light provides an energy source for marine prokaryotes under stress conditions other than nutrient limitation.

Introduction

The discovery of proteorhodopsin (PR) in 2000 challenged the assumption that solar energy is captured mainly through the use of chlorophylls and other analogous pigments (Fuhrman et al., 2008). PR genes are found in various ecosystems (Sabehi et

al., 2004, Venter et al., 2004, Giovannoni et al., 2005, Frigaard et al., 2006, Atamna-Ismaeel et al., 2008, Koh et al., 2010) and they are amongst the most highly expressed and widely distributed proteins in marine bacterial communities(Frias-Lopez et al., 2008). Theoretically, since PR-bearing bacteria can use light energy to supplement energy, it should make them more competitive in situations where other sources of energy are limiting. PR functions as a transmembrane light-driven proton pump which, when exposed to light, creates a proton-motive force that can drive chemical transport processes, motility, and generate ATP (Béjà et al., 2000, Béjà et al., 2001). To show this, marine-derived PR genes were cloned and expressed in *E. coli* which then exhibited increased proton pumping (Béjà et al., 2000). Furthermore, when the recombinant strain was provided with retinal, Walter et al (2007) demonstrated that the recombinant *E. coli* strain showed increased motility under light and that light energized electron transport, despite cellular respiration being impeded by anoxia and sodium azide. The authors suggested that PR can replace respiration as a cellular energy source under some environmental conditions (Walter et al., 2007).

Since the discovery of PR, several workers have attempted to demonstrate the physiological role of PR and its effects on bacterial growth. However, laboratory experiments examining light-stimulated growth in PR-containing bacteria have not produced consistent results. For example, no increase in growth rate and/or final cell yield under illuminated conditions was observed in SAR11 and SAR92 strains (Giovannoni et al., 2005, Stingl et al., 2007). In contrast, light-enhanced growth was

reported for *Dokdonia* sp. MED134 under carbon limited conditions (Gomez-Consarnau et al., 2007). Later it was demonstrated that PR-containing *Vibrio* strain AND4 had increased long-term survival when starved if it was exposed to light based on comparisons to a PR deficient mutant (Gomez-Consarnau et al., 2010). A light/dark growth experiment was conducted on natural mixed communities of oceanic bacteria collected from a low-nutrient ocean environment and most of the marine clades that are thought to possess PR showed different responses to light (Schwalbach et al., 2005).

Currently nearly all hypotheses argue that PR could provide an adaptive advantage to the bacteria under oligotrophic conditions. To date all studies have focused on explaining the physiological role of PR that might affect bacterial growth during periods of low-nutrient or carbon limited conditions (Giovannoni et al., 2005, Gomez-Consarnau et al., 2007, Michelou et al., 2007, Stingl et al., 2007, Lami et al., 2009, DeLong and Béjà, 2010, Gomez-Consarnau et al., 2010, Steindler et al., 2011). Nevertheless, the data to date do not consistently support this hypothesis. Moreover, the non-uniform response between different organisms that all express the PR gene indicates the possibility of multiple functions of PR in these bacteria (Fuhrman et al., 2008). As described above, PR are widespread in natural environments and these environments may have many other conditions in which PR may be advantageous. We hypothesize that PR may be an important provider of light derived energy under stress conditions that are associated with a specific econiche in a nutrient independent manner. Here we report that light-stimulated growth of the PR-containing sea-ice

bacterial species *Psychroflexus torquis* ATCC 700755[14] occurs in concordance with increased PR activity. We also show that PR abundance is dependent on osmotically different conditions rather than carbon limited conditions.

Materials and Methods

Bacterial Strains and Media

Psychroflexus torquis ATCC 700755^{T} was originally isolated from multi-year sea-ice collected from Prydz Bay, Vestfold Hills, Antarctica (S 68°36'30'', E 78°04'19") (Bowman et al., 1998). It was routinely grown in modified marine broth (mMB) (5g of proteose peptone (Oxoid), 2g of yeast extract (Oxoid) 35g of sea salt (Red Sea) in 1L of distilled water aerobically at 5ºC.

Growth Experiments

Experiments were carried out to determine the effect of light under conditions of salinity and nutrient stress. The effect of salinity stress was tested with a series of modified marine broths with different salt concentrations: 17.5 $g.L^{-1}$, 35 $g.L^{-1}$, 52.5 $g.L^{-1}$, and 70 $g.L^{-1}$. 100 ml of triplicate cultures were inoculated with 100 µL working culture (6.0 $\log_{10}$ CFU/ml) and incubated aerobically at 5ºC. Cultures were constantly exposed to three light intensities: high light (27.7 µmol photons m^{-2} s^{-1}); low light (3.7 µmol photons m^{-2} s^{-1}) and complete darkness (0 µmol photons m^{-2} s^{-1}). In all experiments, light was provided from 30W fluorescent lamps. Nutrient stress experiments were conducted under the same temperature and light conditions

described above with mMB diluted 1:10, 1:100, 1:1000 and 1:10000 with artificial sea water such that the salinity was kept constant but the organic content of the medium was diluted. To determine cell numbers direct counts were performed microscopically using a counting chamber.

RNA Extraction

Cells of *P. torquis* were cultured as described above and harvested by centrifugation at 14 000 rpm for 15 min within the early, middle and late exponential phase. The cell pellets were stored in RNAlater (Ambion Inc,) and kept at -20ºC. Total RNA was extracted using the Qiagen RNeasy Mini Kit (Qiagen), according to manufacturer's instructions. In addition an on-column DNase treatment using RNase-free DNase (Qiagen) was performed to remove contaminating genomic DNA. The RNA concentration was determined using a Nanodrop 1000 spectrometer (Thermo Fisher).

Analysis of PR Expression by Quantitative Real-Time PCR (qRT-PCR)

For qRT-PCR analysis 16S rRNA was used as a reference gene. The qRT-PCR was performed in duplicates with 10 ng of the total RNA in a final volume of 25 µL using a one-step QuantiTect SYBR Green RT-PCR Kit (Qiagen). PR primers used in this experiment were 5'-TAT GGC CAT GTT GGC TGC AT-3' (forward) and 5'-CTG AAG CTA GAC CCC ACT GC-3' (reverse) while 16S rRNA primers were 5'-CAG CMG CCG CCG TAA TAC-3' (forward) and 5'-CCG TCA ATT CMT TTR AGT TT-3' (reverse). PCR thermocycling was performed using a Rotor-Gene (Corbett

Research) RG-3000 with reverse transcription performed at 50ºC for 30 min followed by DNA polymerase activation step at 95ºC for 15 min, followed by 35 cycles of amplification at 94ºC for 15s, annealing temperatures 52ºC for 30s and extension at 72ºC for 30s. Only single peaks were observed in melting curves for both the 16S rRNA and PR genes indicating specific amplification with each set of primers. Amplification efficiencies were between 0.85 - 0.98 and the correlation coefficient for a 10-fold dilution series was between 0.96 - 1.00. Relative quantification of PR and 16S rRNA gene mRNA transcripts utilized the Pfaffl equation (Pfaffl, 2001).

Spectral and HPLC Analysis of Pigment

100 μL of *P. torquis* working culture was used to inoculate modified marine agar with the range of salinity conditions described above. Agar plates were incubated aerobically at 5ºC for 23 days. Biomass was harvested and freeze dried using a Dyna-Vac freeze dryer (Technolab，Kingston，TAS, Australia); 4 mg of freeze-dried biomass were extracted with 1ml of 3:3:3:1 methanol : acetone : N,N'-dimethylformamide : water; at 4ºC in the dark (Hagerthey et al., 2006). Absorption spectra were measured using a Thermo Scientific GENESYS 10S UV-Vis，utilizing quartz cuvettes. Spectra were recorded in 0.5-nm steps between 250 and 600nm. The control baseline was established using the same amount of solvent without biomass. The pigment was extracted and analysed by HPLC using the method described by Sabehi et al (2005) to identify the presence of carotenoids and retinal.

Light-Driven Proton Pump Activity Analysis

50ml of mMB with different salinities as outlined above, were inoculated with 50 μL of working culture and incubated aerobically at 5ºC under dark for 23 days. The cells were collected by centrifugation (14 000 rpm, 15min at 4ºC) and then washed twice in artificial seawater. Cell pellets were resuspended in 10 ml of artificial seawater and held for 1 day at 5ºC. Proton pump activity was measured by monitoring pH changes in the suspensions. The suspensions were first placed in darkness and then illuminated by light (~27.7 μmol photons $m^{-2}s^{-1}$) for 30 minutes. pH changes during incubation were monitored and recorded with a pH meter from BioFlo/CelliGen115 Fermentor/Bioreactor (New Brunswick). Measurements were repeated under the same conditions after addition of protonophore carbonylcyanide-3-chlorophenylhydrazone (CCCP) to a final concentration of 10μM (Yoshizawa et al., 2012).

LC/Tandem MS-MS analysis of PR abundance

Protein Extraction

Samples were lysed in 1ml 50mM Tris-HCL buffer (pH 7.0) (Oxoid) by sonication in ice bath. Cell lysis process were conducted by 10s of sonication with 10s rest for 15 times until the cell suspension was clear. Cell suspensions were then centrifuged at 16,000g for 25mins at 4 ºC. The supernatant (with soluble protein) was used to perform a TCA precipitation and the protein pellets were then treated with 0.2M NaOH to make protein more soluble as described by Nandakumar et al (2003). The pellets were then solubilised using 100 μl 50mM Tris-HCL buffer (pH 7.0) (Oxoid).

Protein concentrations were routinely determined using the Bradford assay (Bio-rad). Volumes of sample containing 50 µg of the soluble protein extracts were taken and reduced by reducing agent (50 mM dithiothreitol, 100 mM ammonium bicarbonate) and incubating for 1 h at room temperature. Followed by alkylation by 200 mM iodoacetamide, 100 mM ammonium bicarbonate and incubating for 1h at room temperature. After alkylating, the samples were reduced further of the reducing agent and incubating for 1h at room temperature. 220 µl of digestion buffer (50mM ammonium bicarbonate, 1 mM calcium chloride) was added to the samples and the proteins were digested with sequencing grade modified trypsin (Promega, Madison, WI, USA, V5111), at a sample protein to trypsin ration of 25:1, at 37 °C with gentle shaking, overnight. Digestion was stopped by acidification with 10 µl of 10% formic acid. The samples were centrifuged for 5 min at 14 000 g to remove any insoluble material, and an aliquot (100-200 µl) protein was transferred to HPLC vials for 1D-LC-MS/MS analysis.

1D-LC-MS/MS

The separation of peptides utilized a Surveyor Plus HPLC system fitted in line with an LTQ-Orbitrap XLmass spectrometer (ThermoFisher Scientific). Peptide samples were loaded onto a C_{18} capillary trapping column (Peptide CapTrap, Michrom BioResources) using initially 5% acetonitrile in 0.2% formic acid (solvent A) shfiting up to 90% acetonitrile in 0.2% formic acid (solvent B) over a 5 min period. Sample was subsequently back flushed onto a pre-equilibrated analytical column (Vydac

Everest C18 300 Å, 150 µm ×150 mm; Alltech). Peptides were then separated on an analytical C_{18} nano-column (PicoFrit Column, 15 µm i.d. pulled tip, 10 cm, New Objective) using four linear gradient segments (0-10% solvent B over 7.5 min, 10-25% solvent B over 50 min, 25-55% B over 20 min and then 50–100% solvent B over 15 min), holding at 100% solvent B for a further 15 min before returning to 100% solvent A over 15 min. The LTQ-Orbitrap was fitted with a dynamic nanoelectrospray ion source (Proxeon) containing a 30 µm inner diameter uncoated silica emitter (New Objective). The LTQ-Orbitrap XL was controlled using Xcalibur 2.0 software (ThermoFisher Scientific) and operated in data-dependent acquisition mode whereby the survey scan was acquired in the Orbitrap with a resolving power set to 60,000 (at 400 m/z). MS/MS spectra were concurrently acquired in the LTQ mass analyser on the seven most intense ions from the FT survey scan. Charge state filtering, where unassigned precursor ions were not selected for fragmentation, and dynamic exclusion (repeat count, 1; repeat duration, 30 s; exclusion list size, 500) were used. Fragmentation conditions in the LTQ were: 35% normalized collision energy, activation q of 0.25, 30 ms activation time, and minimum ion selection intensity of 500 counts.

1D-LC-MS/MS Database Searching

The acquired MS/MS data were analysed using Mascot version 2.2.06 (Matrix Science). Proteome Discoverer version 1.2 (Thermo Scientific) was used to extract tandem mass spectra from Xcalibur raw files and submit searches to an in-house

Mascot server according to the following parameters: S-carboxamidomethylation of cysteine residues specified as a fixed modification and cyclisation of N-terminal glutamine to pyroglutamic acid, deamidation of asparagine, hydroxylation of proline, oxidation of methionine specified as variable modifications. Parent ion tolerance of 10 ppm and fragment ion mass tolerances of 0.8 Da were used, and enzyme cleavage was set to trypsin, allowing for a maximum of two missed cleavages.

For protein identification, the Mascot search results were loaded into Scaffold version 3.0.08 to assign probabilities to peptide and protein matches (Searle, 2010). Peptide-spectrum matches were accepted if the peptide was assigned a probability greater than 0.95 as specified by the Peptide Prophet algorithm (Keller et al., 2002). The probability threshold of 0.95 showed good discrimination between the predicted correct and incorrect peptide-spectrum assignments, and only peptides with charge states of +1, +2, and +3 were retained as confident identifications because the Peptide Prophet models were not a good fit to the data for charge states $\geq$ 4. Protein identifications were accepted if the protein contained at least two unique peptides (in terms of amino acid sequence), and the protein was assigned a probability > 0.95 by the Protein Prophet algorithm (Nesvizhskii et al., 2003). This threshold will constrain the protein false discovery rate (FDR) to < 1%. The minimal list of proteins that satisfy the principal of parsimony is reported. Spectral count was used to determine relative protein abundance (Nesvizhskii et al., 2003; Liu et al., 2004).

Results

The Light Stimulated Growth Response of Psychroflexus torquis is Salinity not Nutrient Dependent

To assess whether light is required for PR-mediated proton translocation, and whether it affected growth rate and increased the final cell population, the psychrophilic species *P. torquis* ATCC 700755 was grown under nutrient-limited with constant salinity level and also under nutrient-rich conditions over a range of growth permissive salinities (Bowman et al., 1998). Preliminary carbon limitation experiments showed that *P. torquis* did not significantly exhibit differences in cell abundance under the carbon limited conditions tested (see Appendix A, Figure A.1). Specifically, under constant carbon limitation (70 mg l^{-1} organic nutrient) *P. torquis* grown under different levels of illumination did not significantly differ in cell abundance nor was growth rate altered (Figure 1a). However, when exposed to different levels of salinity under nutrient-rich conditions (5 g l^{-1} organic nutrient) both growth rate and cell abundance were always higher under illumination compared to when grown in the dark (Figure 1b-e). Maximal differences in growth yield between light and dark incubation occurred at sub- and supra-optimal salinities. Specific growth rates and yields were stimulated about 1.5 d^{-1} and 2-fold under 3.7 μmol photons m^{-2} s^{-1} illumination respectively. Under 27.7 μmol photons m^{-2} s^{-1} illumination growth rates were similar but the growth yield differential was markedly reduced suggesting high levels of irradiance become photo-inhibitory to *P. torquis*.

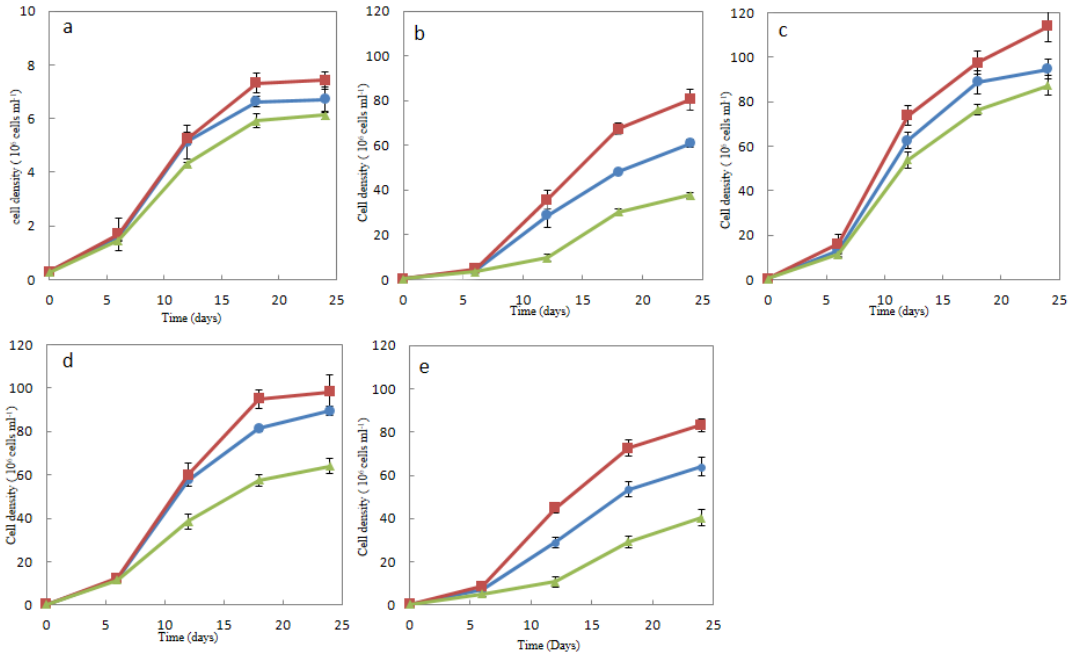


Figure 1 Growth of *P. torquis* grown in 1:100 diluted modified marine broth at 35 $g.L^{-1}$ (a) and in modified marine broth with salinities of 17.5 $g.L^{-1}$ (b), 35 $g.L^{-1}$ (c), 52.5 $g.L^{-1}$ (d), and 70 $g.L^{-1}$ (e) under 3.7 PPF (■), 27.7 PPF (●) and 0 0 PPF (▲). Cell populations were determined microscopically.

P. torquis Pigment Synthesis is Light and Salinity Dependent

Since PR is a retinylidene protein and its function depends on a retinal chromophore derived from carotenoid compounds. The presence of retinal and an unidentified type of carotenoid was confirmed by HPLC (Figure 2a-b). To determine whether illumination stimulates carotenoid biosynthesis, pigment analysis of biomass grown under different levels of illumination and salinities was assessed. Pigment extracts had a major spectral absorbance region between 425 - 525 nm which represents carotenoids. When compared with dark-incubated cultures illumination stimulated pigment production though the effect was primarily influenced by salinity as well as the light intensity (Figure 3a-d). Illumination-driven pigment biosynthesis stimulation

occurred maximally at the optimum salinity for growth, which is equivalent to that of seawater; however at a salinity of 52.5 g.L^{-1} incubation under low irradiance did not correspond to stimulated pigment production (Figure 3c).

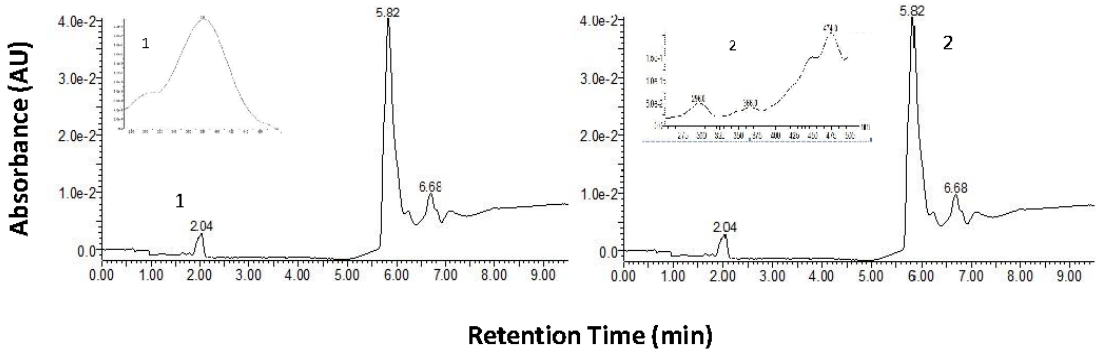


Figure 2 HPLC separation of the carotenoid and retinal from extracted pigment. Insert graphs show the absorption spectra of peaks 1(retinal) and 2 (unknown carotenoid).

Light-Driven Proton Activity of P. torquis is Light Activated and Salinity Dependent

The light-driven proton pump activity of *P. torquis* was also tested at 1ºC under different levels of salinity. When a suspension of *P. torquis* (~ 10^8 cells ml^{-1}) was illuminated， the external pH declined rapidly (Figure 4) indicative of proton extrusion. This was confirmed by the use of the CCCP to uncouple H^+ translocation (Yoshizawa et al., 2012). The light-driven H^+ translocative driven pH changes were influenced by the salinity with maximal H^+ pumping occurring at 52.5 g.L^{-1} (Figure 4) and under 27.7 μmol photons m^{-2} s^{-1} illumination.

PR Gene Expression is Constitutive in P. torquis Regardless of Illumination or Salinity Level

Quantitative RT-PCR was used to determine the expression of the gene encoding PR during growth within early, middle, and late exponential growth phases. Dark incubated cultures were used as controls to allow calculation of the change in PR expression. Different salt concentrations or illumination levels had no significant effect on the expression of the PR gene (Table. 1). In addition, the expression of the PR gene was also unaffected during transition from exponential phase into stationary growth. Based on this data the PR gene appears to be constitutively expressed in *P. torquis* ATCC 700755.

Table 1 Relative quantitative of PR expression

	Early Exponential Phase			Middle Exponential Phase			Late Exponential Phase		
Illumination / Salinity	27.7 PPF	3.7 PPF	0 PPF	27.7 PPF	3.7 PPF	0 PPF	27.7 PPF	3.7 PPF	0 PPF
17.5 $g.L^{-1}$	1.19[a]	1.13	1.00	1.15	1.02	1.00	1.08	0.98	1.00
35 $g.L^{-1}$	1.12	1.15	1.00	1.08	1.10	1.00	1.04	1.05	1.00
52.5 $g.L^{-1}$	1.26	1.23	1.00	1.19	1.17	1.00	1.08	1.01	1.00
70 $g.L^{-1}$	1.09	1.11	1.00	1.05	0.99	1.00	1.03	1.05	1.00

[a]Dark conditions were used as the control for each treatment with fold change in PR mRNA transcriptm abundance compared to 16S rRNA, calculated using the Pfaffl equation (see methods section in main manuscript).

PR cellular Abundance in P. torquis is Influenced by Salinity and Light

We assessed the abundance of PR by using LC/tandem MS-MS analysis and spectral counting of PR peptides detected in trypsinized protein extracts. Based on the accumulated data of almost 700,000 peptide spectra that passed filtration criteria, PR peptides made up approximately 0.4% of the identified peptides. Taking into account the extraction bias that occurs with recovery of transmembrane proteins, PR must be amongst the most abundant proteins in *P. torquis*. The cellular abundance of PR was directly influenced by salinity as well as illumination (Figure 5). The relative abundance of PR was stimulated under both illuminated and dark conditions at seawater salinity levels with the differential effect of light levels being small but still noticeable. At 52.5 $g.L^{-1}$ a higher irradiance level lead to a much more pronounced stimulation in PR abundance relative to both dark and more dimly illuminated cultures. At the salinity extremes PR abundance was not affected by illumination. The results suggest PR abundance is controlled by salinity levels, and at normal sea water salinity levels or slightly above, illumination stimulates PR abundance.

Discussion

PR-containing prokaryotes constitute from 13% to 70% of prokaryotic cells in the marine ecosystem (Venter et al., 2004, Frigaard et al., 2006, Campbell et al., 2008) and thus play an important role in these ecosystems. The current hypothesis to explain the role of PR is that it provides a growth benefit during carbon starvation via H^+ translocation-mediated cellular processes. However, in most cases tested so far, the

presence of PR does not actually increase cell growth rates or yields. Fuhrman et al. (2008) suggested that PR might be used as a mechanism of survival under harsh conditions and that while there are only a few carbon-limited environments many other elements other than carbon can limit the growth of marine bacteria. This indicates that PR may have a broader range of physiological roles than previously reported which lead us to hypothesize that PR may stimulate bacterial growth under stressful conditions, such as osmotic pressure. In this study, we observed light-stimulated growth under osmotic pressure rather than carbon limitation in *P. torquis* ATCC 700755. The results from growth experiments showed a stimulatory influence of PR, similar to that reported for *Dokdonia* sp. MED134 (Kimura et al., 2011), on the growth yield of *P. torquis* under salt stress but not nutrient stress (Figure 1). We also observed that different light levels affected this growth response. According to our study, a light intensity of 3.7 μmol photons m^{-2} s^{-1} is most beneficial to the growth of *P. torquis*.

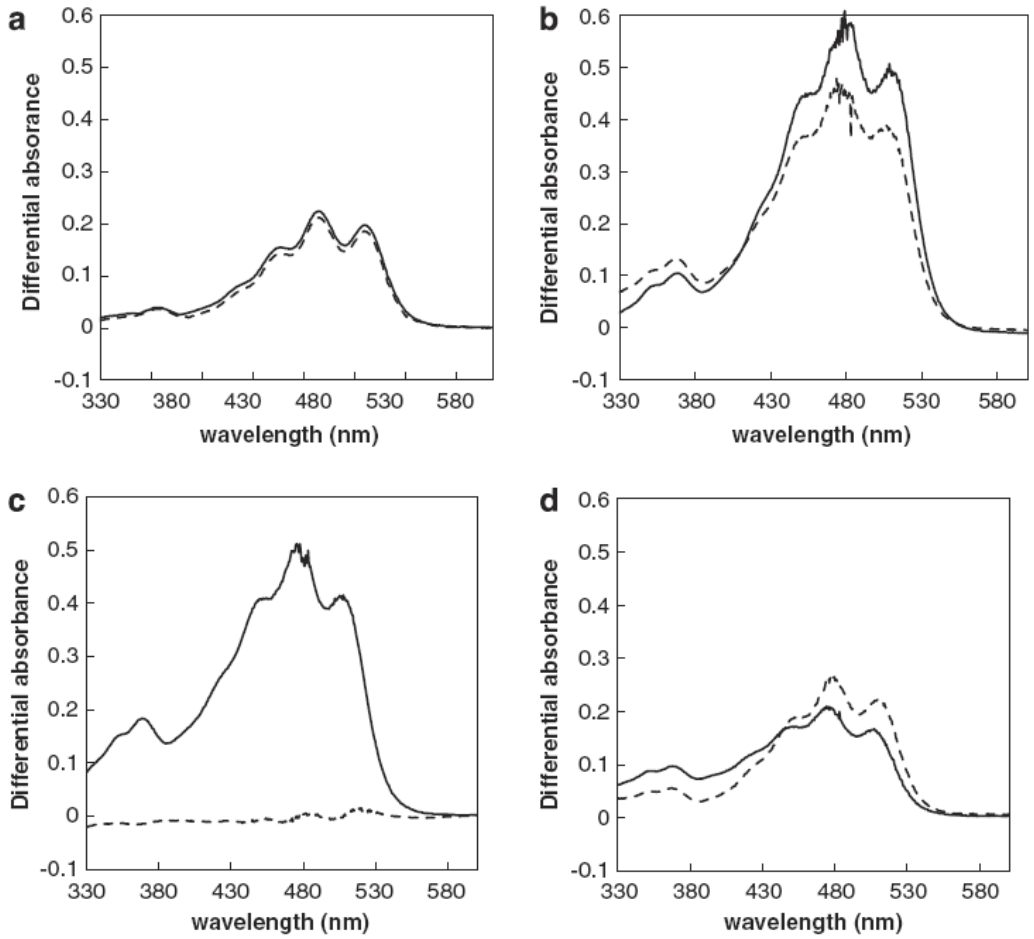


Figure 3 Pigment levels in *P. torquis* ATCC 700755 grown under different light levels (high: 27.7 PPF, low: 3.7 PPF) and salinity levels. The graph shows absorption spectra of pigment extracted from *P. torquis* grown in modified marine broth with 17.5 $g.L^{-1}$ (a), 35 $g.L^{-1}$ (b), 52.5 $g.L^{-1}$ (c), and 70 $g.L^{-1}$ (d). The solid line and dotted lines indicates the pigment extract absorbance levels in high and low light incubated cultures after being substracted from dark incubated cultures.

P. torquis inhabits sea-ice and possesses fastidious growth requirements requiring yeast extract in the media to allow growth (Bowman et al., 1998). It is suspected the species is an algal epiphyte. Sea-ice is a temporally transient and highly variable environment that experiences extremes of internal temperatures and salinity. It is, however a relatively carbon rich environment especially for algal-associated bacteria. As the ice crystals form and move in the water column, they trap inorganic and

organic materials on and between them, with the incorporation of silt, clay particles, grains, diatom and foraminferal detritus in the ice matrices (Grossmann and Dieckmann, 1994, Arrigo et al., 1997). Algal and bacterial cells that subsequently colonize and flourish in the ice matrices also synthesize copious amounts of exopolysaccharide (Reimnitz et al., 1993, Giannelli, 2001, Krembs et al., 2011). The salinity in sea-ice brine channels and pockets formed when salt is extruded during sea-ice formation is very variable and can rise as high as 150 $g.L^{-1}$ or drop to < 10 $g.L^{-1}$ when the sea-ice melts (Thomas and Dieckman, 2002). The optimum salinity for *P. torquis* is approximately 35 $g.L^{-1}$ (Bowman et al., 1998) and we showed that when the salt concentrations were above or below this level greater growth was recorded in the samples grown under light compared to those incubated in the dark. This is a particularly interesting result since *P. torquis* is exposed to fluctuating salt levels in its sea-ice habitat. We hypothesise that PR may create a proton gradient allowing the cell to maintain better sodium homeostasis and PR may also contribute to compatible solute uptake or biosynthesis, the latter being a particularly energy demanding process (Oren, 1999). Kimura et al. (2011) reported that Na^+-translocating NADH-quinone oxidoreductase and Na^+ transporters were greatly up-regulated in *Dokdonia* sp. MED 134 when exposed to light, which suggests that a sodium gradient may have an essential function in light-enhanced growth in PR-containing marine bacteria and it is even possible that some rhodopsins may act as a sodium-pumps (Kwon et al., 2012). Our study focussed on the effect of salinity therefore a Na^+-transporter may be coupled to the functioning of PR in several ways. It may simply be involved in

regulation of osmotic pressure or the Na^+ gradient may be used to generate the motive force for the respiratory electron chain instead of just the a H^+ gradient (Unemoto and Hayashi, 1993, Gonzalez et al., 2008). In sodium-rich environments, using a sodium-motive force may be more advantageous than using the proton-motive force to power energy transduction or, in some marine bacteria, the sodium-motive force couples with the proton-motive force (Kogure 1998, Albers et al., 2001). It is worth noting that *P. torquis* cultures incubated under high light have lower cell abundance and growth rate than low light treatments indicating that high light levels could be inhibitory and thus counteracting the benefits provided by PR. This could be due to photooxidative stress. However, cell abundance and growth rate still exceeded those found for the dark-incubated samples. This response to light intensity may result from *P. torquis* being adapted to relatively shaded conditions. In sea-ice photosynthetically available radiation can be very low, down to ~1 μmol photon m^{-2} s^{-1} due to the ice-sheet thickness and obscuration of light due to sea-ice algal assemblages (McMinn *et al*., 2007).

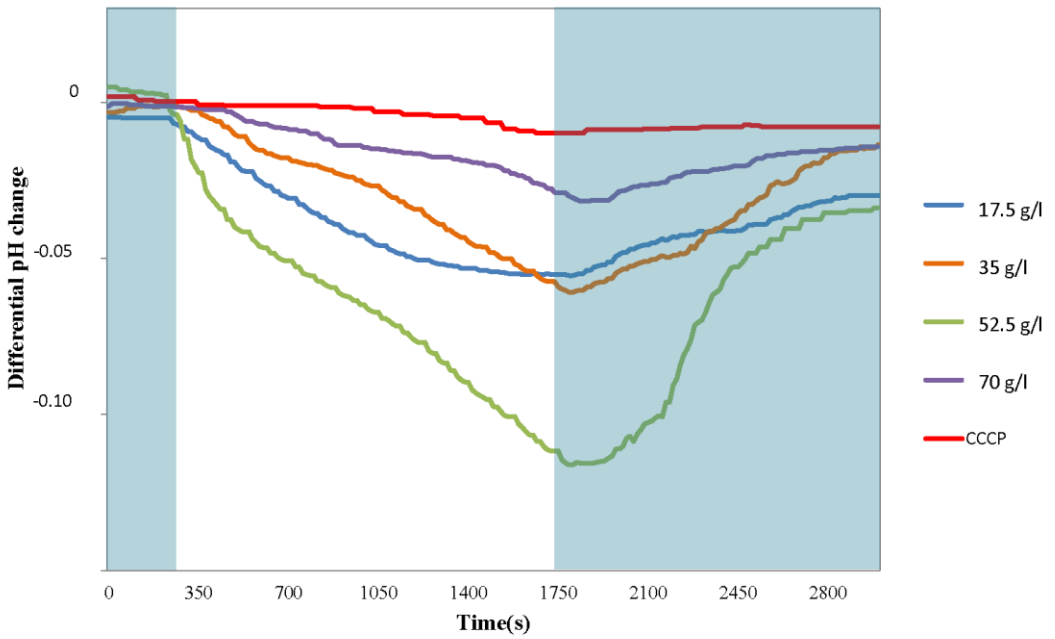


Figure 4 Light-driven proton pump activity in *P. torquis* cell suspensions. The pH changes in the suspension media were measured (initial pH: 7.7±0.5) under light (27.7 µmol photons m^{-2} s^{-1}). The shaded area denotes completely dark conditions during the experiment. The incubation temperature was 1(±0.5)ºC. CCCP was added at a final concentration of 10 µM.

In some studies the expression of PR genes is influenced by light/dark conditions (Gomez-Consarnau et al., 2007, Lami et al., 2009). The recent study performed by Akram et al (2013) suggested that PR expression can be regulated by nutrient limitation in *Vibrio* sp. AND4. However, other researchers reported in laboratory experiments that *in vitro* PR gene expression was constant regardless of illumination (Riedel et al., 2010). Our results also indicate that PR expression in *P. torquis* under different light, salinity and growth stages was expressed at a constant level (Table. 1). Our qPCR results suggested PR gene expression was not regulated either by light intensity or by salinity. High throughput gel-free label-free proteomic analysis

provided a new means to assess PR abundance and our 1D-LC/MS results provided strong evidence that PR abundance is affected by both salinity and illumination levels indicating that in *P. torquis* PR is likely to be post-transcriptionally regulated. However, responses are clearly complicated by interactions between salinity and illumination. Additionally, high levels of PR do not necessarily translate directly to higher cell abundance or growth rate.

Consistent with the protein abundance data, proton pumping activity in *P. torquis* was also strongly affected by salinity (Figure 4 & 5). The data suggests *P. torquis* responds to light more strongly under osmotic pressure compared to the optimum salinity. However, we suspect there are distinct threshold salinities driving this response. Both PR protein level and proton pump activity increased with increasing salinity peaking at 52.5 $g.L^{-1}$ then dropped sharply at a salinity of 70 $g.L^{-1}$. The variations in activity could be indicative of different energy demands at different salinities.

To be fully functioning, PR requires retinal, a chromophore that binds to the final membrane domain of PR (Béjà et al., 2000, Béjà et al., 2001). Retinal is derived from carotenoids, therefore PR is dependent on carotenoid synthesis (McCarren and DeLong, 2007). It can thus be proposed that the level of carotenoids and retinal derivatives can be an indirect indicator of PR abundance. For *P. torquis* the results clearly show that at all salt concentrations the production of carotenoids was higher under light compared to dark incubation conditions (Figure 3). The production of carotenoid was associated with both light and salinity levels and the effect of

illumination level were smaller than the effect of salinity. However, at 52.5 $g.L^{-1}$ light levels strongly affected the production of carotenoids (Figure 3c). Both light-driven proton pump activity and pigment data indicate that light has greater effect at 52.5 $g.L^{-1}$ than the other salinities tested and this corresponded to the large differences in PR abundances we observed using proteomic analysis (Figure 5).

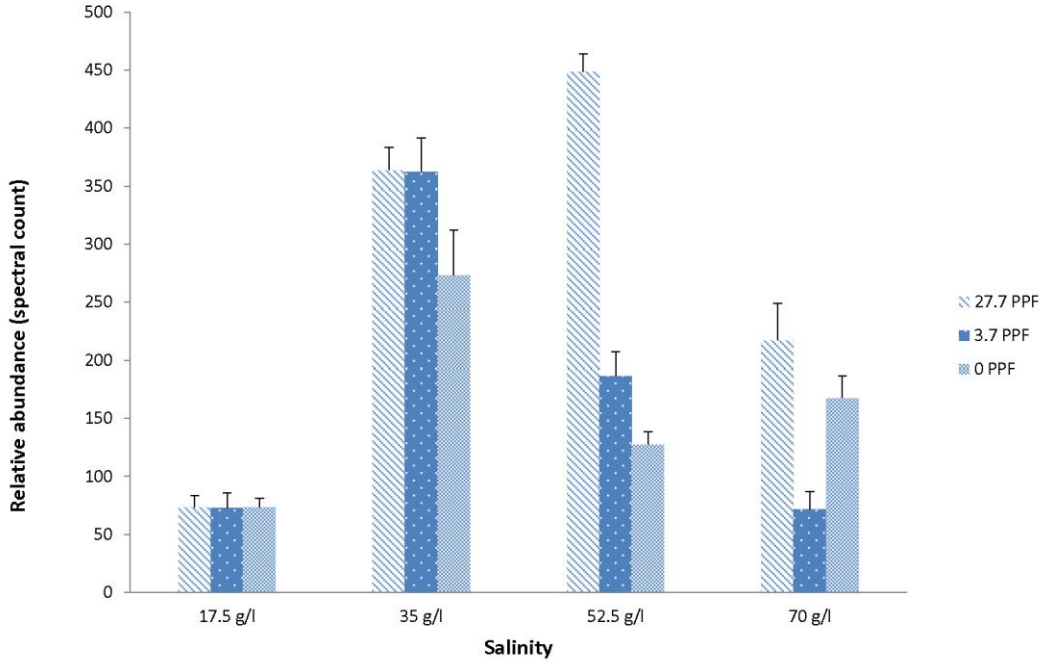


Figure 5 PR abundance under varying light and salinity levels. Normalised spectral count of PR protein detected by 1D LC/MS showed different level of PR protein at different salinity and light conditions (μmol photons m^{-2} s^{-1} [PPF]).

We theorise that PR may represent an adaptation strategy used by microorganisms to their own specific econiche. This study provides evidence that PR promotes the growth of *P. torquis* in conditions of different salinity. Our proteomic data shows that PR is regulated post-transcriptionally, influenced by both illumination and salinity. Thus, we have extended the existing paradigm that PR provides an alternative source

of energy during nutrient stress also include osmotic stress. Furthermore, as many of kinds of stress responses lead to diversion of energy from growth and biosynthesis to maintenance and survival functions, we further suggest that PR may have a broader effect when marine prokaryotes encounter pH, temperature, or other kinds of stress conditions.

Chapter 4

Light and salinity induced proteomic response in a proteorhodopsin-containing sea-ice dwelling flavobacteria

Abstract

Although much effort has been put into the investigation of the microbial ecology of sea ice ecosystems physiological studies of sea ice microbes have not been extensively reported. The extremely psychrophilic proteorhodopsin-containing bacterial species *Psychroflexus torquis* is considered a model sea-ice microorganism, which has adapted to an epiphytic lifestyle. The quantitative proteomic study performed here indicated *P. torquis* responds to changing salinity and illumination conditions by regulating its energy generation, nutrient uptake strategy, and adhesion ability. Furthermore, salinity and illumination intensity were observed to interact highly in affecting the abundance of nutrient uptake transporters, detoxification systems and gliding motility proteins. The data suggests that *P. torquis* responded to changes in both light energy and salinity to modulate membrane and central metabolic proteins which are involved in energy production as well as nutrient uptake and gliding motility processes that would be especially advantageous during the polar summer ice algal bloom. This study provides insights into the life strategy of a model sea-ice bacterial species and demonstrates the complex interrelation of physiological processes connected to proteorhodopsin.

Introduction

Sea ice is a highly variable and dynamic environment covering a maximum 13% of the Earth's surface area (Thomas and Dieckmann, 2002). Although sea ice is dominated by extreme gradients of temperature, salinity and light, sea-ice habitats are a major source of biomass in sub-polar and polar ocean regions (Brown and Bowman, 2001, Thomas and Dieckmann, 2002) and support a diverse microbial community of phytoplankton, bacteria, archaea and viruses. The sea ice microbial community is highly metabolically active and sea ice bacteria have higher metabolic activity than some offshore temperate marine systems. This suggests the potential for substantial biological production that can be exported in polar regions (Koh et al., 2010, Rivkin and Legendre, 2001). However, over the last decade, a severe loss of Arctic Ocean multi-year sea ice has occurred, and the effect of continuous global warming may result in sea ice ecosystems disappearing in many regions. So far, most research has focused on characterizing the overall sea ice microbial community and the factors that influence them (Bowman, 2013).

Here, we present a quantitative proteomic study on the proteorhodopsin-containing sea ice dwelling bacterial species, *Psychroflexus torquis*, using a gel-free label-free based approach. *P. torquis* is an extreme psychrophile belonging to the family *Flavobacteriaceae*, phylum *Bacteroidetes* (Bowman et al., 1998). Our genomic analysis of the *P. torquis* type strain revealed the presence of a proteorhodopsin (PR) gene and a cognate carotenoid monooxygenase involved in synthesis of its retinal co-factor (Feng et al., 2014). PR is a retinal binding transmembrane protein that is

found in aquatic bacteria and can act as a light-driven proton pump (Béjà et al., 2000). PR-bearing bacteria are widely distributed in marine ecosystems and also occur in sea ice microbial communities though the predominance in sea ice is so far unknown (Fuhrman et al., 2008, Koh et al., 2010). A significant amount of PR gene expression in sea-ice samples corresponding to the presence of metabolically active cells has been reported (Koh et al., 2010), which suggests that the PR gene is functional within sea ice. It is hypothesized that PR can provide light derived energy to bacteria under oligotrophic conditions. Another study showed that PR can function as an energy generating system replacing cellular respiration in *E. coli* (Walter et al., 2007). It was also shown that in *Candidatus "*Pelagibacter ubique" the ATP-dependent nutrient transporter activity increased in starved cells under light (Steindler et al., 2011). However, the available data does not consistently show light stimulated growth in PR-containing bacteria (Fuhrman et al., 2008). Thus more studies are needed to determine the ecological role and significance of PR in bacteria. We recently reported (Feng et al., 2013) that the growth of *P. torquis* is stimulated by light under osmotic pressure, and that the stimulatory effect and the PR protein abundance were regulated by salinity levels as well as by light intensity. This demonstrated the possibility that PR might be an adaptive strategy used by microorganisms under stress conditions associated with their specific econiche. However, the complete physiological role of PR-mediated proton transport in the overall cellular metabolism and bacterial growth remains unclear. To our best knowledge, no proteomic studies that describe the physiology and metabolism of PR-containing sea ice dwelling bacteria have been

performed. Our present proteome analysis describes the life style and responses of *P. torquis* to salinity and light providing insight into the role of PR in this microorganism in the sea-ice ecosystem.

Materials and Method

Bacterial Isolation

The bacterial strain used in this study was *Psychroflexus torquis* ATCC 700755^{T}. It was originally isolated from coastal sea ice from Prydz Bay, Vestfold Hills, Antarctica (Bowman et al., 1998).

Culture experiments

Cells were grown on modified marine agar consisting of 5g of proteose peptone (Oxoid), 2g of yeast extract (Oxoid), 15g of agar, and the relevant amount of sea salt (Red Sea) in 1L of distilled water. To achieve four different salinities, 17.5g, 35g (optimal salinity), 52.5g and 70g of sea salt were added per L to achieve the target salinities. For each salinity, cultures were exposed to three light conditions: high light (20-30 μmol photons m^{-2} s^{-1}); low light (3-4 μmol photons m^{-2} s^{-1}) and complete darkness (0 μmol photons m^{-2} s^{-1}). 100 μL of working culture (6.0 $\log_{10}$ CFU/ml) was used for spread plates. Cultures were incubated aerobically at 5 ºC for 3 weeks. The biomass was harvested from agar plates for protein extraction and digestion.

Protein Preparation and Proteomic Measurement

Preparation of protein fractions and explanation of the workflow leading to mass spectrometry was described previously (Al-Naseri et al., 2013, Feng et al., 2014). Full details are given in **Material and Method in Chapter 2.**

Statistical Analysis of Mass Spectrometry Data

The spectral count (SpC) output of the mass spectrometry data, filtered using Protein Prophet, was used to determine relative protein abundance. Fold changes in different protein expression levels were calculated based on the averaged SpC (for three biological replicates). The total spectral count normalization approach was used as described by Bowman et al (2012). Significant differences in protein SpC were normalized and tested using beta-binomial distribution analysis implemented in R as previously described (Al-Naseri et al., 2013, Bowman et al., 2012). CAP analysis (Anderson and Willis, 2003) of SpC data was carried out in Primer 6 (Primer-E, Plymouth, UK) with data transformed to the fourth root. The mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium (http://www.proteomexchange.org) via the PRIDE partner repository (Vizcaino JA, 2013) with the dataset identifier PXD000673

Examination of Gliding Motility

The gliding motility of *P. torquis* was examined under different salinity and illumination levels. Modified marine agar media with the different salinity levels

described above were prepared and sterilized glass slides were imbedded to form a thin layer of agar on top of the slides. 10 μL of working culture (6.0 $\log_{10}$ CFU/ml) was then inoculated on the glass slides and incubated aerobically at 5 ºC for 3 weeks under the illumination conditions described above. Slides were subsequently excised and examined via light microscopy.

Results and Discussion

Proteome Properties and Broad Changes

This study used *Psychroflexus torquis* ATCC 700755^T, which has a completely sequenced genome and contains 3572 predicted genes (Feng et al., 2014). It was originally isolated from Eastern Antarctic coastal sea ice but has a bipolar distribution (Sul et al., 2013). The biomass for proteomic analysis was harvested from agar due to difficulties of separating cells from liquid culture. This is due to the propensity of *P. torquis* ATCC 700755^T to produce prodigious amounts of exopolysaccharide resulting in cultures with substantially increased vsicosity which hindered sedimentation. Furthermore, EPS strongly reacted with the protein detection analysis. Finally, culturing on agar provided evenly distributed illumination conditions compared to broth culture. Analysis of the global proteomes of cultures grown under a range of salinity and illumination conditions produced approximately 750,000 tandem mass spectra (Appendix C. Table S1). A total of 1406 proteins passed identification filtration criteria. TonB-dependent transporters, amino acid metabolism related proteins, central metabolism related proteins, membrane bioenergetics-associated

proteins, and various uncharacterized proteins comprised the majority of high abundance proteins (Fig. 1).

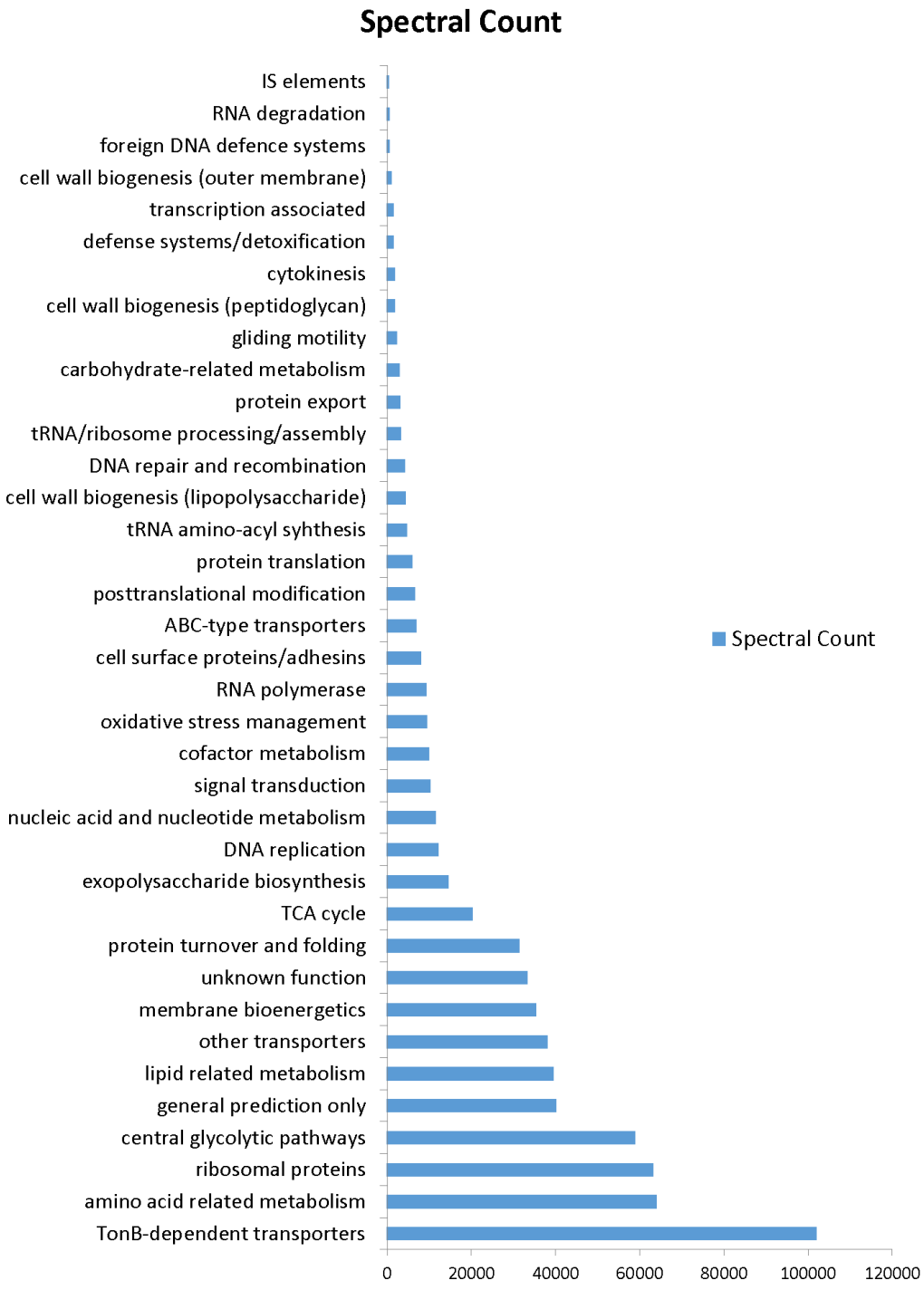


Figure 1 Spectral abundance (that passed filtration criteria) of *P. torquis* proteins observed in sets organized on the basis of functional class. These spectral counts are summed from all samples.

The spectral count data of cells grown under different conditions were normalized and compared using canonical analysis of principal coordinates (CAP). The CAP results (Fig. 2) indicate that the effect of salinity on the proteome is more profound than that induced by light intensity. Overall, sub-optimal salinity levels and high light intensity among all factors were the most clearly separated from other salinity and light conditions tested.

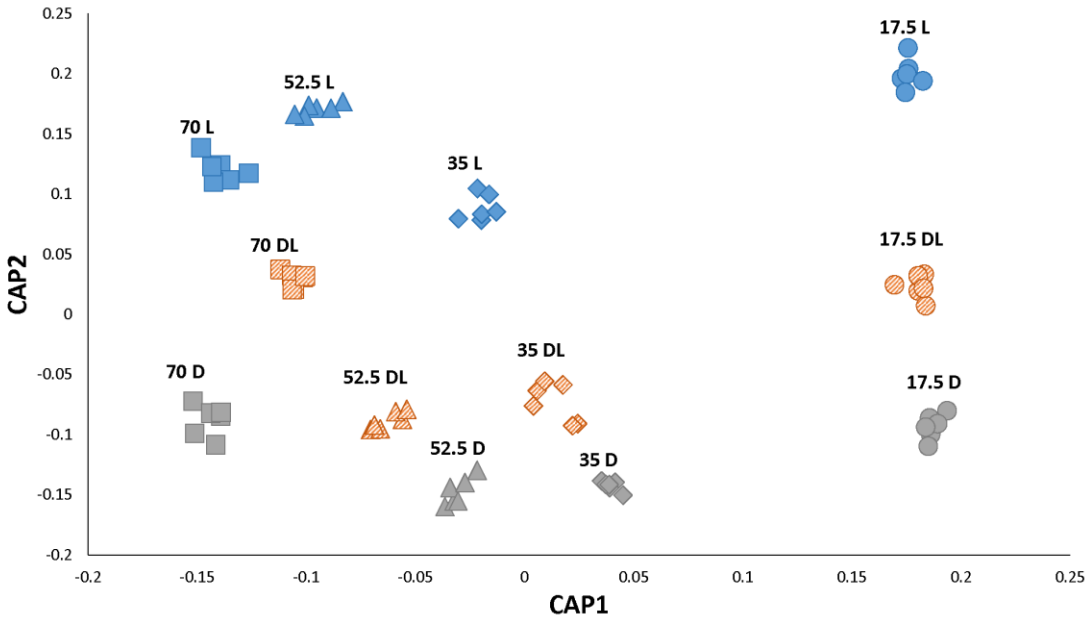


Figure 2 CAP analysis plot comparing the proteomes of *P. torquis* grown under different levels of salinity and light. Flitered spectral count data was fourth-root transformed Cultures were grown in high light 20-30 µmol photons m^{-2} s^{-1} (blue symbols), low light 3-4 µmol photons m^{-2} s^{-1} (diagonal symbols) and dark conditions (grey symbols) in salinities of 17.5g·L^{-1} (circles), 35g·L^{-1} (diamonds), 52.5g·L^{-1} (triangles) and 70g·L^{-1} (squares)

Salinity and Illumination have Different Effects on Membrane Bioenergetics

Sea-ice ecosystems are hotspots for extremophiles, in which strong variations of salinity and light under partially sub-zero temperatures are experienced. Salinity may be a major stressor for sea-ice organisms since they can experience salinities ranging from three times that of seawater to hyposaline conditions when the ice thaws (Thomas and Dieckmann, 2002). Consistent with this hypersaline environment, *P. torquis* has a Na^+-translocating NADH:quinone oxidoreductase (NQR) like other marine and halophilic bacteria (Gonzalez et al., 2008, Gonzalez et al., 2011, Unemoto and Hayashi, 1993). Our results showed that all NQR subunits were significantly down-regulated at a sea salt salinity of 17.5$g \cdot L^{-1}$ (sub-optimal for growth of *P. torquis*). As the salinity increased, the NQR level increased in abundance (Appendix C. Fig. S1).

P. torquis does not encode a bc_1 complex (known as complex III) in the respiratory chain. Instead, like other aerobic members of phylum *Bacteroidetes*, *P. torquis* encodes components of the alternative complex III (ACIII) that is postulated to function as a menaquinol:cytochrome *c* oxidoreductase (Gonzalez et al., 2011, McBride et al., 2009). However, based on structural analysis and the fact that no other proteins can replace its activity, ACIII is presumed to have proton pumping and electron transfer capabilities (Blankenship et al., 2012). Most proteins in the electron transfer chain were only reduced in abundance at low salinity (17.5$g \cdot L^{-1}$) levels, and did not show significant changes under the hypersaline conditions tested (Appendix C.

Fig S1). Interestingly, the overall PR level decreased significantly at 17.5g·L^{-1} compared to 35g·L^{-1} (4.5 fold change on average), and was also reduced at 52.5g·L^{-1} and 70g·L^{-1} by 1.3 fold and 2 fold respectively. This could be due to energy limitation under stressful conditions. The microorganism may potentially devote more effort to other essential life strategies such as energy generation, nutrient uptake, adhesion and gliding rather than produce PR in order to gain a survival advantage. Furthermore, high level of PR level does not necessarily translate to higher cell abundance (Feng et al., 2013). Therefore the regulation of PR level under different salinity is a complicated interaction between different physiological processes.

At salinities of 17.5g L^{-1} and 35g L^{-1} NQR subunits became more abundant under illumination, whereas, at other salinities, NQR levels remained unaffected by illumination. The effect of illumination on ACIII and complex IV was weak (Fig. 3, Appendix C. Table S2), however, the effect on the F_1F_0 ATP synthase complex was stronger. Two of its major subunits, AtpD and AtpE, were significantly down-regulated under light at 17.5g·L^{-1} (0.7 fold and 0.4 fold respectively, *p value* < 0.01). Both AtpB and AtpC were stimulated by light in 17.5g·L^{-1} and 52.5g·L^{-1} sea salt, but exhibited a relatively small change in 70g·L^{-1} (Appendix C. Table S2). In optimal salinity (35g·L^{-1}), the abundance of the ATPase complex was reduced under illumination. In particular, the gamma and delta subunits, which are involved in rotating the F_1-ATPase, were greatly reduced (Appendix C. Table S2). The results suggest that under illumination at optimal salinity, although the electron transport chain retained a higher potential for proton pumping compared to dark conditions,

ATP production by ATPase was limited. Overall, ATP synthase abundance was influenced by both light and salinity. This response may be the result of different proton gradient levels produced by PR and other proton transporters, and also by cellular energy requirements which may explain reduced ATPase subunit protein abundances under illumination at optimal salinity. The accumulated data implies that PR-photophysiology can potentially regulate Na^+-translocating NQR to adapt to different salinity conditions. Under specific levels of light intensity and salinity, NQR and PR proton transport activity likely becomes additive, resulting in growth stimulation (Steindler, et al 2011).

Hypersalinity and Illumination Leads to Increased Transporter Protein Abundances

Functional analysis of the proteome showed that amongst all functionally allied protein sets, TonB-dependent transporters (TBDTs) were the most abundant proteins (Fig. 1). TBDTs are known to utilize the proton motive force to transport nutrients including iron, copper, vitamin B_{12}, and complex carbohydrates across the outer membrane. They may also be involved in degradation and uptake of the products of hydrolytic enzymes (Blanvillain et al., 2007, Giovannoni and Stingl, 2007). It has been proposed that marine flavobacteria rely on proteins and carbohydrate polymers for growth and thus contribute to marine secondary production (Gonzalez et al., 2008). Since sea-ice phytoplankton are the primary producers and *P. torquis* has an epiphytic relationship with them (Bowman et al., 1998), it likely uses TBDTs to take up

nutrients and substrates directly derived from algal products. These high abundance TBDTs include several SusC-like carbohydrate uptake and SusD-like outer membrane proteins that putatively bind polysaccharides, similar to what has been surmised in the related species "*Gramella forsetii*" which also has many TBDT systems (Bauer et al., 2006). These highly abundant transporters were observed to increase in abundance under supra-optimal salinity conditions and decrease under sub-optimal salinity (Appendix C. Fig. S2, Table S2, S3). Increases in transporter abundance under supra-optimal salinity would assist cells under conditions where there is less availability of labile growth substrates due to lower levels of algal biomass.

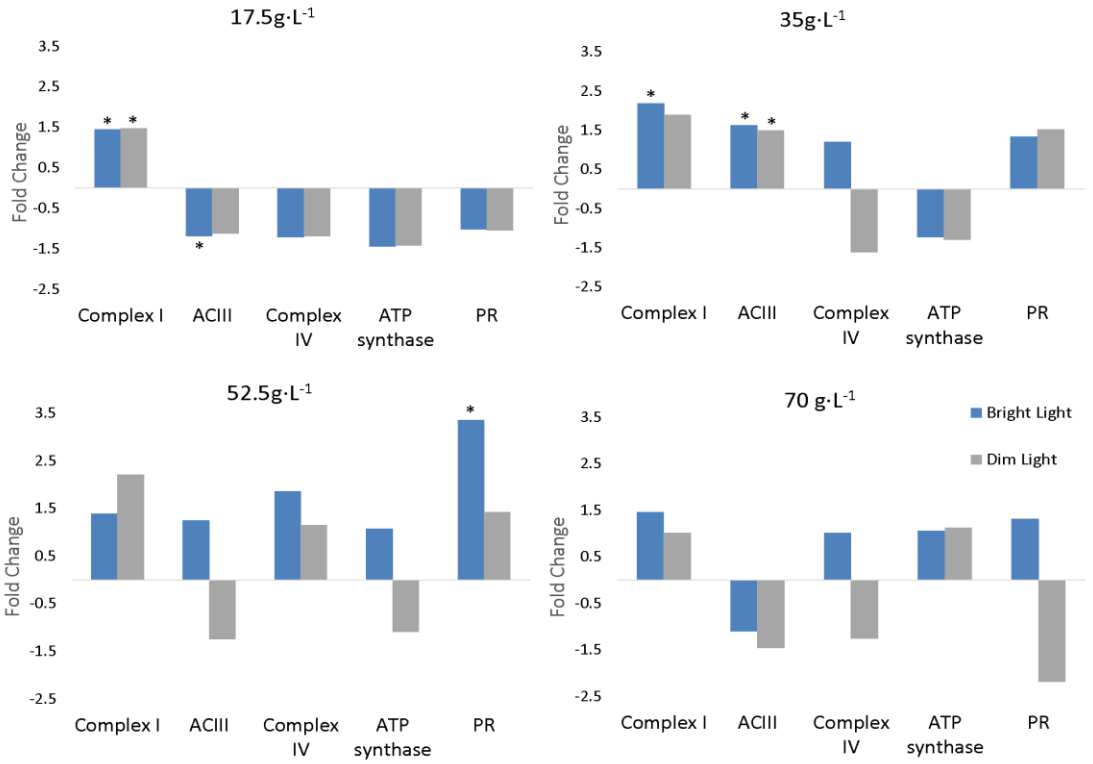


Figure 3 The fold change in relative abundances of proteins belonging to complexes making up the electron transfer chain in *P. torquis* where the fold change is the difference between high light (20-30 µmol photons m^{-2} s^{-1}) and dark (as control) and between low light (3-4 µmol photons m^{-2} s^{-1}) and dark (as control) at different salinity levels. Each bar represents the total abundance of a complex to its cognate complex labelled below it. Asterisks indicate *p* value of the fold change is less than 0.05.

The reduced abundance of transporters under sub-optimal salinity levels, could be a response to ice thawing during which nutrient availability spikes (Smith and Nelson, 1985, Thomas and Dieckmann, 2002). Reduced salinity also corresponded to a lower growth yield. Thus cells shift energy from transporter synthesis to other processes due to greater access to of growth substrates and less active anabolism. The high abundance of TBDTs is also consistent with the concept that enhanced mass transfer of nutrients helps overcome salinity and low temperature stresses (Junge et al., 2004). Proteomic data also revealed that several other transporters were significantly more abundant in cultures exposed to light. In particular 33% of detected efflux transporters, were more abundant under the combination of low salinity and illumination (Appendix C. Table S2, Fig. S2*B*). These efflux systems are presumed to export a wide range of toxic compounds from the cell (Zgurskaya and Nikaido, 1999), and have been found to result in multidrug resistance when they are overexpressed (Ma et al., 1994). Higher temperature and continuous light during summer in polar regions leads to sea ice melting and microbial blooms in sea ice zones. (Bowman, 2013). In competing for nutrients microalgae can secrete compounds that inhibit growth of other microorganisms (Pusceddu et al., 2009). This light stimulated overexpression of efflux transporters could potentially provide *P. torquis* with a survival/growth advantage by transporting noxious compounds out of the cell; however, further analysis is required to determine the exact roles of these proteins.

Glycolysis and TCA Cycle Enzyme Abundance is Stimulated under Salt Stress and Illuminated Conditions

The main energy generating metabolic pathways of *P. torquis* were identified as the glycolysis pathway and the tricarboxylic acid (TCA) cycle. In response to different salinity levels, several enzymes in glycolysis that are involved in energy production became more abundant. In particular, fructose-bisphosphate aldolase (EC 4.1.2.13), pyruvate kinase (EC 2.7.1.40) and enolase (EC 2.4.2.11) were induced under both sub- and supra-optimal salinities (Appendix C. Fig. S3). Glyceraldehyde-3-phosphate dehydrogenase (EC 1.2.1.12), which is involved in the reaction that produces two NADH and one H^+ molecules, was more abundant under sub-optimal salinity levels. Phosphoglycerate kinase (EC 2.7.2.3), which catalyzes ATP generation (to balance consumption of ATP in the pathway), was only induced under supra-optimal salinity levels. In contrast triosephosphate isomerase (EC 5.3.1.1) was less abundant under both sub- and supra-optimal salinity conditions. This may lead to reduced conversion of glyceraldehyde-3-phosphate to glycerone-phosphate, and to an increase in NADH production efficiency. Both pyruvate carboxylase (EC 6.4.1.1) and pyruvate dehydrogenase complex were more abundant at lower salinity, but were not affected by the high salinity levels (Appendix C. Fig S3, Table S3) suggesting a different emphasis in the synthesis of acetyl-CoA between the salinity extremes. The citrate synthase complex, which introduces acetyl-CoA into the TCA cycle, is slightly inhibited. However, the TCA cycle enzymes that are involved in ATP and NADH generation steps all became more abundant (Appendix C. Fig. S3). Previous growth

experiments (Feng et al., 2013) indicated both sub- and supra- salinity levels are stressful for *P. torquis*. Because cells were grown in a nutrient rich medium our findings suggest that changes in glycolysis and the TCA cycle protein abundances are indicative of general responses to salinity stress. The expression pattern of central metabolic pathways indicates a need to generate more energy to overcome the salinity stress. Since light mediates proton pumping via PR and energizes the electron transport chain, some energy requirements imposed by low or high salt stress are alleviated.

Since PR provides light derived ATP to supplement cell energy usage, central metabolic pathways were found to exhibit smaller changes in cells under different levels of illumination compared to cells at different salinity levels (Fig. 4, Appendix C. Fig. S3, Table. S2, S3). Nevertheless, illumination significantly increased the level of the energy generation related enzymes glyceraldehyde-3-phosphate dehydrogenase (EC 1.2.1.12), triosephosphate isomerase (EC 5.3.1.1) and phosphoglycerate kinase (EC2.7.2.3) in glycolysis; and fumarate hydratase (EC 4.2.1.2) in the TCA cycle under all salinity levels tested (Fig. 4). As well, oxoglutarate dehydrogenase SucB (EC 2.3.1.61), which catalyses the production of NADH, significantly increased 2.5-fold at 17.5g·L^{-1} under illumination. The increasing amount of NADH could potentially stimulate the membrane bioenergetics process. TBDTs rely on the proton motive force to import nutrients especially carbohydrates. Therefore, more carbon uptake is expected under illumination resulting in the observed stimulation of various glycolysis and TCA cycle enzymes. It is presumed that efficient harnessing of light

energy can assist biological processes that involve nutrient and energy strategies in *P. torquis*. This could be the basis for prior observations of growth stimulation by light under optimal salinity conditions (Feng et al., 2013).

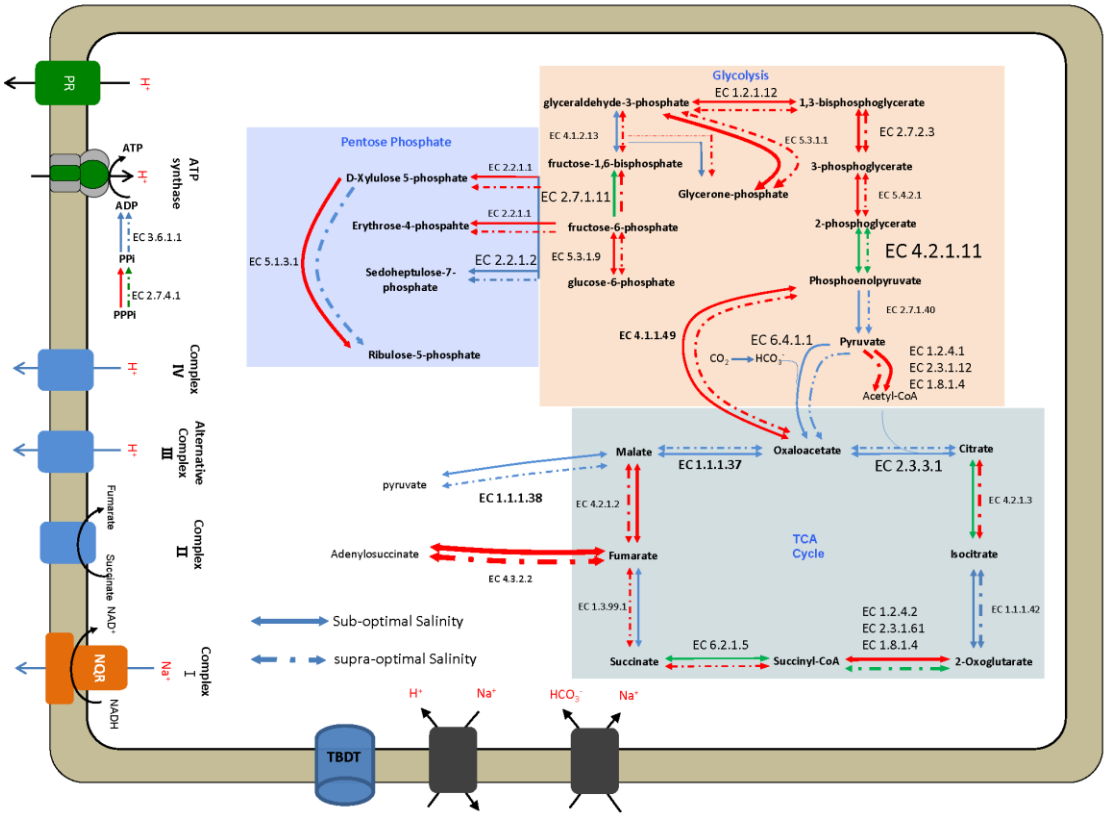


Figure 4 Diagram of central metabolic pathway changes that occur in *P. torquis* under illumination. The fold-change was determined using the dark incubation of each salinity as the control state. Each salinity level is shown separately. Within the metabolic map, red lines indicate an enzyme with greater abundance under light, green lines indicate lower abundance under light, and blue lines indicates an unchanged abundance level. The thickness of lines implies the degree of abundance change. The size of the EC enzyme code provides an indication of the overall abundance of the enzyme.

Certain Anaplerotic CO_2 Fixation Enzymes are More Abundant under Illuminated Conditions

Several proteins that are involved in anaplerotic CO_2 fixation (Tang et al., 2009) could be quantified, including pyruvate carboxylase (EC 6.4.1.1), phosphoenolpyruvate carboxykinase (EC 4.1.1.49), malate dehydrogenase (1.1.1.38) and adenylosuccinate lyase (EC 4.3.2.2), which can act to replenish the TCA cycle. Pyruvate carboxylase that generates oxaloacetate from HCO_3^- and pyruvate (Gonzalez et al., 2008) exhibited no salinity-dependent changes under illumination while its abundance was increased under sub-optimal salinity conditions (Appendix C. Table S2, S3). Phosphoenolpyruvate carboxykinase and adenylosuccinate lyase, which generate oxaloacetate and fumarate, increased 1.3-3.5 fold under illumination (Fig. 4, Appendix C. Table S2). In a previous study it was shown that *Polaribacter* sp. MED152 fixed more HCO_3^- under illuminated conditions (Gonzalez et al., 2008). A subsequent report on light-stimulated growth of *Dokdonia* sp. MED134, however showed significant down-regulation of pyruvate carboxylase expression under illuminated conditions (Kimura et al., 2011). These differences could be due to strain variation as well as poor correlations between changes in mRNA transcript and protein abundance (Kocharunchitt et al., 2012). In general, PR proton pumping is an independent process from organic matter respiration, however provision of supplementary ATP and reductant due to more active electron transport could enhance carbon flow through the TCA cycle, thus requiring increased uptake of CO_2.

Putative Light Sensing Proteins are Stimulated under Illuminated Conditions

The genomic analysis of *P. torquis* revealed that it harbors genes coding for the enzymes responsible for the production of carotenoids (*crtEBIY*) and retinal (*blh*) which are chromophores for PR (Feng et al., 2014). We observed an extremely low level of these proteins in our proteomic survey except for phytoene desaturase which was presented in relatively constant amounts under different growth conditions. Previous studies also demonstrated the presence of both carotenoids and retinal in *P. torquis* and their levels were affected by light intensity and salinity (Feng et al., 2013). We assume that these proteins were lost during protein extraction possibly due to complexes formed with carotenoid pigments (Holt and Krogmann, 1981) or they are expressed at very low levels.

P. torquis contains several regulatory systems that potentially sense light and invoke regulatory responses. These include several proteins with PAS and GAF domains known to be common components of phytochromes. PAS domains are reported to respond to cellular energy levels, oxygen levels, redox potential and light, whereas GAF domains often work as phototransducers (Anantharaman et al., 2001, Gonzalez et al., 2008, Taylor and Zhulin, 1999). Another gene related to light sensing containing the BLUF domain that specifically responds to blue light was also found (Gomelsky and Klug, 2002). Like other PR containing marine bacteria the genome also encodes deoxyribodipyrimidine photolyases and DASH family cryptochromes coding genes (Gonzalez et al., 2008, Gonzalez et al., 2011). In addition to light

sensors, several histidine kinase and two-component system sensors serve as a basic environmental stimulation response coupling mechanism that helps organisms to sense and respond to the environment (Stock et al., 2000). In this study, most proteins in the signal transduction group were only detected at low abundances. Nevertheless, our data is sufficiently sensitive to reveal PAS and GAF domain protein levels are always higher when cells were grown exposed to light, additionally these signal transduction systems were also influenced by salinity. In general, these proteins are less abundant under optimal salinity levels (Fig. 5). BLUF domain, cryptochromes and deoxyribodipyrimidine photolyase proteins were detected only at low to very low abundance levels under all conditions suggesting that these proteins do not couple with the responses linked to activities associated with the green-light absorbing PR that *P. torquis* contains. Some secondary transduction enzymes including two-component histidine kinases showed increasing expression levels under illumination and have the same trend of response to salinity as the PAS and GAF domain-containing proteins (Fig. 5). These sensory systems may therefore also be involved in illumination-influenced regulation of gene expression (Swartz et al., 2007).

P. torquis Photo-inhibition under High Light Levels is Linked to Oxidative Stress and Lower Transporter Abundance

Illumination experiments were conducted at two levels (20-30 μmol photons m^{-2} s^{-1} and 3-4 μmol photons m^{-2} s^{-1}). Previous studies have reported that high light levels

can inhibit bacterial activity and this is also the case for *P. torquis* (Feng et al., 2013, Straza and Kirchman, 2011). The consensus is that strong and continual light exposure creates photooxidative stress; however, this stress is unlikely to be derived from UV irradiation-related damage since the light source used had low UV energy emittance levels and the petri dish cover on cultures would have further impeded UV light exposure. The proteome data suggests that under higher level of light, the protein abundance changes in the central metabolic pathway and membrane bioenergetics pathway are generally more substantial (Appendix C. Table S2). In addition, the antioxidant enzyme peroxiredoxin was induced by light, implying that *P. torquis* has a mechanism that responds to photooxidative stress. Nevertheless, some TBDTs were less abundant under high illumination levels (Appendix C. Fig. S2B). This reduction could partly explain light inhibition of growth since chemoheterotrophic *P. torquis* relies on uptake of nutrients via TBDTs.

Illumination and Sub-Optimal Salinity Promotes the Abundance of Adhesion Proteins

P. torquis harbors several types of putative adhesion proteins such as fasciclin repeat domain-containing proteins and putative ice-binding adhesion-like proteins (Feng et al., 2014). These are potentially associated with adhesion to algal and ice surfaces aiding the colonization of living phytoplankton cells and recruitment into forming sea-ice (Krembs and Deming, 2008). We found two fasciclin domain containing proteins that exhibited 20-fold and 6-fold increases under illumination at sub-optimal

salinity, and a Por secretion system protein (Sato et al., 2010) that was 15-fold more abundant under the same conditions, yet they all have smaller scales of stimulation at higher salinities. This finding provides a hypothesis that adhesion proteins are involved in maintaining phytoplankton colonization at the time of sea-ice thaw and formation (Bowman, 2013). These associations are likely to be especially important during and after the sea-ice microbial community is dispersed into the water column.

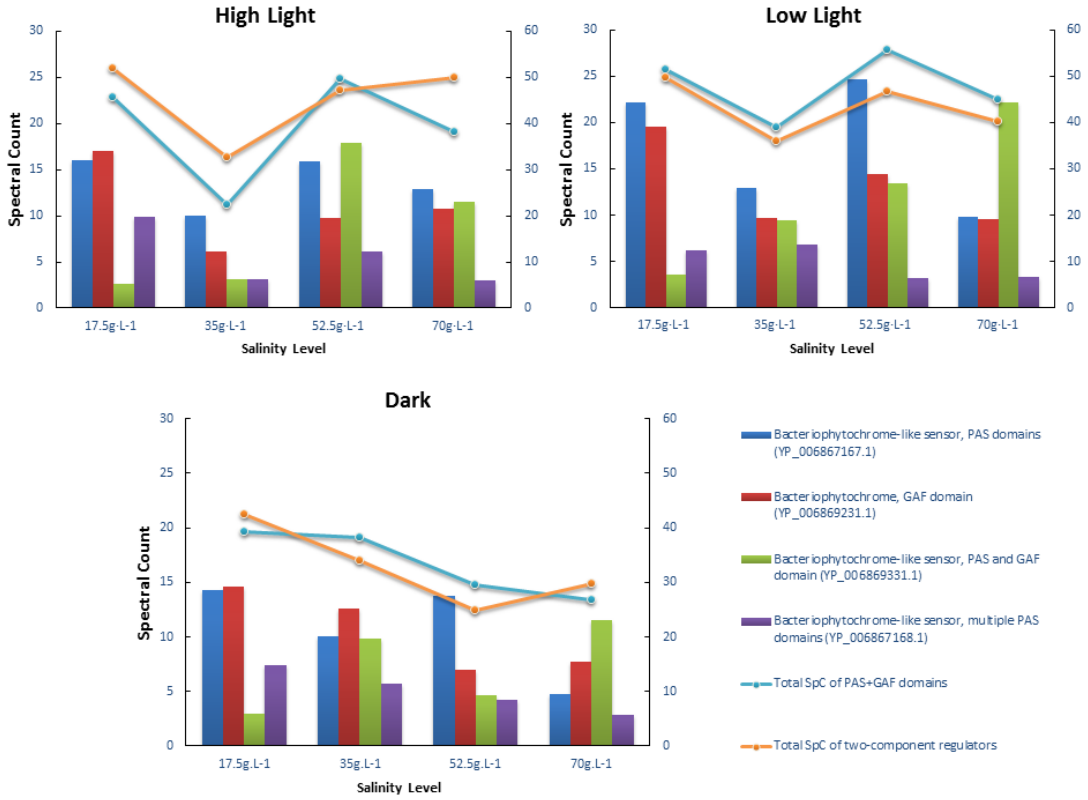


Figure 5 Comparison of total abundance of light sensing protein between different salinity and light conditions. As well as detail of each light sensing protein cross different salinity of light levels. Illumination intensity is: high light (20-30 μmol photons m^{-2} s^{-1}); low light (3-4 μmol photons m^{-2} s^{-1}) and complete darkness (0 μmol photons m^{-2} s^{-1}).

Dark Conditions Promote Gliding whilst Illumination Suppresses Gliding

Previous studies have shown that after transferring PR into *E. coli* illumination can increase the swimming motility of *E. coli* (Béjà et al., 2000, Walter et al., 2007). In our proteomic data four proteins critical for efficient gliding motility (GldJ, GldL, GldM and GldN) (Braun et al., 2005) were abundant and could be readily measured. Illumination did not affect the abundance of these proteins at the optimal growth salinity, but illumination suppressed the level of GldJ and GldN by 0.3-fold and 0.7-fold respectively at sub-optimal salinity and stimulated GldJ's abundance 4-fold at the highest salinity. GldM was increased under illumination 2.5-fold at sub-optimal salinity but did not change at other salinity levels (Appendix C. Table S2). Our direct observation of *P. torquis*' gliding motility on agar showed obvious increasing motility under dark at salinities of 17.5g·L^{-1}, 35g·L^{-1} and 70g·L^{-1}. However, the motility was largely suppressed by illumination at 17.5g·L^{-1} and 35g·L^{-1} while some gliding ability was retained at 52.5g·L^{-1} and 70g·L^{-1}. The overall effect of salinity on gilding motility is smaller than the effect of light (Fig. 6) indicating overexpression of GldM may inhibit gliding and GldJ and GldN are potentially important for gliding in *P. torquis*. Thus, darkness may induce the motility of *P. torquis* potentially aiding nutrient uptake during the polar winter when algal primary production is negligible. It is reasonable to expect there would be greater gliding motility in the natural sea ice environment, since the experiment was conducted on nutrient rich agar where as in natural sea ice nutrient levels are lower.

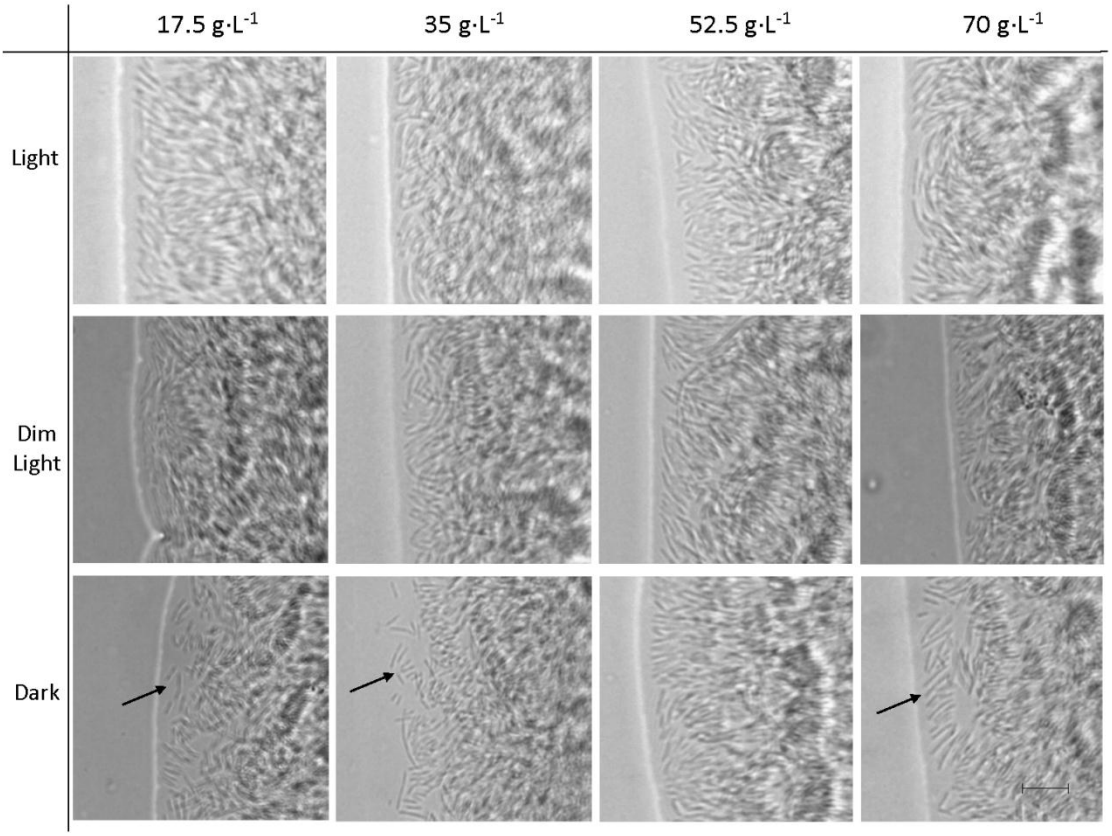


Figure 6 Gliding motility of *P. torquis* under different salinity and illumination levels. Light intensities were: high light (20-30 µmol photons m^{-2} s^{-1}); low light (3-4 µmol photons m^{-2} s^{-1}) and complete darkness (0 µmol photons m^{-2} s^{-1}). Arrows point to cells gilding away from the colony. The scale bar indicates 100 µm.

Exopolysaccharides (EPS) Formation is Salinity Regulated

EPS is one of the major outputs of bacterial secondary production in sea ice communities and contributes to the overall carbon cycle in the Antarctic environment (Nichols et al., 2005). EPS is important for sea ice inhabitation providing a structural network for microbial associations and acting as nutrient ligands for growth at low temperature and high salinity (Nichols et al., 2005). However, there are few reports on the mechanism of the effect of EPS on bacterial tolerance to high salinity, such as that found in the brine channels of sea ice. In this study we found the production of EPS in

P. torquis was clearly salinity regulated. Compared to optimal salinity both low and high levels of salinity induced the enzymes involved in EPS biosynthesis as well as some lipopolysaccharide biosynthesis proteins (Appendix C. Table S3). A recent study has shown that EPS from a sea ice diatom can alter the physical properties of sea ice resulting in the retention of more salt in the sea ice (Krembs et al., 2011). Secretion of EPS by *P. torquis* could theoretically act to exclude salt to achieve optimal salinity in the direct vicinity of cells. This could explain why EPS production and the associated enzymes were slightly increased under hypersaline conditions.

Conclusions

The proteome analysis of *P. torquis* revealed a number of proteins that changed in response to salinity and illumination. These included proteins associated with energy generation, adhesion, gliding motility, transporters and predicted light sensors that were correlated with salinity and light induced effects on growth rate and cell yields. We propose that the physiological role of proteorhodopsin may not involve it acting simply as a proton pump. It has complex interrelation with other physiological processes, and needs to be more extensively studied. The inferred protein abundance data can be related to the capability of *P. torquis* to adjust its energy production and nutrient uptake strategy to gain a survival advantage within a highly dynamic sea-ice ecosystem leading to an apparent interaction between the effects of salinity and exposure to light. The protein expression patterns of *P. torquis* shows it can use light to provide an advantage in carrying out colonization of phytoplankton; to take up

more nutrients; to optimize energy production; and to stay within or even modify the sea-ice habitat. Analysis of the proteome of *P. torquis* provides insights into its life strategies and survival and provides a basis for the study of other sea-ice associated microorganisms. Finally, the data also suggests similar types of environmentally and physiologically connected responses may occur in other PR-containing microorganisms.

Chapter 5

General Conclusion

The focus of the research presented in this thesis was a model sea-ice bacterial species that possesses PR. *Psychroflexus torquis* is an extreme psychrophile that dwells as an algal epiphyte and has peculiar traits that suggests it has become highly adapted to sea-ice. The primary goal of this thesis was to examine the sea-ice-associated lifestyle and the role of PR within the *P. torquis* type strain ATCC 700755^{T}. This commenced by analysing the whole genome of *P. torquis* to gain an understanding of how it is adapted to the highly dynamic sea-ice environment and how it may have acquired traits that provide this capacity. This was done by comparing the complete genome of ATCC 700755^{T}, generated in this study, to the draft genome of its closely related non-psychrophilic sister species, *Psychroflexus gondwanensis* ACAM 44^{T}, which dwells in hypersaline lakes in Antarctica. Comparison of the genomes of *P. torquis* and *P. gondwanensis* showed that horizontal gene transfer has occurred much more extensively in *P. torquis*. Approximately 35% of *P. torquis*' genome includes regions putatively subjected to HGT. Several genes in these regions can be linked to the sea ice lifestyle such as exopolysacchardes and polyunsaturated fatty acid biosynthesis, putative ice-binding proteins and light sensing proteins. However, most of these traits are absent in the closely related sister species, *P. gondwanensis* indicating that *P. torquis*' adaptation to sea-ice environment is driven by HGT which may suggest that sea ice is a hotspot for HGT. Therefore high levels of HGT should be expected in

other sea ice associated bacteria. Genome sequence analysis of the xanthorhodopsin-containing sea-ice *Alphaproteobacteria Octadecabacter arcticus* and *Octadecabacter antarcticus* (Vollmer et al. 2013) also suggest high levels of HGT since the genomes of these species show high levels of plasticity, possessing low levels of synteny, and are very rich in insertional elements (Vollmers et al. 2013).

Further proteome analysis of *P. torquis* showed that several genes associated with these features were highly translated, especially those responsible for exopolysaccharides and polyunsaturated fatty acid production providing evidence of the importance of these traits to *P. torquis*. This study suggests *P. torquis* is an excellent model to study sea-ice functional biology and evolutionary processes.

Analysis of the *P. torquis* genome revealed several genes that can provide *P. torquis* competitive advantage in the complex sea-ice ecosystem, however, the exact function of many is unknown. These include for example genes coding putative ice-binding proteins and metacaspase family proteins. Hypothetically these less characterized genes might be important for existence with algal partners in mutually beneficial or competitive arrangements as well as for persistence in sea-ice, for example attachment to ice crystals (via ice-binding proteins).

The Antarctic microbial ecosystem provides useful models to answer questions in evolutionary ecology due to the relative isolation of the continent. Given that Antarctica is open to invasion from elsewhere, scientists suggest that newly arrived microbes could evolve within this region to eventually become new species, or new arrivals and gene transfer from lower latitudes could constantly displace these nascent

endemic microbes. Our study provided direct evidence regarding this matter in that *P. torquis* appears to represent a taxon adapted to the specific sea-ice niche while being related to other species that have evolved successfully to inhabit other ecosystems with hypersaline characteristics including salt lakes, solar salterns, and presumably via use of brine or salt from these locations spread to foods with low water activity (i.e. *Psychroflexus halocasei* isolated from semi-hard cheese) (Seiler et al. 2012). Future studies on more microbes, especially in ancient microbial ecosystems, will help develop general models in evolutionary ecology that describe what occurs under severe environmental pressures. Studies at the molecular level may also provide insights not only about genetic similarities between Arctic and Antarctic microbiota, but also about the evolutionary ecology of microorganisms throughout the global biosphere.

Since the effect of sea-ice thaw (change salinity levels), changes in algal productivity and the seasonal changes in solar insolation results in huge yearly variations in parameters all which may affect bacterial growth rates and yields. Thus the degree PR plays in the physiology of *P. torquis* was investigated utilizing conditions relevant to sea-ice including different nutrient levels, salinity, and illumination levels. Conventional microbiology and molecular biology techniques were used to characterize and infer the physiological responses of *P. torquis* ATCC 700755^{T} to varying light and salinity levels. During this phase of study, the physiological role of PR in *P. torquis* was examined (Chapter 3). PR genes are widely distributed and are among the most highly expressed proteins in marine bacterial communities (Sabehi et

al., 2004, Venter et al., 2004, Giovannoni et al., 2005, Frigaard et al., 2006, Atamna-Ismaeel et al., 2008, Koh et al., 2010). It has been demonstrated that bacteria can use light energy to supplement energy via PR when under oligotrophic conditions (Béjà et al., 2000, Gomez-Consarnau et al., 2007, Walter et al., 2007, Gomez-Consarnau et al., 2010). However, this does not explain why PR is also abundant in some comparatively nutrient rich ecosystems such as the sea ice environment. Therefore we hypothesized that PR may provide an adaptation strategy used by microorganisms for specific econiches.

A set of experiments were carried out to characterize the growth of PR-containing *P. torquis* as well as the PR gene expression and PR protein abundance under light and different salinity levels. *P. torquis* showed light-stimulated growth under salinity stress conditions. The maximum light-stimulated growth occurred at 52.5 $g.L^{-1}$ sea salt and the minimum effect occurred at optimal salinity (35 $g.L^{-1}$). Although PR gene expression was consistent at all salinity levels, PR protein levels and proton pump activities peaked under light at salinity of 52.5 $g.L^{-1}$. This suggests that the PR protein levels in *P. torquis* is post-transcriptionally regulated by both light and salinity. This study extends the current hypothesis that light provides energy for marine prokaryotes under stress conditions by providing evidence that light can also provide energy to assist prokaryotic growth under salinity stress, therefore changing the perspective of PR associated photophysiology. Further studies can build upon the findings of this chapter to explore the complete ecological role of PR. Although the numbers of culturable PR-containing microorganisms are limited, a surprisingly high proportion

come from Antarctic ecosystems (e.g. *Glaciecola*, flavobacteria). Other microbial rhodopsins exist such as bacteriorhodopsins, halorhodopsins, xanthorhodopsins and actinorhodopsins that can help halophilic archaea to adapt high salinity. Therefore a greater array of conditions needs to be tested for PR- and other rhodopsin family-containing microorganisms, especially those stress factors important within a specific environment, in order to understand the complete physiological role of PR and related proteins. Since PR is a light driven proton pump that is widely distributed in marine microbes, there is a possibility that PR was used to adapt to a possible ancient acid ocean environment (Whitfield 2009). This might explain why most PR-containing strains tested under current ocean pH do not show clear growth or survival response to the changes of illumination. Other possibilities of why PR does not provide a growth or survival advantage to some bacteria might be the result of excessive light levels used in experiments. Excessive levels of irradiance may have a negative effect on growth due to photooxidative stress or other mechanisms. Thus light intensity and its application might be a new direction to further explore the function of PR. Fundamentally, we still cannot completely reproduce the natural environment, therefore it is difficult to understand the real role of PR by only using culture based methods thus new approaches must be conceived. The development of metagenomic and more advanced proteomic techniques provides the possibility to further examine the genes and proteins that surround the function of PRs. Understanding the roles of PR might help us predict some potentially important global change impacts, for example stemming from ocean acidification, on ocean carbon cycles, marine

microbial ecology, energy production, and community structure changes.

The final phase of this thesis (Chapter 4) investigated the global proteomic response of *P. torquis* to different levels of light and salinity via nano-LC MS/MS in order to better define how these factors influence the physiology of *P. torquis.* This chapter aimed to provide details of how proteins involved in membrane and central metabolic pathways, nutrient uptake and gliding motility processes might be advantageous during the polar summer ice algal bloom.

Gel-free proteomics served as the primary analytical tool in this study. Proteomics can potentially provide direct information about cellular functionality, physiological processes, compared to genome expression prediction and transcriptomic analysis. However, 1D- LC-MS/MS based proteomics still has its limits. First of all, so far there is no existing method to comprehensively obtain all protein from cells in a single extract; cytosolic proteins predominate while membrane proteins are under-sampled. Secondly, loss of protein can occur during protein extraction and digestion. In this study the presence of EPS and carotenoids was a problem and in order to obtain analyzable extracts, some loss of proteins associated with these fractions of the cell was observed. Thirdly, due to the nature of the instrumentation used, peptide detection coverage has limits. This is mainly the result of the huge dynamic range in protein concentrations in the cell, in the order of a million-fold variation, thus low abundance proteins are not detected. Therefore the proteome obtained here was not complete. To gain more information, future studies should combine several techniques together such as proteomic, transcriptomics and

metabolomics in effect a “systems biology” approach to complement each other and provide greater detail and to better verify findings. It can be assumed that instrument and bioinformatic technology will improve to provide greater coverage of proteomes and absolute quantification.

Despite these limitations, the results of this study show that salinity and illumination have quantifiably different effects on membrane bioenergetics proteins, transporters, central metabolic pathways, adhesion proteins and gliding associated proteins. Protein data was used to infer physiological and metabolic aspects relevant to *P. torquis'* capacity to adapt to the highly dynamic sea-ice environment. The data revealed the strategies that this microorganism may use to survive during winter and thrive in summer. Proteomic analysis demonstrated that *P. torquis* can regulate transporters and gliding associated proteins at different salinities. This tendency may allow it to compete for nutrients at different times as conditions change. Indeed the overall data revealed that the proteome was highly dynamic and that *P torquis* is capable of altering its physiology within the defined set of physiological boundaries that it can grow at. Overall, this phase of the study suggests that environmentally stimulated physiological responses could also occur in other PR-containing microorganisms within their own specific econiches, which suggest PR can be a general light-driven stress adaptation system for survival in cold environments. Therefore, study of more bacteria in a similar way may provide a broader understanding of how bacteria dynamically adjust to survive in natural ecosystems. Future studies should focus on the interactions of sea ice microorganisms with other bacteria or algae using

metagenomic, metatranscriptomic and metaproteomic approaches to provide a clearer picture of the highly dynamic sea-ice ecosystem.

References

Akram N, Palovaara J, Forsberg J, Lindh MV, Milton DL, Luo H et al. (2013). Regulation of proteorhodopsin gene expression by nutrient limitation in the marine bacterium *Vibrio* sp. AND4. *Environ Microbiol.* **15**: 1400-1415

Albers SV, Van de Vossenberg J, Driessen AJM, Konings WN. (2001) Bioenergetics and solute uptake under extreme conditions. *Extremophiles.* **5**: 285-294.

Al-Naseri A, Bowman JP, Wilson R, Nilsson RE, Britz ML. (2013). Impact of lactose starvation on the physiology of *Lactobacillus casei* GCRL163 in the presence or absence of Tween 80. *J Proteome Res.* **12**: 5313-5322.

Anantharaman V, Koonin EV, Aravind L. (2001). Regulatory potential, phyletic distribution and evolution of ancient, intracellular small-molecule-binding domains. *J Mol Biol.* **307**: 1271-1292.

Anderson MJ, Willis TJ. (2003). Canonical analysis of principal coordinates: A useful method of constrained ordination for ecology. *Ecology.* **84**: 511-525.

Arrigo KR, Worthen DL, Lizotte MP, Dixon P, Dieckmann GS. (1997). Primary production in Antarctic sea ice. *Science.* **276**: 394-397

Atamna-Ismaeel N, Sabehi G, Sharon I, Witzel KP, Labrenz M, Jurgens K et al. (2008). Widespread distribution of proteorhodopsins in freshwater and brackish ecosystems. *ISME J.* **2**: 656-662.

Bauer M, Kube M, Teeling H, Richter M, Lombardot T, Allers E et al. (2006). Whole genome analysis of the marine Bacteroidetes '*Gramella forsetii*' reveals adaptations to degradation of polymeric organic matter. *Environ Microbiol.* **8**:

2201-2213.

Bayer-Giraldi M, Uhlig C, John U, Mock T, Valentin K. (2010). Antifreeze proteins in polar sea ice diatoms: diversity and gene expression in the genus *Fragilariopsis*. *Environ Microbiol*. **12**:1041-1052.

Béjà O, Arvind L, Koonin EV, Suzuki MT, Hadd A, Nguyen LP et al. (2000). Bacterial rhodopsin: Evidence for a new type of phototrophy in the sea. *Science.* **289**: 1902-1906.

Béjà O, Spudich EN, Spudich JL, Leclerc M, DeLong EF. (2001). Proteorhodopsin phototrophy in the ocean. *Nature.* **411**: 786-789.

Berge J, Varpe Ø, Moline MA, Wold A, Renaud PE, Daase M, Falk-Petersen S. (2012). Retention of ice-associated amphipods: possible consequences for an ice-free Arctic Ocean. *Biol Lett*. **8**: 1012-1015.

Bielawski JP, Dunn KA, Sabehi G, Béjà O. (2004) Darwinian adaptation of proteorhodopsin to different light intensities in the marine environment. *Proc Natl Acad Sci USA.* **101**: 14824-14829

Bland C, Ramsey TL, Sabree F, Lowe M, Brown K, Kyrpides NC, Hugenholtz P. (2007). CRISPR recognition tool (CRT): a tool for automatic detection of clustered regularly interspaced palindromic repeats. *BMC Bioinformat*. **8**: 209.

Blankenship RE, Majumder EW, King JD. (2012). Alternative Complex III from phototrophic bacteria and its electron acceptor auracyanin. *BBA. Bioenergetics.* **1817**: S29-S29.

Blanvillain S, Meyer D, Boulanger A, Lautier M, Guynet C, Denance N et al. (2007). Plant carbohydrate scavenging through TonB-dependent receptors: A feature shared by phytopathogenic and aquatic bacteria. *PLoS One.* **2**: e224

Bluhm, BA, Gebruk AV, Gradinger R, Hopcroft RR, Huettmann, F, Kosobokova KN et al. (2011). Arctic marine biodiversity: An update of species richness and examples of biodiversity change. *Oceanogr*. **24**: 232–248.

Borg J, Campos A, Diema C, Omeñaca N, de Oliveira E, Guinovart J, Vilaseca M. (2011). Spectral counting assessment of protein dynamic range in cerebrospinal fluid following depletion with plasma-designed immunoaffinity columns. *Clin Proteomics*. **8**: 1-15.

Bowman JP. (2008). Genomic analysis of psychrophilic prokaryotes. In: Margesin R., Schinner F, Marx JC, Gerday C, editors. *Psychrophiles: From Biodiversity to Biotechnology*. Springer: New York. p. 265-284.

Bowman JP, Hages E, Nilsson RE, Kocharunchitt C, Ross T. (2012). Investigation of the *Listeria monocytogenes* Scott A acid tolerance response and associated physiological and phenotypic features via whole proteome analysis. *J Proteome Res.* **11**: 2409-2426.

Bowman JP, McCammon SA, Brown MV, Nichols DS, McMeekin, TA. (1997). Diversity and association of psychrophilic bacteria in Antarctic sea ice. *Appl Environ Microbiol.* **63**: 3068-3078.

Bowman JP, McCammon SA, Lewis T, Skerratt JH, Brown JL, Nichols DS et al. (1998). *Psychroflexus torquis* gen. nov., sp. nov., a psychrophilic species from

Antarctic sea-ice, and reclassification of *Flavobacterium gondwanense* (Dobson et al. 1993) as *Psychroflexus gondwanense* gen. nov., comb. nov. *Microbiol.* **144**: 1601-1609.

Bowman JS, Rasmussen S, Blom N, Deming JW, Rysgaard S, Sichertz-Ponten T. (2012). Microbial community structure of Arctic multiyear sea ice and surface seawater by 454 sequencing of the 16S rRNA gene. *ISME J.* **6**: 11-20

Braun TF, Khubbar MK, Saffarini DA, Mcbride MJ. (2005). *Flavobacterium johnsoniae* gliding motility genes identified by mariner mutagenesis. *J Bacteriol.* **187**: 6943-6952.

Brinkmeyer, R, Knittel Katrin, Jürgens J, Weyland H, Amann R, Helmke E. (2003). Diversity and structure of bacterial communities in Arctic versus Antarctic pack ice. *Appl Environ Microbiol*. **69**: 6610-6619

Brown MV Bownman JP. (2001). A molecular phylogenetic survey of sea-ice microbial communities (SIMCO). *FEMS Microbiol Ecol.* **35**: 267-275.

Burke GR, Moran NA. 2011. Massive genomic decay in *Serratia symbiotica*, a recently evolved symbiont of aphids. *Genome Biol Evol*. **3**: 195-208.

Campbell BJ, Waidner LA, Cottrell MT, Kirchman DL. (2008). Abundant proteorhodopsin genes in the North Atlantic Ocean. *Environ Microbiol.* **10**: 99-109.

Carver T, Harris SR, Berriman M, Parkhill J, McQuillan JA. (2012). Artemis: an integrated platform for visualization and analysis of high-throughput sequence-based experimental data. *Bioinformat*. **28**: 464-469.

Carver T, Thomson N, Bleasby A, Berriman M, Parkhill J. (2009). DNAPlotter: circular and linear interactive genome visualization. *Bioinformat*. **25**: 119-120.

Choi CJ, Berges JA. (2013). New types of metacaspases in phytoplankton reveal diverse origins of cell death proteases. *Cell Death Dis*. **4**: e490

Church, MJ, Duchklow, HW, Karl, DA. (2004). Light dependence of [H-3]leucine incorporation in the oligotrophic North Pacific ocean. *Appl Environ Microbiol*. **70**: 4079-4087.

Church, MJ., Hutchins, DA. Ducklow, HW. (2000). Limitation of bacterial growth by dissolved organic matter and iron in the Southern ocean. *Appl Environ Microbiol*. **66**: 455-466.

Collins RE, Deming JW. (2013). An inter-order horizontal gene transfer event enables the catabolism of compatible solutes by *Colwellia psychrerythraea* 34H. *Extremophiles*. **17**: 601-610.

Corkrey R, Olley J, Ratkowsky D, McMeekin T, Ross T. (2012). Universality of thermodynamic constants governing biological growth rates. *PLoS One*. **7**: e32003.

Cottrell MT, Mannino A, Kirchman DL. (2006). Aerobic anoxygenic phototrophic bacteria in the Mid-Atlantic Bight and the North Pacific Gyre. *Appl Environ Microbiol*. **72**: 557-564.

D'Amico S, Collins T, Marx JC, Feller G, Gerday C. (2006). Psychrophilic microorganisms: challenges for life. *EMBO Rep*. **7**: 385-389

D'Aoust DY, Kushner DJ. (1971). Structural changes during lysis of a psychrophilic marine bacterium. *J Bacteriol*. **108**: 916–927.

Delcher AL, Harmon D, Kasif S, White O, Salzberg SL. (1999). Improved microbial gene identification with GLIMMER. *Nucleic Acids Res*. **27**: 4636-4641

DeLong EF, Béjà O. (2010). The light-driven proton pump proteorhodopsin enhances bacterial survival during tough times. *PLoS Biol.* **8**: e1000359.

Dioumaev AK, Brown LS, Shih J, Spudich EN, Spudich JL, Lanyi JK. (2002). Proton transfers in the photochemical reaction cycle of proteorhodopsin. *Biochemistry*. **41**: 5348-5358.

Dioumaev AK, Wang JM, Balint Z, Varo G, Lanyi JK. (2003). Proton transport by proteorhodopsin requires that the retinal Schiff base counterion Asp-97 be anionic. *Biochemistry*. **42**: 6582-6587.

Ewert M, Deming JW. (2011). Selective retention in saline ice of extracellular polysaccharides produced by the cold-adapted marine bacterium *Colwellia psychrerythraea* strain 34H. *Ann Glaciol*. **52**: 111-117.

Fanzmann PD. (1996). Examination of Antarctic prokaryotic diversity through molecular comparisons. *Biodivers Conserv*. **5**: 1295-1305.

Feng S, Powell SM, Wilson R, Bowman JP. (2013). Light stimulated growth of proteorhodopsin bearing sea-ice psychrophile *Psychroflexus torquis* is salinity-dependent. *ISME J*. **7**: 2206-2213.

Feng S, Powell SM, Wilson R, Bowman JP. (2014). Extensive gene acquisition in the extremely psychrophilic bacterial species *Psychroflexus torquis* and the link

to sea-ice ecosystem specialism. *Genome Biol Evol.* **6**: 133-148.

Fisher MM, Wilcox LW, Graham LE. (1998). Molecular characterization of epiphytic bacterial communities on charophycean green algae. *Appl Environ Microbiol*. **64**: 4384–4389.

Frias-Lopez J, Shi Y, Tyson, GW, Coleman ML, Schuster SC, Chisholm SW et al. (2008). Microbial community gene expression in ocean surface waters. *Proc Natl Acad Sci USA.* **105**: 3805-3810.

Friedrich T, Geibel S, Kalmbach R, Chizhov I, Ataka K, Heberle J, Engelhard M, Bamberg E. (2002). Proteorhodopsin is a light-driven proton pump with variable vectoriality. *J Mol Biol.* **321**: 821-838.

Frigaard NU, Martinez A, Mincer TJ, DeLong EF. (2006). Proteorhodopsin lateral gene transfer between marine planktonic Bacteria and Archaea. *Nature.* **439**: 847-850.

Fuhrman JA, Schwalbach MS, Stingl U. (2008). Opinion - Proteorhodopsins: an array of physiological roles? *Nat Rev Microbiol*. **6**: 488-494.

Gao F, Luo H, Zhang CT. (2012). DoriC 5.0: an updated database of *oriC* regions in both bacterial and archaeal genomes. *Nucleic Acids Res*. **41**:D90-D93.

Giannelli V, Thomas DN, Haas C, Kattner G, Kennedy H, Dieckmann GS. (2001). Behavior of dissolved organic matter and inorganic nutrients during experimental sea-ice fomation. *Ann Glaciol.* **33:** 317-321.

Giovannoni SJ, Bibbs L, Cho JC, Stapels MD, Desiderio R, Vergin KL et al. (2005). Proteorhodopsin in the ubiquitous marine bacterium SAR11. *Nature*. **438**:

82-85.

Giovannoni SJ, Stingl U. (2007). The importance of culturing bacterioplankton in the 'omics' age. *Nat Rev Microbiol.* **5**: 820-826.

Gomelsky M, Klug G. (2002). BLUF: a novel FAD-binding domain involved in sensory transduction in microorganisms. *Trends in Biochem Sci*. **27**: 497-500

Gomez-Consarnau L, Akram N, Lindell K, Pedersen A, Neutze R, Milton DL et al. (2010). Proteorhodopsin phototrophy promotes survival of marine bacteria during starvation. *PLoS Biol.* **8**: e1000358.

Gomez-Consarnau L, Gonzalez JM, Coll-Llado M, Gourdon P, Pascher T, Neutze R et al. (2007). Light stimulates growth of proteorhodopsin-containing marine Flavobacteria. *Nature.* **445**: 210-213.

Gonzalez JM, Fernandez-Gomez B, Fernandez-Guerra A, Gomez-Consarnau L, Sanchez O, Coll-Llado M et al. (2008). Genome analysis of the proteorhodopsin-containing marine bacterium *Polaribacter sp* MED152 (Flavobacteria). *Proc Natl Acad Sci USA.* **105:** 8724-8729

Gonzalez JM, Pinhassi J, Fernandez-Gomez B, Coll-Llado M, Gonzalez-Velazquez M, Puigbo P et al. (2011). Genomics of the proteorhodopsin-containing marine flavobacterium *Dokdonia sp* strain MED134. *Appl Environ Microbiol.* **77**: 8676-8686.

Goris J, Konstantinidis KT, Kelppenbach, JA, Coenye T, Vandamme P, Tiedje JM. (2007). DNA-DNA hybridization values and their relationship to whole-genome sequence similarities. *Int J Syst Evol Microbiol*. **57**: 81-91.

Grossmann S, Diekmann GS. (1994). Bacterial standing stock, activity and carbon production of sea ice in the Weddell Sea, Antarctica. *Appl Environ Microbiol*. **60**: 2746-2753.

Hagerthey SE, Louda JW, Mongkronsri P. (2006). Evaluation of pigment extraction methods and a recommended protocol for periphyton chlorophyll a determination and chemotaxonomic assessment. *J Phycol*. **42**: 1125-1136.

Hamana K, Niitsu M. (2001). Large production of an aromatic amine, 2-phenylethylamine, in a psychrophilic marine bacterium, *Psychroflexus torquis*. *J Gen Appl Microbiol*. **47**: 103-105.

Hassler C, Schoemann V, Mancuso-NicholsCA, Butler EC, Boyd P. (2011). Saccharides enhance iron bioavailability to Southern Ocean phytoplankton. *Proc Natl Acad Sci USA*. **108**: 1076-1081.

Hoffman PF, Schrag DP. (2002). The snowball Earth hypothesis testing the limits of global change. *Terra Nova*. **14**:129-155.

Holt TK, Krogmann DW. (1981). A carotenoid-protein from cyanobacteria. *BBA. Bioenergetics.* **637**: 408-414.

Hunken M, Harder J, Kirst GO. (2008). Epiphytic bacteria on the Antarctic ice diatom *Amphiprora kufferathii* Manguin cleave hydrogen peroxide produced during algal photosynthesis. *Plant Biol*. **10**:519-526.

Huston AL, Krieger-Brockett BB, Deming JW. (2000). Remarkably low temperature optima for extracellular enzyme activity from Arctic bacteria and sea ice. *Environ Microbiol*. **2**:383-388.

Hyatt D et al. (2010). Prodigal: prokaryotic gene recognition and translation initiation site identification. *BMC Bioinformati.* **11**:119.

Imlay JA. (2003). Pathways of oxidative damage. *Ann Rev Microbiol*. **57**: 395-418.

James SR, Burton HR, McMeekin TA, Mancuso CA. (1994). Seasonal abundance of *Halomonas meridiana*, *Halomonas subglaciescola*, *Flavobacterium gondwanense* and *Flavobacterium salegens* in four Antarctic lakes. *Ant Sci*. **6**:325-332.

Jiao, NZ., Feng, FY. Wei, B. (2006). Proteorhodopsin - A new path for biological utilization of light energy in the sea. *Chin Sci Bull.* **51**: 889-896.

Junge K, Eicken H, Deming JW. (2004). Bacterial activity at -2 to -20ºC in Arctic wintertime sea ice. *Appl Environ Microbiol*. **70**: 550-557.

Kawano T, Reinhard P, Uozumi N, Miyake C, Asada K, Kolattukudy PE, Muto S. (2000). Aromatic monoamine-induced immediate oxidative burst leading to an increase in cytosolic Ca^{2+} concentration in tobacco suspension culture. *Plant Cell Physiol*. **41**: 1251-1258.

Keller A, Nesvizhskii AI, Kolker E, Aebersold R. (2002). Empirical statistical model to estimate the accuracy of peptide identifications made by MS/MS and database search. *Anal Chem.* **74**: 5383-5392.

Kimura H, Young CR, Martinez A, DeLong EF. (2011). Light-induced transcriptional responses associated with proteorhodopsin-enhanced growth in a marine flavobacterium. *ISME J.* **5**: 1641–1651.

Kirchman DL. (1990). Limitaion of bacterial-growth by dissolved organic-matter in

the sub-Arctic pacific. *Mar Ecol Prog Ser.* **62**: 47-54.

Kirchman DL, Hanson TE. (2013). Bioenergetics of photoheterotrophic bacteria in the oceans. *Environ Microbiol Rep.* **5**: 188-199.

Kocharunchitt C, King T, Gobius K, Bowman JP, Ross T. (2012). Integrated transcriptomic and proteomic analysis of the physiological response of *Escherichia coli* O157:H7 sakai to steady-state conditions of cold and water activity stress. *Mol Cell Proteomics.* **11**: M111.009019.

Kogure K. (1998). Bioenergetics of marine bacteria. *Curr Opin Biotechnol.* **9**: 278-282.

Koh EY, Atamna-Ismaeel N, Martin A, Cowie ROM, Béjà O, Davy SK et al. (2010). Proteorhodopsin-bearing bacteria in Antarctic sea-ice. *Appl Environ Microbiol.* **76**: 5918-5925.

Kottmeier ST, Grossi SM, Sullivan CW. (1987). Sea ice microbial communities. VIII. Bacterial production in annual sea ice of McMurdo Sound, Antarctica. *Mar Ecol Prog Ser*. **35**: 175-186.

Kottmeier ST, Sullivan CW. (1987). Late winter primary production and bacterial production in sea ice and seawater west of the Antarctic Peninsula. *Mar Ecol Prog Ser*. **36**: 287-298.

Krembs C, Deming JW. (2008). The role of exopolymers in microbial adaptation to sea ice. In: Margesin, R., Schinner, F., Marx, JC., Gerday, C (eds) Psychrophiles: from Biodiversity to Biotechnology. Springer Berlin Heidelberg, pp 247-264.

Krembs C, Eicken H, Deming JW. (2011). Exopolymer alteration of physical properties of sea-ice and implications for ice habitability and biogeochemistry in a warmer Arctic. *Proc Natl Acad Sci USA.* **108**: 3653-3658.

Kuo CH, Ochman H. (2010). The extinction dynamics of bacterial pseudogenes. *PLoS Biol*. 6:e1001050

Kwon SK, Kim BK, Song JY, Kwak MJ, Lee CH, Yoon JH et al. (2012). Genomic makeup of the marine Flavobacterium *Nonlabens* (*Donghaeana*) *dokdonensis* and identification of a novel class of rhodopsins. *Genome Biol Evol* **5**: 187-199.

Lakatos M, Lanyi JK, Szakacs J, Varo G. (2003). The photochemical reaction cycle of proteorhodopsin at low pH. *Biophysical Journal.* **84**: 3252-3256.

Lambrecht M, Okon Y, Vande Broek A, Vanderleyden J. (2000). Indole-3-acetic acid: a reciprocal signaling molecule in bacteria–plant interactions. *Trends Microbiol.* **8**: 298-300.

Lami R, Cottrell MT, Campbell BJ, Kirchman DL. (2009). Light-dependent growth and proteorhodopsin expression by Flavobacteria and SAR11 in experiments with Delaware coastal waters. *Environ Microbiol.* **11**: 3201-3209.

Lanyi JK. (2004). Bacteriorhodopsin. *Ann Rev Physiol.* **66**: 665-688.

Lerat E, Ochman H. (2005). Recognizing the pseudogenes in bacterial genomes. *Nucleic Acids Res*. **33**: 3125-3132.

Liu HB, Sadygov RG, Yates JR. (2004). A model for random sampling and estimation of relative protein abundance in shotgun proteomics. *Anal Chem*.

76: 4193-4201.

Lowe TM, Eddy SR. (1997). tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence. *Nucleic Acids Res*. **25**:955-964.

Ma D, Cook DN, Hearst JE, Nikaido H. (1994). Efflux pumps and drug resistance in Gram-negative bacteria. *Trends Microbiol* **2**: 489-493.

Macdonald FA, Schmitz MD, Crowley JL, Roots CF, Jones DS, Maloof AC et al. (2010). Calibrating the Cryogenian. *Science*. **327**: 1241–1243.

Madeo F, Herker E, Maldener C, Wissing S, Lächelt S, Herlan M et al. (2002). A caspase-related protease regulates apoptosis in yeast. *Mol Cell*. **9**: 911-917.

Mancuso Nichols, CA. (2005). Exopolysaccharide production by Antarctic marine bacteria. PhD thesis, University of Tasmania, Australia.

Man-Aharonovich D, Sabehi G, Sineshchekov OA, Spudich EN, Spudich JL, Béjà O. (2004). Characterization of RS29, a blue-green proteorhodopsin variant from the Red Sea. *Photochem Photobiol Sci*. **3**: 459-462.

Man DL, Wang WW, Sabehi G, Aravind L, Post AF, Massana R et al. (2003). Diversification and spectral tuning in marine proteorhodopsins. *Embo J,* **22**: 1725-1731.

Martinez A, Bradley AS, Waldbauer JR, Summons RE, Delong EF. (2007). Proteorhodopsin photosystem gene expression enables photophosphorylation in a heterologous host. *Proc Natl Acad Sci USA,* **104,** 5590-5595.

Marx JG, Carpenter SD, Deming JW. (2009). Production of cryoprotectant extracellular polysaccharide substances (EPS) by the marine psychrophilic

bacterium *Colwellia psychrerythraea* strain 34H under extreme conditions. *Can J Microbiol*. **55**: 63-72.

McBride MJ, Xie G, Martens EC, Lapidus A, Henrissat B, Rhodes RG et al. (2009). Novel features of the polysaccharide-digesting gliding bacterium *Flavobacterium johnsoniae* as revealed by genome sequence analysis. *Appl Environ Microbiol* **75**: 6864-6875.

McBride MJ, Zhu Y. (2013). Gliding motility and Por secretion system genes are widespread among members of the phylum *Bacteroidetes*. *J Bacteriol*. **195**: 270-278.

McCarren J, DeLong EF. (2007). Proteorhodopsin photosystem gene clusters exhibit co-evolutionary trends and shared ancestry among diverse marine microbial phyla. *Environ Microbiol* **9**: 846-858.

McMinn A, Martin A. (2013). Dark survival in a warming world. *Proc RoyalRoy Soc B: Biol Sci* **280**: 20122909.

McMinn A, Ryan KG, Ralph PJ, Pankowski A. (2007). Spring sea-ice photosynthesis, primary productivity and biomass distribution in eastern Antarctica, 2002-2004. *Mar Biol* **151**: 985-995.

Methé BA, Nelson KE. Deming JW, Momen B, Melamud E, Zhang X, Moult J, Madupu R et al. (2005). The psychrophilic lifestyle as revealed by the genome sequencing of *Colwellia psychrerythraea* 34H through genome and proteomic analyses. *Proc Natl Acad Sci USA*. **102**: 10913-10918.

Michelou VK, Cottrell MT, Kirchman DL. (2007). Light-stimulated bacterial

production and amino acid assimilation by cyanobacteria and other microbes in the North Atlantic Ocean. *Appl Environ Microbiol* **73**: 5539-5546.

Moran MA, Miller WL. (2007). Resourceful heterotrophs make the most of light in the coastal ocean. *Nat Rev Microbiol*. **5**: 792-800.

Mosier AC, Murray AE, Fritsen CH. (2007). Microbiota within the perennial ice cover of Lake Vida, Antarctica. *FEMS Microbiol Ecol*. **59**: 274-288.

Nandakumar MP, Shen J, Raman B, Marten MR. (2003). Solubilization of trichloroacetic acid (TCA) precipitated microbial proteins via NaOH for two-dimensional electrophoresis. *J Proteome Res*. **2**: 89-93.

Nesvizhskii AI, Keller A Kolker, Aebersold R. (2003). A statistical model for identifying proteins by tandem mass spectrometry. *Anal Chem*. **75**: 4646-4658.

Nichols CAM, Guezennec J, Bowman JP. (2005). Bacterial exopolysaccharides from extreme marine environments with special consideration of the Southern Ocean, sea ice, and deep-sea hydrothermal vents: a review. *Marine Biotechnol.* **7**: 253-271.

Nichols DS, Brown JL, Nichols PD, McMeekin TA. (1997). Production of eicosapentaenoic acid and arachidonic acids by an Antarctic bacterium: response to growth temperature. *FEMS Microbiol Lett.* **152**: 349-354.

Nishida T, Hori R, Morita N, Okuyama H. (2010). Membrane eicosapentaenoic acid is involved in the hydrophobicity of bacterial cells and affects the entry of hydrophilic and hydrophobic compounds. *FEMS Microbiol Lett*. **306**: 91-96.

Norman L, Thomas DN, Stedmon CA, Granskog MA, Papadimitriou S, Krapp R et al. (2011). The characteristics of dissolved organic matter (DOM) and chromophoric dissolved organic matter (CDOM) in Antarctic sea ice. *Deep-Sea Res Pt II*. **59**: 1075-1091.

Ochman H, Lawrence JG, Groisman EA. (2000). Lateral gene transfer and the nature of bacterial innovation. *Nature*. **40**: 299-304

Oestrhe D, Stoecken W. (1971). Rhodopsin-like protein from purple membrane of *Halobacterium-Halobium*. *Nature New Biol*. **233**: 149-152.

Oren A. (1999). Bioenergetic aspects of halophilism. *Microbiol Mol Biol Rev.* **63**: 334.

Ou K, Ong C, Koh SY, Rodrigues F, Slim SH, Wong, D et al. (2005). Integrative genomic, transcriptional, and proteomic diversity in natural isolates of the human pathogen *Burkholderia pseudomallei*. *J Bacteriol*. **187**: 4276-4285.

Ou HY, Chen LL, Lonnen J, Chaudhuri RR, Thani AB, Smith R et al. (2006). A novel strategy for the identification of genomic islands by comparative analysis of the contents and contexts of tRNA sites in closely related bacteria. *Nucleic Acids Res*. **34**: e3-e3.

Petersen TN, Brunak S, Heijne G, Nielsen, H. (2011). SignalP 4.0: discriminating signal peptides from transmembrane regions. *Nature Methods*. **8**: 785-786.

Pfaffl MW. (2001). A new mathematical model for relative quantification in real-time RT-PCR. *Nucleic Acids Res*. **29**: e45-e45

Polyak L. (2010). History of sea ice in the Arctic. *Quat Sci Rev*. **29**: 1757–1778.

Poulin M, Daugbjerg N, Gradinger R, Ilyash L, Ratkova T, Quilfeldt CV. (2011). The pan-Arctic biodiversity of marine phytoplankton and sea-ice unicellular eukaryotes: A first-attempt assessment. *Mar Biodivers*. **41**:13–28.

Puigbò P, Bravo IG, Garcia-Vallvé S. (2008). PE-CAI: a novel server to estimate an expected value of Codon Adaptation Index (eCAI). *BMC Bioinformat*. **9**: 65.

Pusceddu A, Dell'Anno A, Vezzulli L, Fabiano M, Saggiomo V, Cozzi S. (2009). Microbial loop malfunctioning in the annual sea ice at Terra Nova Bay (Antarctica). *Polar Biol* **32**: 337-346.

Ratkowskey DA, Lowry RK, McMeekin TA, Stokes AN, Chandler RE. (1983). Model for bacterial culture growth rate throughout the entire biokinetic temperature range. *J. Bacteriol*. **154**: 1222-1226.

Rauch A, Bellew M, Eng J, Fitzgibbon M, Holzman T, Hussey P et al. (2006). Computational proteomics analysis system (CPAS): an extensible, open-source analytic system for evaluating and publishing proteomic data and high throughput biological experiment. *J Proteome Res*. **5**: 112-121.

Raymond, JA, Fritsen C, Shen K. (2007). An ice-binding protein from an Antarctic sea ice bacterium. *FEMS Microbiol Ecol*. **61**: 214-221.

Reimnitz E, Clayton JR, Kempema EW, Payne JR, Weber WS. (1993). Interaction of rising frazil with suspended particles–tank experiments with applications to nature. *Cold Reg Sci Technol*. **21**: 117-135.

Riedel T, Tomasch J, Buchholz I, Jacobs J, Kollenberg M, Gerdts G et al. (2010). Constitutive expression of the proteorhodopsin gene by a flavobacterium

strain representative of the proteorhodopsin-producing microbial community in the North Sea. *Appl Environ Microbiol.* **76**: 3187-3197.

Riley M, Staley JT, Danchin A, Wang TZ, Brettin TS, Hauser L et al. (2008). Genomics of an extreme psychrophile, *Psychromonas ingrahamii*. *BMC Genomics*. **9**: 210.

Rivkin RB, Legendre L. (2001). Biogenic carbon cycling in the upper ocean: Effects of microbial respiration. *Science.* **291**: 2398-2400.

Rusch DB, Halpern AL, Sutton G, Heidelberg KB, Williamson S, Yooseph S, Wu DY et al. (2007). The Sorcerer II global ocean sampling expedition: Northwest Atlantic through Eastern tropical Pacific. *PLoS Biol.* **5**: 398-431.

Rusk N. (2011). From pseudogenes to proteins. *Nature Methods*. **8**: 448–449.

Russell NJ, Nichols DS. (1999). Polyunsaturated fatty acids in marine bacteria--a dogma rewritten. *Microbiology.* **145**: 767-779.

Sabehi G, Béjà O, Suzuki MT, Preston CM, DeLong EF. (2004). Different SAR86 subgroups harbour divergent proteorhodopsins. *Environ Microbiol.* **6**: 903-910.

Sabehi G, Loy A, Jung KH, Partha R, Spudich JL, Isaacson T et al. (2005). New insights into metabolic properties of marine bacteria encoding proteorhodopsins. *PLoS Biol.* **3**: 1409-1417.

Sabehi G, Massana R, Bielawski JP, Rosenbergy M, Delong EF, Béjà O. (2003). Novel proteorhodopsin variants from the Mediterranean and Red Seas. *Environ Microbiol.* **5**: 842-849.

Sato K, Naito M, Yukitake H, Hirakawa H, Shoji M, McBride MJ et al. (2010). A protein secretion system linked to bacteroidete gliding motility and pathogenesis. *Proc Natl Acad Sci USA.* **107**: 276-281.

Sanchez C. (2011). Horizontal gene transfer eukaryotes under a new light. *Nat Rev Microbiol.* **9**: 228-229.

Schwalbach MS, Brown M, Fuhrman JA. (2005). Impact of light on marine bacterioplankton community structure. *Aquat Microb Ecol.* **39**: 235-245.

Schwalbach MS, Fuhrman JA. (2005). Wide-ranging abundances of aerobic anoxygenic phototrophic bacteria in the world ocean revealed by epifluorescence microscopy and quantitative PCR. *Limnol Oceanogr.* **50**: 620-628.

Schwartzman D, Lineweaver CH. (2005). Temperature, biogenesis and biospheric self-organization. In: Kleidon A Lorenz R, editors. *Non-equilibrium Thermodynamics and the Production of Entropy: Life, Earth and Beyond*. New York: Springer. p. 207-217.

Searle BC. (2010). Scaffold: A bioinformatic tool for validating MS/MS-based proteomic studies. *Proteomics*. **10**: 1265-1269.

Seiler H, Bleicher A, Busse H J, Hüfner J, Scherer S. (2012). *Psychroflexus halocasei* sp. nov., isolated from a microbial consortium on a cheese. *Int J Syst Evol Microbiol*. **62**: 1850-1856.

Steindler L, Schwalbach MS, Smith DP, Chan F, Giovannoni SJ. (2011). Energy starved *Candidatus* Pelagibacter ubique substitutes light-mediated ATP

production for endogenous carbon respiration. *PLoS One* **6**: e19725.

Sharma AK, Spudich JL, Doolittle WF. (2006). Microbial rhodopsins: functional versatility and genetic mobility. *Trends Microbiol.* **14**: 463-469.

Shi LC, Laje EMR, Ahmed MAM, Brown LS, Ladizhansky V. (2009). Solid-state NMR study of proteorhodopsin in the lipid environment: Secondary structure and dynamics. *BBA - Biomembranes.* **1788**: 2563-2574.

Shulse CN, Allen EE. (2011). Diversity and distribution of microbial long-chain fatty acid biosynthetic genes in the marine environment. *Environ Microbiol*. **13**: 684-695.

Slamovits CH, Okamoto N, Burri L, James ER, Keeling PJ. (2011). A bacterial proteorhodopsin proton pump in marine eukaryotes. *Nat Commun.* **2**: 183

Smith WO, Nelson DM. (1985). Phytoplankton bloom produced by a receding ice edge in the Ross Sea: spatial coherence with the density field. *Science.* **227**: 163-166.

Spudich JL, Yang CS, Jung KH, Spudich EN. (2000). Retinylidene proteins: Structures and functions from archaea to humans. *Ann Rev Cell Dev Biol.* **16:** 365

Staley JT, Gosink JJ. (1999). Poles apart: Biodiversity and biogeography of sea ice bacteria. *Ann Rev Microbiol*. **53**: 189-215.

Stingl U, Desiderio RA, Cho JC, Vergin KL, Giovannoni SJ. (2007). The SAR92 clade: an abundant coastal clade of culturable marine bacteria possessing proteorhodopsin. *Appl Environ Microbiol.* **73**: 2290-2296.

Straza TRA, Kirchman DL. (2011). Single-cell response of bacterial groups to light and other environmental factors in the Delaware Bay, USA. *Aquat Microb Ecol.* **62**: 267-277.

Stock AM, Robinson VL, Goudreau PN. (2000). Two-component signal transduction. *Ann Rev Biochem.* **69**: 183-215.

Struvay C, Feller G. (2012). Optimization to low temperature activity in psychrophilic enzymes. *Int J Mol Sci*. **13**: 11643-11665

Sul WJ, Oliver TA, Ducklow HW, Amaral-Zettler LA, Sogin ML. (2013). Marine bacteria exhibit a bipolar distribution. *Proc Natl Acad Sci USA.* **110**: 2342-2347.

Swartz TE, Tseng TS, Frederickson MA, Paris G, Comerci DJ, Rajashekara G et al. (2007). Blue-light-activated histidine kinases: Two-component sensors in bacteria. *Science.* **317**: 1090-1093.

Tang KH, Feng XY, Tang YJJ, Blankenship RE. (2009). Carbohydrate metabolism and carbon fixation in *Roseobacter denitrificans* OCh114. *PLoS One.* **4**: e7233

Taylor BL, Zhulin IB. (1999). PAS domains: Internal sensors of oxygen, redox potential, and light. *Microbiol Mol Biol Rev.* **63**: 479.

Thomas DN, Dieckmann GS. (2002). Antarctic sea-ice - a habitat for extremophiles. *Science.* **295**: 641-644.

Turner J, Comiso JC, Marshall GJ, Lachlan-Cope TA, Bracegirdle T, Maksym T et al. (2009). Non-annular atmospheric circulation change induced by stratospheric

ozone depletion and its role in the recent increase of Antarctic sea ice extent. *Geophys Res Lett*. **36**: L08502

Unemoto T, Hayashi M. (1993). Na^+-translocating NADH-quinone reductase of marine and halophilic bacteria. *J Bioenerg Biomembr.* **25**: 385–391.

Usui K, Hiraki T, Kawamoto J, Kurihara T, Nogi Y, Kato C et al. (2012). Eicosapentaenoic acid plays a role in stabilizing dynamic membrane structure in the deep-sea piezophile *Shewanella violacea*: A study employing high-pressure time-resolved fluorescence anisotropy measurement. *Biochim Biophys Acta.* **1818**: 574-583.

Venter JC, Remington K, Heidelber JF, Halpern AL, Rusch D, Eisen JA et al.(2004). Environmental genome shotgun sequencing of the Sargasso Sea. *Science.* **304**: 66-74.

Vercammen D, Declerq W, Vandenabeele P, Van Breusegem F. (2007). Are metacaspases caspases? *J Cell Biol*. **179**: 375-380.

Vollmers J, Voget S, Dietrich S, Gollnow K, Smits M, Meyer K et al. (2013). Poles Apart: Arctic and Antarctic *Octadecabacter* strains share high genome plasticity and a new type of xanthorhodopsin. *PLoS One*. **8**: e63422

Vizcaino JA, Cote RG, Csordas A, Dianes JA, Fabregat A, Foster JM. (2013). The Proteomics Identifications (PRIDE) database and associated tools: status in 2013. *Nucleic Acids Res.* **41**: 1063-1069.

Walter JM, Greenfield D, Bustamante C, Liphardt J. (2007). Light-powering *Escherichia coli* with proteorhodopsin. *Proc Natl Acad Sci USA.* **104**:

2408-2412.

Wang WW, Sineshchekov OA, Spudich EN, Spudich JL. (2003). Spectroscopic and photochemical characterization of a deep ocean proteorhodopsin. *Journal of Biological Chemistry.* **278**: 33985-33991.

Ward NL, Challacombe JF, Janssen PH, Henrissat B, Coutinbo P,M, Wu M et al. (2009). Three genomes from the phylum *Acidobacteria* provide insight into the lifestyles of these microorganisms in soils. *Appl Environ Microbiol*. **75**: 2046-2056.

Wells LE, Deming JW. (2006). Modelled and measured dynamics of viruses in Arctic winter sea-ice brines. *Environ Microbiol*. **8**: 1115-1121.

Whitaker RJ. (2006). Allopatric origins of microbial species. *Philos Trans R Soc Lond B Biol Sci*. **361**: 1975–1984.

Whitfield J. (2009). Origin of life: Nascence man. *Nature*. **459**: 316-319.

Whitman WB, Coleman DC, Wiebe W. (1998). Prokaryotes: The unseen majority. *Proc Natl Acad Sci US.* **95**: 6578-6583.

Whitney TJ, Gardner DG, Mott ML, Brandon M. (2010). Identifying the molecular basis of functions in the transcriptome of the social amoeba *Dictyostelium discoideum*. *Genet Mol Res*. **9**: 394-415.

Yau S, Lauro FM, Williams TJ, Demaere MZ, Brown MV, Rich J et al. (2013). Metagenomic insights into strategies of carbon conservation and unusual sulfur biogeochemistry in a hypersaline Antarctic lake. *ISME J*. **7**: 1944-1961.

Yoshizawa S, Kawanabe A, Ito H, Kandori H, Kogure K. (2012). Diversity and

functional analysis of proteorhodopsin in marine Flavobacteria. *Environ Microbiol.* **14**: 1240-1248.

Zeldovich KB, Berezovsky IN, Shakhnovich EI. (2007) Protein and DNA sequence determinants of thermophilic adaptation. *PLoS Comput Biol*. **3**: e5.

Zgurskaya HI, Nikaido H. (1999). Bypassing the periplasm: Reconstitution of the AcrAB multidrug efflux pump of *Escherichia coli*. *Proc Natl Acad Sci USA.* **96**: 7190-7195.

Zubkov MV, Tarran GA, Fuchs BM. (2004). Depth related amino acid uptake by *Prochlorococcus* cyanobacteria in the Southern Atlantic tropical gyre. *FEMS Microbiology Ecology,* **50**: 153-161.

Appendix A

Light stimulated growth of proteorhodopsin bearing sea-ice psychrophile *Psychroflexus torquis* is salinity-dependent (Chapter 2)

Figure A.1

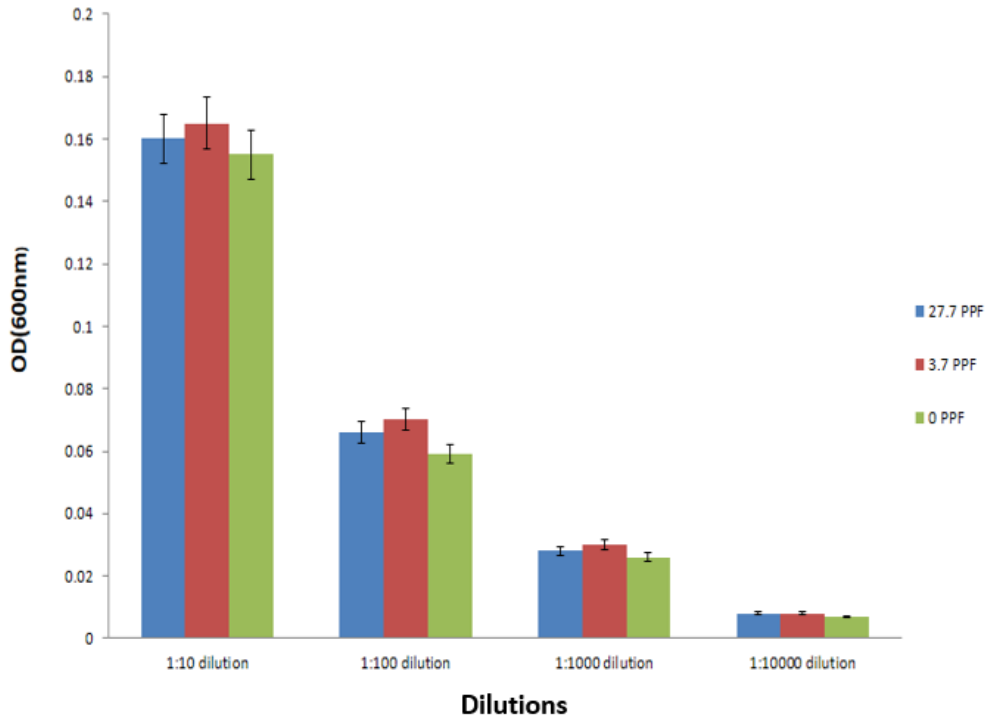
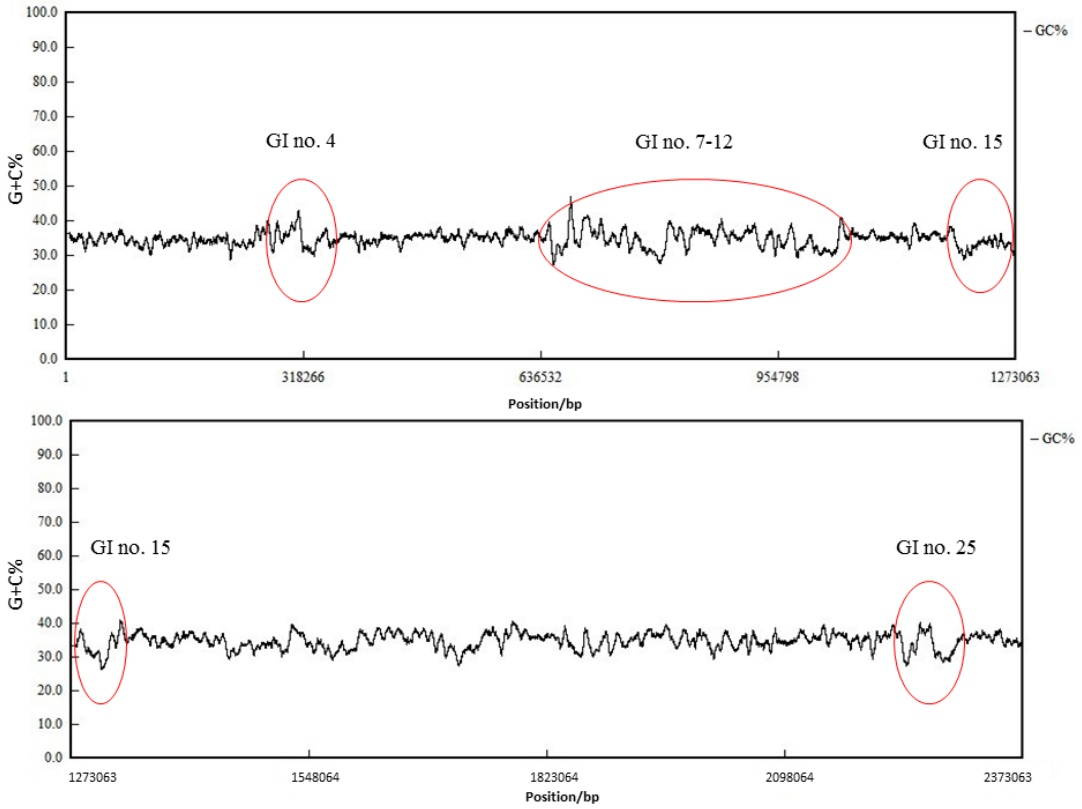


Figure A.1. P. Growth of *P. torquis* in 1:10, 1:100, 1:1000 and 1:10000 diluted marine broth under three different illumination levels (27.7 PPF, 3.7PPF and Dark) at salinity 35 $g.L^{-1}$. Cell populations were determined by optical density.

Appendix B

Extensive gene acquisition in the extremely psychrophilic bacterial species *Psychroflexus torquis* and the link to sea-ice ecosystem specialism (Chapter 3)

Figure B.1



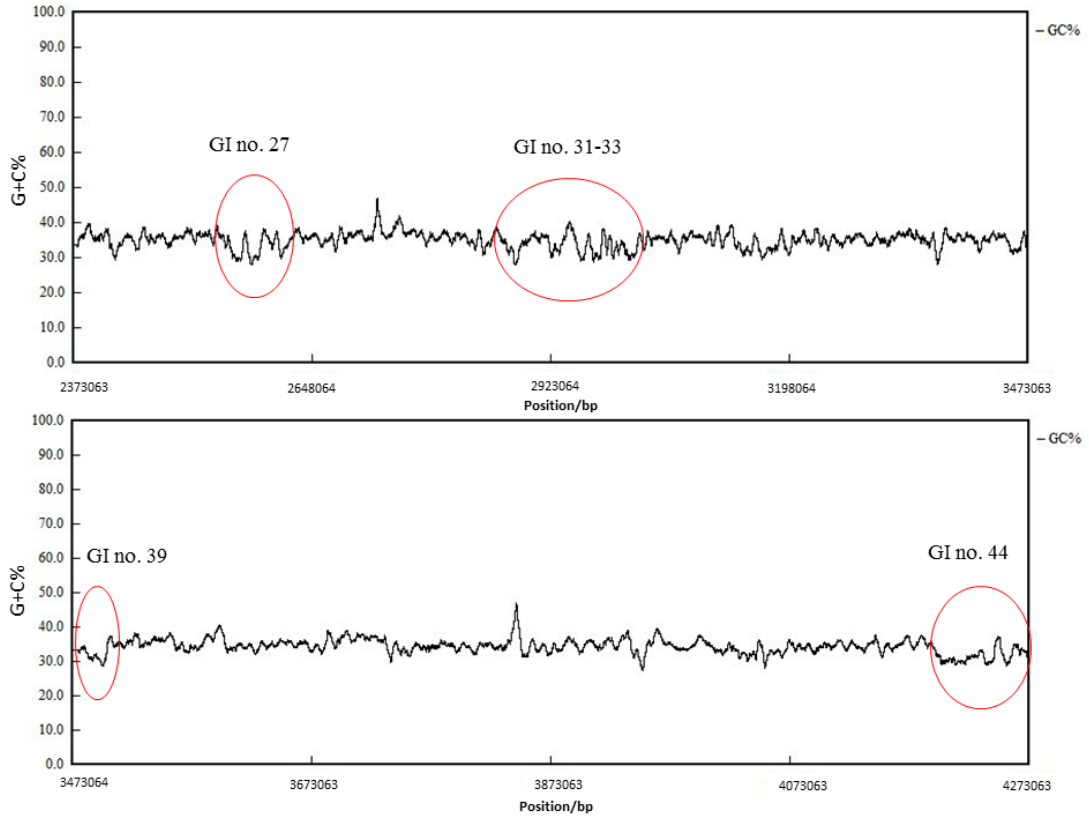


Figure B.1. Overall GC percentage of *P. torquis* plotted using DNAMAN. The average GC% of *P. torquis* is 34.5%. There are 15 GIs located in regions where the GC% varies from 26% to 50% (red circled regions). These suggest an external origin and ancestry of these genomic regions

Supplementary data available online”

Table B.1

http://gbe.oxfordjournals.org/content/suppl/2013/12/27/evt209.DC1/Table_S1_Ptorquis_genome_supplementary_data_Feng_etal.xlsx

Table B.2

http://gbe.oxfordjournals.org/content/suppl/2013/12/27/evt209.DC1/TableS2_Pgondwanensis_genome_data_supplementary.xlsx

Table B.3

http://gbe.oxfordjournals.org/content/suppl/2013/12/27/evt209.DC1/TableS3_Ptorquis_pseudogene_data_supplementary.xlsx

Appendix C

Light and salinity induced proteomic response in a proteorhodopsin-containing sea-ice dwelling *flavobacterium*

Figure C.1

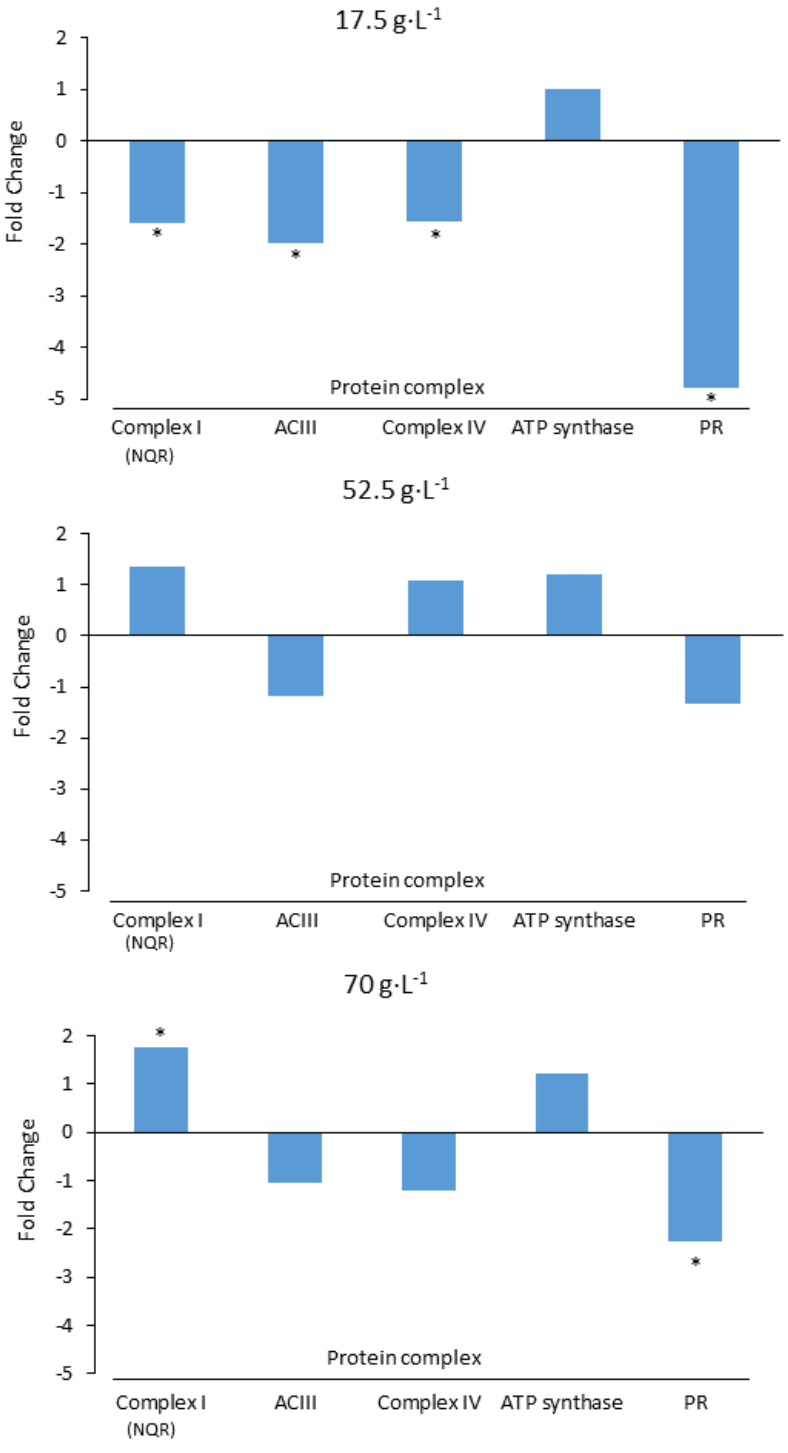
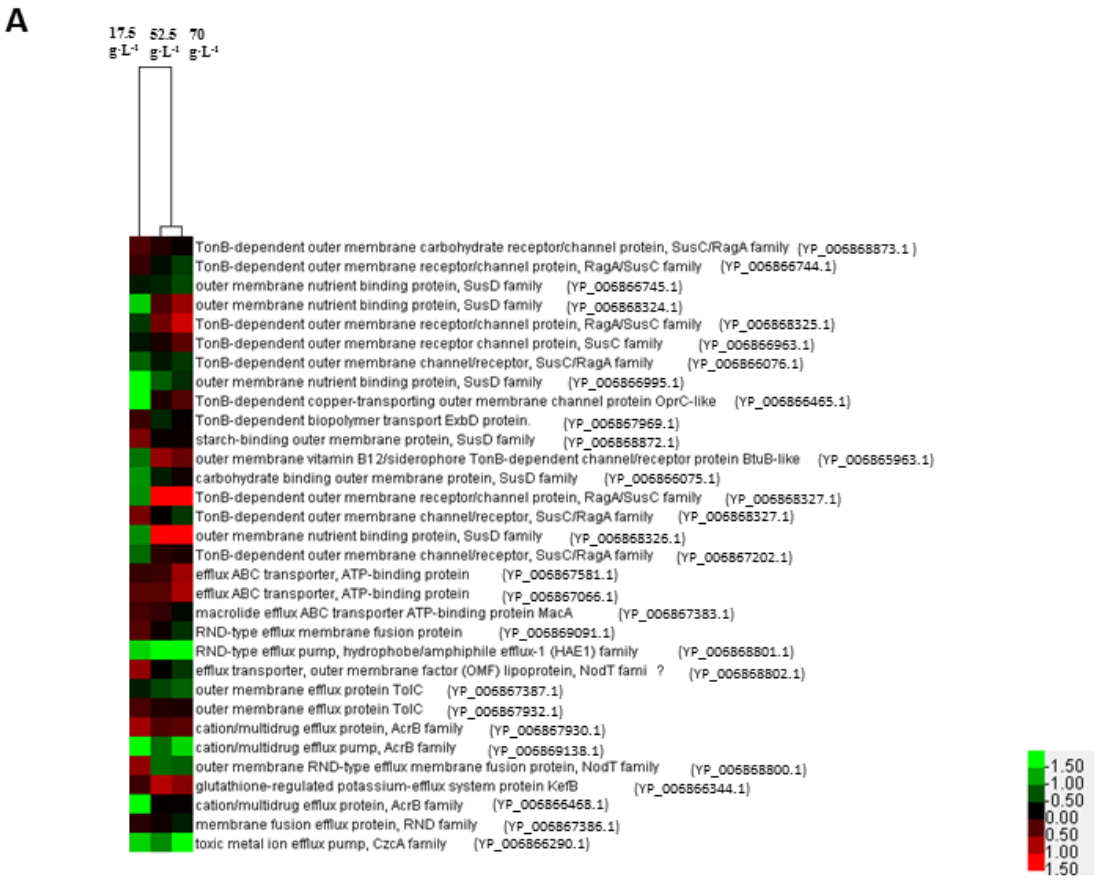


Figure C.1 Fold change of electron transport chain proteins under different salinities. The spectral counts from all three light conditions for each salinity level were pooled together to compare with the optimal salinity ($35 g.L^{-1}$). Asterisks indicate *P* value of the fold change is less than 0.05.

Figure C.2


A
17.5 g·L⁻¹
52.5 g·L⁻¹
70 g·L⁻¹
TonB-dependent outer membrane carbohydrate receptor/channel protein, SusC/RagA family (YP_006868873.1)
TonB-dependent outer membrane receptor/channel protein, RagA/SusC family (YP_006866744.1)
outer membrane nutrient binding protein, SusD family (YP_006866745.1)
outer membrane nutrient binding protein, SusD family (YP_006868324.1)
TonB-dependent outer membrane receptor/channel protein, RagA/SusC family (YP_006868325.1)
TonB-dependent outer membrane receptor channel protein, SusC family (YP_006866963.1)
TonB-dependent outer membrane channel/receptor, SusC/RagA family (YP_006866076.1)
outer membrane nutrient binding protein, SusD family (YP_006866995.1)
TonB-dependent copper-transporting outer membrane channel protein OprC-like (YP_006866465.1)
TonB-dependent biopolymer transport ExbD protein. (YP_006867969.1)
starch-binding outer membrane protein, SusD family (YP_006868872.1)
outer membrane vitamin B12/siderophore TonB-dependent channel/receptor protein BtuB-like (YP_006865963.1)
carbohydrate binding outer membrane protein, SusD family (YP_006866075.1)
TonB-dependent outer membrane receptor/channel protein, RagA/SusC family (YP_006868327.1)
TonB-dependent outer membrane channel/receptor, SusC/RagA family (YP_006868327.1)
outer membrane nutrient binding protein, SusD family (YP_006868326.1)
TonB-dependent outer membrane channel/receptor, SusC/RagA family (YP_006867202.1)
efflux ABC transporter, ATP-binding protein (YP_006867581.1)
efflux ABC transporter, ATP-binding protein (YP_006867066.1)
macrolide efflux ABC transporter ATP-binding protein MacA (YP_006867383.1)
RND-type efflux membrane fusion protein (YP_006869091.1)
RND-type efflux pump, hydrophobe/amphiphile efflux-1 (HAE1) family (YP_006868801.1)
efflux transporter, outer membrane factor (OMF) lipoprotein, NodT fami ? (YP_006868802.1)
outer membrane efflux protein TolC (YP_006867387.1)
outer membrane efflux protein TolC (YP_006867932.1)
cation/multidrug efflux protein, AcrB family (YP_006867930.1)
cation/multidrug efflux pump, AcrB family (YP_006869138.1)
outer membrane RND-type efflux membrane fusion protein, NodT family (YP_006868800.1)
glutathione-regulated potassium-efflux system protein KefB (YP_006866344.1)
cation/multidrug efflux protein, AcrB family (YP_006866468.1)
membrane fusion efflux protein, RND family (YP_006867386.1)
toxic metal ion efflux pump, CzcA family (YP_006866290.1)
-1.50
-1.00
-0.50
0.00
0.50
1.00
1.50

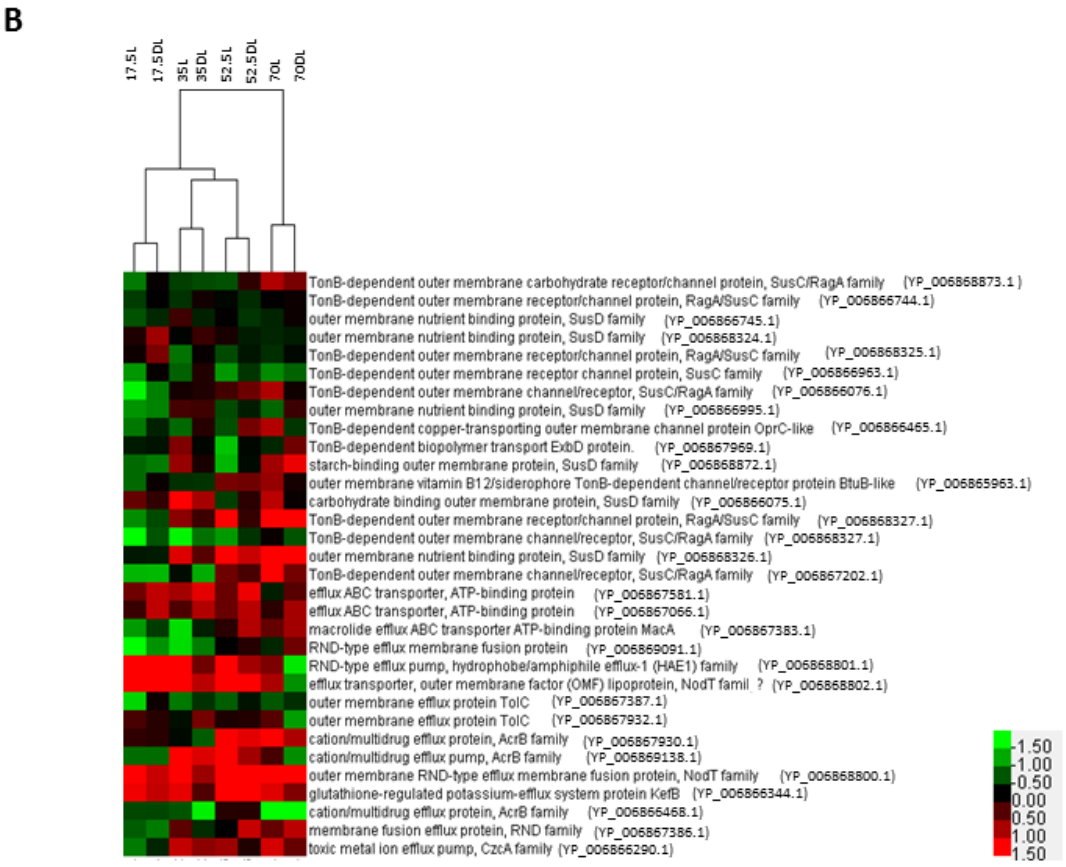


Figure C.2. Heat map generated from proteomic data for some TonB-dependent transporters and efflux transporters. Hierarchical clustering was performed based on the spectral counts. (A) represents the response to different salinities. The optimal salinity was used as control and spectral counts from all three light conditions for each salinity level were pooled together and compared to the optimal salinity. (B) represents response of proteins to different illumination level compare to dark at their corresponding salinity.

Figure C.3

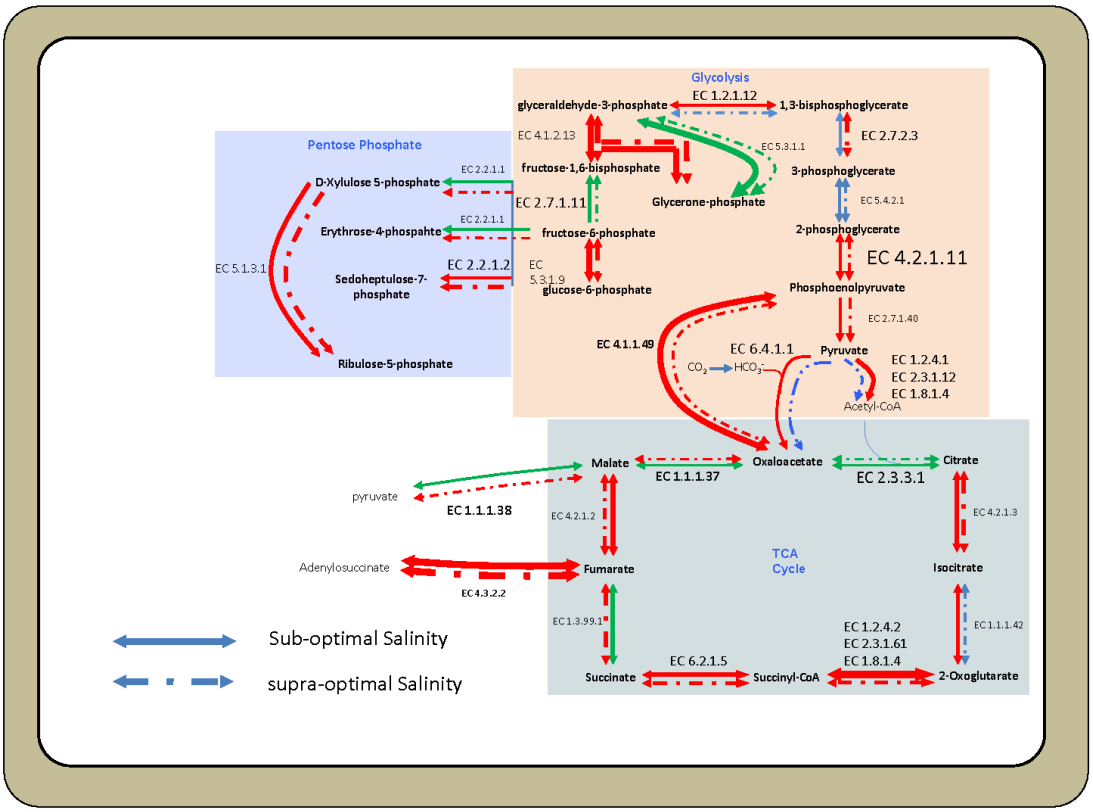


Figure C.3. Diagram of central metabolic pathway change under different salinities. The optimal salinity (35g.L^{-1}) was used as control to compare with sub- and supra-optimal salinities. Within the metabolic map, red line indicates up-regulation, green line indicates down-regulation, blue line indicates unchanged. The thickness of the lines implies the degree of change. The size of enzyme code indicates the abundance of the enzyme.

Table S1: Spectral count and number of proteins identified by LC-MS/MS from each sample.

Spectra and Protein / Treatment	Spectral Count	Number of Proteins Identified
17.6 $g.L^{-1}$ (High Light*) No. A	19039	790
17.6 $g.L^{-1}$ (High Light) No. B	19303	758
17.6 $g.L^{-1}$ (High Light) No. C	18720	672
17.6 $g.L^{-1}$ (Low Light) No. A	18411	804
17.6 $g.L^{-1}$ (Low Light) No. B	19787	709
17.6 $g.L^{-1}$ (Low Light) No. C	19491	725
17.6 $g.L^{-1}$ (Dark) No. A	19160	664
17.6 $g.L^{-1}$ (Dark) No. B	18737	793
17.6 $g.L^{-1}$ (Dark) No. C	14511	553
35 $g.L^{-1}$ (High Light) No. A	14818	581
35 $g.L^{-1}$ (High Light) No. B	17104	648
35 $g.L^{-1}$ (High Light) No. C	17340	653
35 $g.L^{-1}$ (Low Light) No. A	11278	410
35 $g.L^{-1}$ (Low Light) No. B	16611	509
35 $g.L^{-1}$ (Low Light) No. C	19610	777
35 $g.L^{-1}$ (Dark) No. A	17878	570
35 $g.L^{-1}$ (Dark) No. B	15033	489
35 $g.L^{-1}$ (Dark) No. C	18671	708
52.5 $g.L^{-1}$ (High Light) No. A	16165	579
52.5 $g.L^{-1}$ (High Light) No. B	13216	424
52.5 $g.L^{-1}$ (High Light) No. C	18243	742
52.5 $g.L^{-1}$ (Low Light) No. A	12942	574
52.5 $g.L^{-1}$ (Low Light) No. B	18139	751
52.5 $g.L^{-1}$ (Low Light) No. C	18466	788
52.5 $g.L^{-1}$ (Dark) No. A	16658	591
52.5 $g.L^{-1}$ (Dark) No. B	12425	427
52.5 $g.L^{-1}$ (Dark) No. C	18788	753
70 $g.L^{-1}$ (High Light) No. A	15477	676
70 $g.L^{-1}$ (High Light) No. B	17805	769
70 $g.L^{-1}$ (High Light) No. C	18142	798
70 $g.L^{-1}$ (Low Light) No. A	18010	573
70 $g.L^{-1}$ (Low Light) No. B	16258	727
70 $g.L^{-1}$ (Low Light) No. C	13009	684
70 $g.L^{-1}$ (Dark) No. A	17902	578
70 $g.L^{-1}$ (Dark) No. B	17046	685
70 $g.L^{-1}$ (Dark) No. C	18772	726

*High Light is 27.7 PPF and Low light is 3.7 PPF

Table S2. *P. torquis* proteome response to different levels illumination compare to dark under different salinities

Proteins (Reference No./ Name)	17.5g·L^{-1}				35g·L^{-1}				52.5g·L^{-1}				70g·L^{-1}			
	High Light		Low Light		High Light		Low Light		High Light		Low Light		High Light		Low Light	
	FC	p value	FC	p value	FC	p value	FC	p value	FC	p value	FC	p value	FC	p value	FC	p value
Central Metabolism																
Glycolysis																
YP_006868884.1, Phosphoglucose isomerase	1.82	**0.04**	1.33	0.19	0.49	0.13	0.66	0.31	1.13	0.64	2.41	**0.04**	1.67	0.21	1.59	0.27
YP_006868881.1, Phosphohexokinase	0.62	**<0.01**	0.90	0.45	0.72	0.09	1.22	0.39	1.10	0.36	1.42	**0.01**	1.27	0.07	1.16	0.23
YP_006865907.1, Aldolase	0.86	0.74	0.38	0.06	0.45	**0.02**	0.87	0.78	0.58	0.09	3.74	**0.03**	1.08	0.48	0.72	0.74
YP_006868950.1, Triosephosphate isomerase	3.31	**<0.01**	1.47	0.11	1.34	0.08	0.96	0.95	0.93	**<0.01**	1.11	0.93	0.96	0.53	1.58	0.66
YP_006865925.1,Glyceraldehyde-3-phosphate dehydrogenase	1.87	**<0.01**	1.52	**0.04**	0.93	0.84	0.98	0.27	1.07	0.59	1.47	0.16	0.97	0.93	1.32	0.08
YP_006867462.1, Phosphoglycerate kinase	3.75	**0.01**	1.36	0.31	2.71	**<0.01**	0.96	0.72	0.78	0.32	1.25	0.92	0.52	0.16	2.29	0.24
YP_006866377.1, Phosphoglycerate mutase	1.57	0.17	0.77	0.63	1.43	**0.03**	0.26	0.37	0.76	1	2.00	0.49	1.07	0.93	1.46	0.31
YP_006866708.1, Enolase Eno	0.78	**0.01**	0.85	0.03	0.74	0.45	0.84	0.59	0.71	0.07	0.72	0.12	0.78	0.39	1.01	0.61
YP_006867615.1, Pyruvate kinase PykF	0.75	0.22	1.10	0.62	0.65	0.19	0.75	0.25	0.69	0.51	1.61	0.29	0.57	0.02	0.82	0.39
YP_006867982.1, Fuctose bisphosphatase	0.83	0.64	1.18	0.67	0.61	0.76	0.29	0.06	0.28	**<0.01**	0.95	0.47	0.51	0.05	0.90	0.71
YP_006868022.1, Pyruvate dehydrogenase α	2.74	**<0.01**	2.21	**<0.01**	2.16	**0.01**	1.91	0.10	2.90	**0.05**	2.76	**<0.01**	1.49	0.25	1.14	0.46
YP_006868492.1, Pyruvate dehydrogenase β	1.26	0.50	0.67	**0.01**	2.38	**<0.01**	1.57	0.46	2.15	0.29	2.43	**<0.01**	2.57	<0.01	2.29	**0.01**
YP_006866261.1, Pyruvate dehydrogenase AcoAB	1.59	0.18	1.41	0.30	1.91	0.07	1.35	1	1.63	0.21	2.20	0.02	1.63	0.35	2.08	0.26
YP_006868023.1, pyruvate dehydrogenase AceF	1.16	0.64	0.47	1	1.35	0.11	1.13	0.78	2.28	0.34	3.41	**0.01**	3.67	<0.01	2.98	**0.05**
YP_006868660.1, Dihydrolipoamide dehydrogenase	1.12	0.11	1.35	**<0.01**	0.70	**0.01**	0.84	0.16	1.36	**0.04**	1.19	**0.01**	0.96	0.81	0.91	0.37
TCA Cycle																
YP_006868127.1, Malate dehydrogenase	0.88	0.35	1.49	**0.04**	0.77	**0.02**	0.81	0.13	1.02	0.96	0.98	0.15	0.87	1.69	0.99	0.98
YP_006866707.1, Citrate synthase	0.82	**0.01**	0.96	0.61	0.92	0.78	1.65	0.45	0.90	0.53	1.04	0.47	0.89	0.68	1.24	0.53
YP_006868474.1, Aconitate hydratase	1.06	0.86	0.71	0.15	0.97	0.56	1.48	0.14	1.20	0.44	1.89	0.44	1.79	**0.01**	2.29	0.13
YP_006868735.1, Isocitrate dehydrogenase	1.11	0.55	0.93	0.64	0.67	0.15	0.86	0.53	1.31	0.24	2.16	**<0.01**	1.04	0.83	0.68	0.09
YP_006867957.1, oxoglutarate dehydrogenase SucA	1.43	0.07	1.36	0.07	0.25	**<0.01**	0.56	0.58	0.73	0.49	1.56	0.35	0.76	0.53	0.84	0.37
YP_006867958.1, oxoglutarate dehydrogenase SucB	5.43	**<0.01**	2.53	**<0.01**	0.87	0.53	1.51	0.16	0.85	0.80	0.95	0.80	0.41	**0.03**	0.76	0.37

Proteins (Reference No./ Name)	17.5g·L^{-1}				35g·L^{-1}				52.5g·L^{-1}				70g·L^{-1}			
	High Light		Low Light		High Light		Low Light		High Light		Low Light		High Light		Low Light	
	FC	p value	FC	p value	FC	p value	FC	p value	FC	p value	FC	p value	FC	p value	FC	p value
YP_006868521.1, Succinyl-CoA synthetase	1.18	0.93	0.45	**<0.01**	0.74	0.38	1.03	0.86	0.96	0.92	1.49	0.09	1.07	0.67	1.57	0.17
YP_006867595.1, Succinate dehydrogenase SdhA	0.82	0.37	1.03	0.76	2.33	**0.01**	1.48	0.28	1.37	0.23	1.00	0.42	1.18	0.54	1.42	0.21
YP_006867596.1, Succinate dehydrogenase SdhB	1.50	0.10	1.47	0.10	4.28	**0.02**	1.05	0.71	1.86	0.34	0.96	0.97	0.96	0.91	1.16	0.64
YP_006867594.1, Succinate dehydrogenase SdhC	0.95	0.83	0.74	0.38	0.93	0.88	0.61	0.83	0.57	0.17	0.97	0.75	1.41	0.11	1.70	**0.03**
YP_006866281.1, Fumarate hydratase	1.94	**0.03**	1.19	0.66	1.48	**0.05**	1.23	0.78	1.23	0.53	2.03	0.15	1.94	0.16	1.38	0.46
Anaplerotic metabolism																
YP_006867408.1, Pyruvate carboxylase	0.97	0.86	0.98	0.97	1.09	0.45	0.77	0.27	0.72	0.35	1.35	0.11	1.20	0.20	1.07	0.57
YP_006866763.1,Phosphoenolpyruvate carboxykinase	1.69	**0.01**	1.31	0.23	1.26	0.84	1.31	0.92	1.00	0.62	1.83	0.49	2.43	**0.02**	1.88	0.46
YP_006868831.1, Adenylosuccinate lyase	3.71	**<0.01**	2.01	**0.03**	2.03	**0.01**	2.29	**0.04**	3.15	**0.03**	2.76	**<0.01**	1.02	0.92	0.51	0.11
YP_006866302.1, Adenylosuccinate synthetase	1.92	0.08	1.30	0.51	1.39	0.12	2.09	0.46	0.77	0.48	0.91	0.41	1.23	0.55	2.61	**0.02**
Membrane Bioenergetics																
YP_006865934.1, NADH dehydrogenase	1.39	0.35	1.63	0.48	1.78	0.93	1.28	0.49	2.52	**0.02**	1.86	0.17	1.81	0.11	1.27	0.90
YP_006868191.1, NQR Subunit A	1.25	0.30	1.27	0.31	1.25	0.31	2.52	0.47	1.05	0.77	1.61	0.11	1.61	0.19	0.92	0.40
YP_006868192.1, NQR Subunit B	0.90	1	0.89	1	1.40	1	1.14	0.19	1.33	0.05	1.70	0.98	1.70	0.58	0.82	0.50
YP_006868193.1, NQR Subunit C	1.81	**0.05**	0.89	1	4.17	1	2.58	0.17	1.75	**0.01**	1.30	0.12	1.30	0.56	0.69	0.33
YP_006868194.1, NQR Subunit D	1.21	0.25	0.89	1	1.75	1	1.71	0.15	1.27	0.09	1.27	1	1.27	0.61	0.85	0.26
YP_006868195.1, NQR Subunit E	0.90	1	1.18	0.25	1.71	0.25	1.13	1	0.99	0.17	1.71	0.24	1.71	0.17	1.15	1
YP_006868196.1, NQR Subunit F	1.82	0.31	1.98	0.09	2.98	0.09	2.13	0.32	1.53	**0.04**	1.27	**<0.01**	1.27	0.29	1.13	0.57
YP_006867701.1, ACIII protein ActA	0.57	0.23	0.62	0.27	2.15	0.27	0.67	0.40	0.71	**0.03**	0.92	0.66	0.98	0087	0.92	0.68
YP_006867702.1, ACIII protein ActB	0.75	0.11	0.81	0.09	1.56	0.09	1.47	0.90	1.12	0.55	0.65	0.24	0.90	0.31	0.59	0.11
YP_006867703.1, ACIII protein ActC	1.53	0.11	1.28	0.41	2.93	0.40	3.03	0.63	1.99	0.14	1.15	0.65	0.89	0.97	0.87	0.89
YP_006867704.1, ACIII protein ActD	1.16	0.32	1.27	0.08	1.11	0.08	0.96	0.21	1.56	0.73	1.07	0.59	0.88	0.50	0.78	0.11
YP_006867705.1, ACIII protein ActE	0.77	0.53	1.15	0.59	1.49	0.59	0.96	0.74	0.53	0.87	0.62	**0.04**	0.76	0.32	1.00	0.87
YP_006867706.1, ACIII protein ActF	2.20	**0.02**	1.16	0.67	2.73	0.67	1.96	0.68	3.16	0.21	3.77	0.53	1.42	0.62	2.61	0.40

Proteins (Reference No./ Name)	17.5g·L^{-1}				35g·L^{-1}				52.5g·L^{-1}				70g·L^{-1}			
	High Light		Low Light		High Light		Low Light		High Light		Low Light		High Light		Low Light	
	FC	p value	FC	p value	FC	p value	FC	p value	FC	p value	FC	p value	FC	p value	FC	p value
YP_006867708.1, Cytochrome c oxidase subunit I	1.15	0.83	1.14	0.58	1.08	0.58	0.58	0.59	1.63	0.62	1.05	0.69	1.13	0.79	0.71	0.33
YP_006867175.1, Cytochrome c oxidase subunit II	0.74	**0.02**	0.72	**0.01**	1.13	**0.01**	0.44	0.08	2.07	0.53	1.15	0.51	0.89	0.74	0.73	0.29
YP_006867725.1, Cytochrome c oxidase subunit III	0.96	0.75	1.46	0.75	2.41	0.75	2.34	0.78	1.42	0.37	1.27	0.99	1.29	0.32	1.27	0.39
YP_006867726.1, Cytochrome c oxidase subunit IV	0.69	0.22	0.68	0.22	1.04	0.22	1.13	1	1.22	1	1.20	0.58	1.38	0.23	1.15	1
YP_006868230.1, Cytochrome cbb3 CcoN & CcoO	0.90	1	0.89	1	2.16	1	0.64	0.20	2.94	0.26	2.21	**0.02**	1.92	0.23	0.91	0.70
YP_006868232.1, Cytochrome cbb3 CcoP	1.30	0.25	1.77	**0.05**	1.04	**0.05**	1.13	1	0.99	1	0.98	1	1.04	1	1.15	1
YP_006867547.1, ATP synthase alpha chain AtpA	1.07	0.63	1.00	0.27	0.95	0.92	0.91	0.73	0.99	0.29	0.86	0.08	1.04	0.65	1.26	0.45
YP_006867175.1, ATP synthase beta chain AtpD	0.68	**<0.01**	0.71	**<0.01**	0.82	0.48	0.77	0.36	1.12	0.89	0.88	0.29	1.03	0.77	1.12	0.72
YP_006867548.1, ATP synthase gamma chain AtpG	0.91	0.62	0.76	0.21	0.75	0.32	0.52	0.09	0.85	0.51	1.46	0.51	0.86	0.66	0.71	0.26
YP_006867546.1, ATP synthase delta chain AtpH	0.97	0.98	0.96	0.96	0.44	**0.04**	0.39	**0.04**	0.62	0.39	1.12	0.56	0.74	0.31	1.08	0.91
YP_006867174.1, ATP synthase epsilon chain AtpC	2.11	**0.04**	0.59	0.26	1.46	0.84	0.91	0.63	1.91	**0.04**	1.76	0.20	0.82	0.57	0.26	**0.04**
YP_006867543.1, ATP synthase A chain AtpB	3.92	**<0.01**	2.41	**0.02**	1.71	0.94	0.79	0.76	1.09	0.48	2.47	0.83	0.75	0.40	1.50	0.52
YP_006867545.1, ATP synthase B chain AtpF	1.81	**0.05**	0.89	1	1.15	0.39	0.76	0.79	0.82	0.27	0.46	0.60	1.58	0.97	0.95	0.32
YP_006867544.1, ATP synthase C chain Atp E	0.36	**<0.01**	0.45	**0.01**	0.31	0.34	0.71	0.87	0.97	0.76	0.94	0.76	1.49	0.09	1.07	0.68
YP_006865968.1, Proteorhodopsin	0.98	0.63	0.95	0.27	1.38	0.12	1.52	0.18	3.35	0.15	1.42	**0.02**	1.31	0.56	0.45	0.31
Transporters	0.60	**<0.01**	1.03	0.89	0.76	0.11	0.73	0.16	0.71	0.22	1.21	0.52	2.05	**<0.01**	1.64	**0.03**
YP_006868873.1, TBDT carbohydrate receptor SusC/RagA family	0.77	**0.01**	1.00	0.96	0.81	**0.04**	1.08	0.49	0.97	0.84	0.83	0.52	1.01	0.78	1.05	0.62
YP_006866744.1, TBDT receptor SusC/RagA family	0.71	**<0.01**	0.85	**<0.01**	1.28	**0.03**	0.87	0.16	1.00	0.95	0.81	0.11	0.86	0.12	1.05	0.52
YP_006866745.1, Outer membrane nutrient binding protein SusD family	1.13	0.68	1.91	**<0.01**	0.98	0.82	1.26	0.18	1.11	0.77	0.89	0.70	0.85	0.21	0.88	0.22
YP_006868324.1, Outer membrane nutrient binding protein SusD family	1.06	0.81	1.65	**<0.01**	0.62	**0.03**	1.03	0.87	0.73	0.28	0.91	0.69	0.82	0.43	0.96	0.92
YP_006868325.1, TBDT Outer membrane receptor RagA/SusC family																

Proteins (Reference No./ Name)	17.5g·L^{-1}				35g·L^{-1}				52.5g·L^{-1}				70g·L^{-1}			
	High Light		Low Light		High Light		Low Light		High Light		Low Light		High Light		Low Light	
	FC	p value	FC	p value	FC	p value	FC	p value	FC	p value	FC	p value	FC	p value	FC	p value
YP_006866963.1, TBDT Outer membrane receptor SusC family	0.53	**<0.01**	1.03	0.58	0.67	**0.04**	1.14	0.27	0.53	**<0.01**	0.81	0.51	0.56	**0.03**	0.66	0.12
YP_006866076.1, TBDT Outer membrane receptor SusC/RagA family	0.36	**<0.01**	0.59	**<0.01**	1.08	0.99	1.14	0.48	1.31	0.22	1.55	**0.05**	2.04	**0.03**	1.07	0.94
YP_006866995.1, Outer membrane nutrient binding protein SusD family	0.54	**<0.01**	0.58	**<0.01**	1.32	0.16	1.29	0.31	0.73	0.90	0.89	0.90	0.65	**0.02**	1.27	0.27
YP_006868771.1, TBDT outer membrane receptor	0.42	**0.02**	0.64	0.24	0.51	**<0.01**	1.16	0.50	0.49	**<0.01**	0.56	**<0.01**	0.61	**0.04**	1.03	0.57
YP_006868872.1, Starch-binding outer membrane protein SusD family	0.65	0.14	0.60	0.09	1.84	0.10	1.11	0.75	0.47	**0.04**	0.95	0.61	1.93	0.12	2.63	0.12
YP_006866075.1, Carbohydrate binding outer membrane protein SusD family	1.51	0.48	1.22	0.62	3.61	**<0.01**	1.97	0.12	0.77	0.57	1.18	0.89	2.14	0.25	1.03	0.96
YP_006868327.1, TBDT Outer membrane receptor SusC/RagA family	0.56	0.08	0.73	0.50	1.77	0.13	1.27	0.58	5.09	**0.01**	1.23	0.59	35.27	**<0.01**	5.98	**<0.01**
YP_006867195.1, TBDT Outer membrane receptor SusC/RagA family	0.35	**<0.01**	0.71	0.30	0.36	0.34	0.62	0.96	0.53	0.58	0.81	0.97	1.04	0.79	0.71	0.44
YP_006868326.1, Outer membrane nutrient binding protein, SusD family	0.90	1	0.89	1	2.46	**<0.01**	1.42	0.22	5.16	**<0.01**	2.26	0.10	10.23	<**0.01**	3.85	**<0.01**
YP_006867202.1, TBDT Outer membrane receptor SusC/RagA family	0.47	**0.05**	0.46	**0.05**	0.95	0.95	0.49	**0.02**	1.57	0.24	1.34	0.50	4.09	**<0.01**	1.61	0.21
YP_006868657.1, TBDT outer membrane receptor	0.45	**0.03**	0.44	**0.03**	0.47	0.06	0.87	0.71	0.61	0.66	0.71	0.79	2.89	**0.03**	4.18	0.12
YP_006869089.1, ABC-type multidrug or Lipoprotein, ATP-binding component	0.36	**<0.01**	0.47	**<0.01**	0.69	0.21	0.90	0.74	0.72	0.44	0.71	0.54	1.18	0.53	2.61	0.26
YP_006869074.1, ABC-type multidrug or Lipoprotein, ATP-binding component	0.09	**<0.01**	0.19	**<0.01**	0.55	**0.01**	0.88	0.66	1.60	0.25	1.96	0.23	1.20	0.44	2.94	0.60
YP_006867001.1, ABC-type exporter, ATP-binding permease protein MsbA	1.77	0.11	1.55	**0.04**	2.69	**<0.01**	1.32	0.10	2.61	**<0.01**	1.46	0.51	1.16	0.50	0.63	0.50

Proteins (Reference No./ Name)	17.5g·L^{-1}				35g·L^{-1}				52.5g·L^{-1}				70g·L^{-1}			
	High Light		Low Light		High Light		Low Light		High Light		Low Light		High Light		Low Light	
	FC	p value	FC	p value	FC	p value	FC	p value	FC	p value	FC	p value	FC	p value	FC	p value
YP_006869086.1, ABC-type permease ADOP	0.24	**<0.01**	0.46	**<0.01**	0.35	0.91	0.37	0.71	0.95	0.95	0.88	0.96	0.79	0.93	1.00	0.76
YP_006869079.1, ABC-type permease ADOP	0.18	**<0.01**	0.46	**<0.01**	0.33	0.08	0.50	0.38	0.67	0.34	1.04	0.59	0.78	0.99	0.90	0.99
YP_006869077.1, ABC-type permease ADOP	0.12	**<0.01**	0.40	**<0.01**	0.34	0.22	0.39	0.63	0.81	0.99	0.66	0.75	1.55	0.22	2.07	0.25
YP_006867538.1, ABC-type ATPase component	0.34	**<0.01**	0.75	0.35	0.52	**0.04**	0.88	0.75	0.78	0.49	1.56	0.25	0.68	0.54	0.58	0.34
YP_006867668.1, Bicarbonate ABC-type transporter	0.77	0.66	1.03	0.75	2.24	**0.01**	1.45	0.18	1.36	0.38	0.89	0.49	1.96	0.16	0.92	0.46
YP_006867581.1, Efflux ABC transporter	1.57	0.26	2.18	**<0.01**	1.95	1.33	2.54	**0.03**	1.53	0.24	2.65	**0.01**	0.88	0.70	1.42	0.36
YP_006867066.1, Efflux ABC transporter	1.25	0.55	2.18	**<0.01**	1.37	0.53	2.06	0.13	1.53	0.24	2.13	0.07	1.18	0.72	1.97	0.09
YP_006869080.1, ABC-type permease ADOP	0.12	**<0.01**	0.16	**<0.01**	0.26	**0.04**	0.53	0.45	1.15	0.63	1.79	0.73	2.28	0.09	1.71	0.28
YP_006867383.1,Macrolide efflux ABC transporter	0.52	0.11	0.78	0.47	0.40	**0.02**	0.88	0.83	1.18	0.64	2.08	0.07	1.51	0.28	1.94	0.27
YP_006869088.1, ABC-type permease ADOP	0.10	**<0.01**	0.27	**<0.01**	1.41	0.64	1.53	0.37	1.84	**0.05**	4.06	**0.04**	1.76	0.37	0.85	0.26
YP_006867384.1,Macrolide efflux ABC transporter	0.61	**1**	0.69	0.33	3.15	**<0.01**	1.42	0.22	1.84	**0.05**	1.81	0.07	1.04	1	1.15	1
YP_006868801.1, RND-type efflux pump	5.06	**<0.01**	3.25	**<0.01**	3.29	**<0.01**	1.56	**0.04**	6.65	**<0.01**	1.94	**0.05**	1.66	1.15	1.38	0.61
YP_006868802.1, Efflux transporter, NodT family	4.47	**<0.01**	2.99	**<0.01**	4.07	**<0.01**	2.14	0.13	5.46	**0.01**	5.25	**<0.01**	1.99	0.06	0.55	**0.01**
YP_006867387.1, Outer membrane efflux TolC	0.41	**0.04**	1.06	0.57	0.60	**0.01**	0.81	0.28	0.68	0.14	0.92	0.09	0.69	0.23	0.82	0.35
YP_006867932.1, Outer membrane efflux TolC	1.35	0.19	1.17	0.49	0.94	0.65	1.61	0.11	1.12	0.77	1.15	0.47	1.42	0.22	0.52	**0.02**
YP_006867930.1, Multidrug/ cation efflux protein	1.21	0.47	1.18	0.31	0.96	0.61	0.68	0.67	2.79	**0.04**	2.55	**0.04**	3.10	**<0.01**	2.02	0.07
YP_006869138.1, Multidrug/ cation efflux protein	0.64	0.25	0.64	0.24	3.96	**<0.01**	2.46	**0.03**	3.58	**0.01**	1.94	**0.03**	1.34	0.37	0.55	0.25
YP_006868800.1, RND-type efflux pump	5.72	**0.04**	2.27	0.23	3.62	**<0.01**	1.88	0.39	3.68	**<0.01**	3.50	**<0.01**	4.52	**<0.01**	4.36	**<0.01**
YP_006867386.1, Efflux protein, RND family	0.68	0.32	0.60	1	1.64	0.37	0.93	0.71	1.00	0.99	2.25	**0.05**	1.51	0.28	2.14	0.13
YP_006866290.1, Toxic metal ion efflux pump	0.60	0.35	0.86	0.79	2.22	0.12	1.77	0.16	1.89	0.08	1.14	0.57	2.38	**0.05**	1.52	0.21
Light Sensing proteins																
YP_006867167.1, Bacteriophytochrom like protein histidinekinase/PAS domains	1.12	0.72	1.55	0.14	0.99	0.42	1.28	0.79	1.15	0.80	1.78	0.27	2.71	**0.01**	2.06	0.17

Proteins (Reference No./ Name)	17.5g·L^{-1}				35g·L^{-1}				52.5g·L^{-1}				70g·L^{-1}			
	High Light		Low Light		High Light		Low Light		High Light		Low Light		High Light		Low Light	
	FC	p value	FC	p value	FC	p value	FC	p value	FC	p value	FC	p value	FC	p value	FC	p value
YP_006869231.1, Bacteriophytochrom like protein Histidine kinase/GAF domains	1.16	0.59	1.33	0.31	0.48	0.14	0.76	0.81	1.41	0.71	2.06	1	1.39	0.36	1.25	0.92
YP_006869331.1, Bacteriophytochrom like protein PAS and GAF domains	0.90	1	1.20	0.25	0.31	**0.01**	0.96	0.78	3.82	**<0.01**	2.87	**0.01**	0.99	0.91	1.91	0.64
YP_006867168.1, Bacteriophytochrom like protein Histidine kinase/multiple PAS domains	1.31	0.45	0.82	0.66	0.54	0.06	1.19	0.67	1.44	0.31	0.75	0.21	1.04	1	1.15	1
YP_006867637.1, Two component sensor with histidine kinase and TPR repeat domains	0.64	0.25	0.89	0.82	1.30	0.24	2.28	**0**	2.78	0.11	3.10	**0.01**	1.34	0.38	0.61	0.20
YP_006866773.1, Two component regulator	1.27	0.44	1.08	0.75	0.85	0.97	1.45	0.35	1.27	0.23	1.52	0.11	1.02	0.96	0.69	0.11
YP_006866774.1, Two component sensor Histidine kinase	2.10	1	2.10	1	1.04	1	1.71	0.08	2.51	**0.01**	2.66	**0.01**	1.72	0.17	1.32	0.56
Oxidative stress management																
YP_006866424.1, Peroxiredoxin	0.57	0.15	0.45	**0.02**	0.88	0.95	1.48	0.24	2.33	**0.02**	1.48	0.34	4.68	**<0.01**	3.53	**<0.01**
YP_006866442.1, Peroxiredoxin	1.60	0.21	1.79	0.15	0.39	0.09	0.66	0.28	0.45	**0.02**	1.05	0.98	1.79	0.27	2.47	0.06
YP_006868959.1, Peroxiredoxin, 2-cys family AhpC	5.87	**<0.01**	2.59	**<0.01**	1.05	0.60	1.02	0.89	0.71	0.14	0.68	0.44	0.99	0.45	1.12	0.97
Adhesion Proteins																
YP_006868804.1, Secreted protein with fasciclin repeat domain	1.63	**<0.01**	1.29	**0.01**	1.21	0.11	0.73	0.23	1.31	0.46	0.65	0.27	1.85	0.09	1.37	0.22
YP_006865932.1, Cell surface protein, putative	1.67	**<0.01**	2.36	**<0.01**	0.56	0.19	0.73	0.34	0.99	0.63	1.32	0.09	0.57	0.31	0.69	0.70
YP_006866709.1, Secreted protein with fasciclin domain	20.7	**<0.01**	3.41	**<0.01**	5.05	0.08	1.23	0.55	2.31	**0.02**	1.58	0.79	3.96	0.06	1.54	0.35
YP_006867848.1, Secreted cell surface protein Por secretion system C-terminal sorting domain	15.4	**<0.01**	6.11	**<0.01**	6.64	**<0.01**	2.05	0.10	2.09	0.66	1.61	0.27	1.58	0.32	1.24	0.85
YP_006865954.1, PEP-CTERM system TPR-repeat lipoprotein	1.83	**0.05**	1.45	0.26	1.70	**0.03**	2.87	**0.04**	1.57	0.24	3.22	0.14	6.70	**<0.01**	2.40	0.06
YP_006867733.1, Secreted protein with fasciclin domain	5.76	**<0.01**	2.57	**<0.01**	0.74	0.37	0.74	0.46	0.99	1	0.96	1	1.04	1	1.61	0.21

Proteins (Reference No./ Name)	17.5g·L^{-1}				35g·L^{-1}				52.5g·L^{-1}				70g·L^{-1}			
	High Light		Low Light		High Light		Low Light		High Light		Low Light		High Light		Low Light	
	FC	p value	FC	p value	FC	p value	FC	p value	FC	p value	FC	p value	FC	p value	FC	p value
YP_006868541.1, Secreted adhesin-like protein, cleaved adhesin superfamily, Por secretion system	0.21	**<0.01**	0.35	**<0.01**	0.51	**<0.01**	0.53	**<0.01**	0.54	**<0.01**	0.70	**<0.01**	0.82	0.29	1.45	**0.02**
Gliding motility associated proteins																
YP_006866693.1, Gliding motility protien GldM	2.93	**<0.01**	2.11	**<0.01**	1.07	0.65	1.26	0.27	0.94	0.85	0.74	0.38	1.00	0.75	1.69	**0.02**
YP_006866692.1, Gliding motility protien GldN	0.66	**<0.01**	0.86	0.15	0.85	0.82	0.90	0.44	1.36	0.50	2.16	**<0.01**	1.49	**0.04**	0.86	0.75
YP_006866694.1, Gliding motility protien GldL	1.12	0.72	1.66	0.18	1.68	0.16	0.87	0.83	0.44	**0.03**	0.47	**0.05**	2.32	**0.01**	2.33	**0.04**
YP_006868018.1, Gliding motility protien GldJ	0.33	**<0.01**	0.38	**<0.01**	2.52	0.26	1.06	0.93	0.75	0.44	1.19	0.71	4.73	**<0.01**	3.81	**0.01**

Table S3. *P. torquis* proteome response to low and high salinities compare with optimal salinity (35g·L^{-1}).

Proteins (Reference No./ Name)	17.5g·L^{-1}		52.5g·L^{-1}		70g·L^{-1}	
	FC	p value	FC	p value	FC	p value
Central Metabolism						
Glycolysis						
YP_006868884.1, Phosphoglucose isomerase	2.73	**<0.01**	2.71	**<0.01**	3.21	**<0.01**
YP_006868881.1, Phosphohexokinase	0.67	**0.01**	1.02	0.56	0.71	**0.03**
YP_006865907.1, Aldolase	2.91	**<0.01**	1.85	0.44	2.76	**<0.01**
YP_006868950.1, Triosephosphate isomerase	0.13	**<0.01**	0.58	**<0.01**	0.90	0.40
YP_006865925.1,Glyceraldehyde-3-phosphate dehydrogenase	1.13	0.58	1.04	0.68	0.93	0.47
YP_006867462.1, Phosphoglycerate kinase	1.19	0.84	0.85	0.47	1.45	0.88
YP_006866377.1, Phosphoglycerate mutase	0.61	0.61	0.71	0.31	0.69	0.79
YP_006866708.1, Enolase Eno	1.53	**<0.01**	1.06	0.48	1.27	**0.05**
YP_006867615.1, Pyruvate kinase PykF	1.89	**<0.01**	1.36	0.47	1.42	**0.02**
YP_006867982.1, Fuctose bisphosphatase	1.02	0.39	1.56	0.10	1.54	0.06
YP_006868022.1, Pyruvate dehydrogenase α	1.26	0.16	0.86	0.38	0.99	0.69
YP_006868492.1, Pyruvate dehydrogenase β	1.54	**0.02**	1.35	0.33	1.28	0.34
YP_006866261.1, Pyruvate dehydrogenase AcoAB	0.96	0.62	0.76	0.14	0.86	0.39
YP_006868023.1, pyruvate dehydrogenase AceF	1.46	**0.02**	1.33	0.23	1.36	0.09
YP_006868660.1, Dihydrolipoamide dehydrogenase	0.99	0.47	0.89	0.45	0.69	**0.01**
TCA Cycle						
YP_006868127.1, Malate dehydrogenase	0.73	**<0.01**	1.39	**<0.01**	1.36	**<0.01**
YP_006866707.1, Citrate synthase	0.53	**0.04**	0.76	0.65	0.73	0.43
YP_006868474.1, Aconitate hydratase	1.88	**<0.01**	2.03	**<0.01**	2.35	**<0.01**
YP_006868735.1, Isocitrate dehydrogenase	1.61	**<0.01**	1.17	0.48	0.97	0.65
YP_006867957.1, Oxoglutarate dehydrogenase SucA	3.37	**<0.01**	1.19	0.83	1.30	0.07
YP_006867958.1, Oxoglutarate dehydrogenase SucB	2.65	**<0.01**	0.55	**<0.01**	0.72	**0.04**

Proteins (Reference No./ Name)	17.5g·L^{-1}		52.5g·L^{-1}		70g·L^{-1}	
	FC	p value	FC	p value	FC	p value
YP_006868521.1, Succinyl-CoA synthetase	2.26	**<0.01**	1.39	**0.03**	1.78	**<0.01**
YP_006867595.1, Succinate dehydrogenase SdhA	0.28	**<0.01**	0.76	0.65	0.73	0.43
YP_006867596.1, Succinate dehydrogenase SdhB	0.07	**<0.01**	0.81	0.96	0.60	0.53
YP_006867594.1, Succinate dehydrogenase SdhC	0.47	**<0.01**	1.86	**0.01**	1.43	**0.01**
YP_006866281.1, Fumarate hydratase	2.34	**<0.01**	1.43	0.38	1.46	0.29
Anaplerotic metabolism						
YP_006867408.1, Pyruvate carboxylase	1.42	**<0.01**	1.03	0.84	0.92	0.34
YP_006866763.1,Phosphoenolpyruvate carboxykinase	2.86	**<0.01**	1.22	0.54	1.46	0.12
YP_006868831.1, Adenylosuccinate lyase	4.27	**<0.01**	1.42	0.35	1.94	**0.01**
YP_006866302.1, Adenylosuccinate synthetase	0.66	0.13	0.92	0.62	0.99	0.94
Membrane Bioenergetics						
YP_006865934.1, NADH dehydrogenase	0.83	**<0.01**	0.58	**<0.01**	0.65	**0.02**
YP_006868191.1, NQR Subunit A	0.60	0.08	2.00	**<0.01**	3.20	**<0.01**
YP_006868192.1, NQR Subunit B	0.28	**<0.01**	0.72	0.71	0.80	0.33
YP_006868193.1, NQR Subunit C	0.47	**<0.01**	1.00	0.61	0.96	0.69
YP_006868194.1, NQR Subunit D	0.68	0.06	0.79	0.08	0.90	0.45
YP_006868195.1, NQR Subunit E	0.79	0.49	0.92	0.55	0.96	0.98
YP_006868196.1, NQR Subunit F	0.53	0.13	1.31	0.27	1.56	**<0.01**
YP_006867701.1, ACIII protein ActA	1.21	0.43	1.31	0.77	1.25	0.32
YP_006867702.1, ACIII protein ActB	0.36	**<0.01**	0.56	0.09	0.43	**0.03**
YP_006867703.1, ACIII protein ActC	2.41	**<0.01**	1.62	**0.02**	2.02	**<0.01**
YP_006867704.1, ACIII protein ActD	0.98	0.90	0.53	**<0.01**	0.62	**0.02**
YP_006867705.1, ACIII protein ActE	0.32	**0.01**	0.90	0.48	0.66	0.91
YP_006867706.1, ACIII protein ActF	0.22	**<0.01**	1.15	0.51	0.59	**0.04**

Proteins (Reference No./ Name)	17.5g·L^{-1}		52.5g·L^{-1}		70g·L^{-1}	
	FC	p value	FC	p value	FC	p value
YP_006867708.1, Cytochrome c oxidase subunit I	0.27	**0.01**	1.75	0.25	0.75	0.80
YP_006867175.1, Cytochrome c oxidase subunit II	0.70	0.10	1.11	0.29	0.79	0.34
YP_006867725.1, Cytochrome c oxidase subunit III	0.46	**<0.01**	1.04	0.78	0.74	**0.04**
YP_006867726.1, Cytochrome c oxidase subunit IV	0.96	0.25	1.51	**<0.01**	1.06	0.24
YP_006868230.1, Cytochrome cbb3 CcoN & CcoO	0.41	**<0.01**	1.01	0.62	0.91	0.91
YP_006868232.1, Cytochrome cbb3 CcoP	1.23	**0.02**	1.02	1	0.96	1
YP_006867547.1, ATP synthase alpha chain	0.83	**<0.01**	1.19	0.13	1.20	**0.03**
YP_006867175.1, ATP synthase beta chain	0.70	0.10	1.11	0.29	0.79	0.34
YP_006867548.1, ATP synthase gamma chain	1.00	0.85	0.90	0.20	1.36	0.07
YP_006867546.1, ATP synthase delta chain	0.92	0.79	1.27	0.32	1.43	**0.03**
YP_006867174.1, ATP synthase epsilon chain	1.34	0.21	0.94	0.56	1.57	0.13
YP_006867543.1, ATP synthase A chain	0.17	**0.01**	1.76	0.65	0.40	0.06
YP_006867545.1, ATP synthase B chain	0.46	0.06	2.01	0.12	1.50	0.18
YP_006867544.1, ATP synthase C chain	0.45	**<0.01**	1.27	0.08	0.76	0.19
YP_006865968.1, Proteorhodopsin	0.20	**<0.01**	0.74	0.13	0.43	**0.03**
Transporters	1.38	**<0.01**	1.15	0.32	1.03	0.53
YP_006868873.1, TBDT carbohydrate receptor SusC/RagA family	1.24	**<0.01**	0.93	0.97	0.77	**0.01**
YP_006866744.1, TBDT receptor SusC/RagA family	0.90	0.17	0.86	0.14	0.74	**<0.01**
YP_006866745.1, Outer membrane nutrient binding protein SusD family	0.43	**<0.01**	1.35	**<0.01**	1.86	**<0.01**
YP_006868324.1, Outer membrane nutrient binding protein SusD family	0.79	0.81	1.59	**<0.01**	2.35	**<0.01**
YP_006868325.1, TBDT Outer membrane receptor RagA/SusC family						

Proteins (Reference No./ Name)	17.5g·L^{-1}		52.5g·L^{-1}		70g·L^{-1}	
	FC	p value	FC	p value	FC	p value
YP_006866963.1, TBDT Outer membrane receptor SusC family	0.91	0.86	1.11	0.09	1.45	**<0.01**
YP_006866076.1, TBDT Outer membrane receptor SusC/RagA family	0.67	**<0.01**	0.90	0.48	0.79	0.11
YP_006866995.1, Outer membrane nutrient binding protein SusD family	0.32	**<0.01**	0.68	**0.01**	0.83	
YP_006868771.1, TBDT outer membrane receptor	0.11	**<0.01**	3.67	**<0.01**	6.11	**<0.01**
YP_006868872.1, Starch-binding outer membrane protein SusD family	1.65	**0.01**	1.04	0.95	1.05	0.91
YP_006866075.1, Carbohydrate binding outer membrane protein SusD family	0.54	**0.02**	0.88	0.14	1.07	0.80
YP_006868327.1, TBDT Outer membrane receptor SusC/RagA family	0.56	**0.01**	4.34	**0.01**	10.8	**0.01**
YP_006867195.1, TBDT Outer membrane receptor SusC/RagA family	1.62	**0.01**	1.01	0.80	0.80	0.74
YP_006868326.1, Outer membrane nutrient binding protein, SusD family	0.67	0.13	1.51	0.07	1.61	**0.04**
YP_006867202.1, TBDT Outer membrane receptor SusC/RagA family	0.65	0.07	1.18	0.49	1.13	0.83
YP_006868657.1, TBDT outer membrane receptor	0.72	0.43	1.46	0.18	2.42	0.06
YP_006869089.1, ABC-type multidrug or Lipoprotein, ATP-binding component	0.89	0.95	1.65	0.31	1.17	0.63
YP_006869074.1 ABC-type multidrug or Lipoprotein, ATP-binding component	2.48	0.09	1.48	0.12	0.92	0.41
YP_006867001.1, ABC-type exporter, ATP-binding permease protein MsbA	1.14	0.06	1.28	0.15	0.75	0.47

Proteins (Reference No./ Name)	17.5g·L^{-1}		52.5g·L^{-1}		70g·L^{-1}	
	FC	p value	FC	p value	FC	p value
YP_006869086.1, ABC-type permease ADOP	1.00	0.09	1.19	0.95	0.54	0.51
YP_006869079.1, ABC-type permease ADOP	1.26	0.11	0.90	0.86	1.47	0.36
YP_006869077.1, ABC-type permease ADOP	1.24	0.14	1.27	0.19	1.89	**<0.01**
YP_006867538.1, ABC-type ATPase component	1.15	0.17	1.02	0.47	1.22	0.31
YP_006867668.1, Bicarbonate ABC-type transporter	0.60	**0.01**	0.83	0.18	1.47	0.36
YP_006867581.1, Efflux ABC transporter	1.24	0.14	1.27	0.19	1.89	**<0.01**
YP_006867066.1, Efflux ABC transporter	1.43	**0.03**	1.42	0.08	2.01	**<0.01**
YP_006869080.1, ABC-type permease ADOP	1.68	0.26	1.12	0.57	0.38	**0.04**
YP_006867383.1,Macrolide efflux ABC transporter	1.32	0.17	1.22	0.37	0.96	0.98
YP_006869088.1, ABC-type permease ADOP	2.97	**0.02**	1.14	0.86	0.70	0.20
YP_006867384.1,Macrolide efflux ABC transporter	1.31	0.28	0.86	0.47	0.51	**<0.01**
YP_006868801.1, RND-type efflux pump	0.42	0.07	0.34	0.01	0.25	**<0.01**
YP_006868802.1, Efflux transporter, NodT family	1.82	**0.05**	1.00	0.42	0.79	**<0.01**
YP_006867387.1, Outer membrane efflux TolC	0.86	0.30	0.75	**0.04**	0.66	**0.01**
YP_006867932.1, Outer membrane efflux TolC	1.35	**0.02**	1.14	0.61	1.31	0.49
YP_006867930.1, Multidrug/ cation efflux protein	1.93	**0.01**	1.37	0.38	1.45	0.96
YP_006869138.1, Multidrug/ cation efflux protein	0.16	**0.01**	0.64	0.17	0.41	**0.01**
YP_006868800.1, RND-type efflux pump	1.80	0.29	0.64	0.23	0.67	0.33
YP_006867386.1, Efflux protein, RND family	1.17	0.58	1.07	0.91	0.88	0.54
YP_006866290.1, Toxic metal ion efflux pump	0.42	**0.03**	0.56	0.12	0.32	**<0.01**
Light Sensing proteins						
YP_006867167.1, Bacteriophytochrom like protein histidinekinase/PAS domains	1.59	0.07	1.64	0.44	0.83	0.42

Proteins (Reference No./ Name)	17.5g·L^{-1}		52.5g·L^{-1}		70g·L^{-1}	
	FC	p value	FC	p value	FC	p value
YP_006869231.1, Bacteriophytochrom like protein histidine kinase/GAF domains	1.80	**<0.01**	1.10	0.71	0.98	0.64
YP_006869331.1, Bacteriophytochrom like protein PAS and GAF domains	0.41	**0.03**	1.60	0.11	2.01	0.08
YP_006867168.1, Bacteriophytochrom like protein histidine kinase/multiple PAS domains	1.49	0.07	0.85	0.52	0.58	**0.01**
YP_006867637.1, Two component sensor with histidine kinase and TPR repeat domains	1.96	**<0.01**	1.66	0.08	1.44	**0.04**
YP_006866773.1, Two component regulator	2.07	**0.03**	0.78	0.12	0.81	0.22
YP_006866774.1, Two component sensor Histidine kinase	1.37	0.07	1.80	**0.03**	2.03	**<0.01**
YP_006866424.1, Peroxiredoxin	1.32	0.31	1.44	0.16	1.10	0.97
YP_006866442.1, Peroxiredoxin	2.24	**<0.01**	1.53	0.13	1.20	0.67
YP_006868959.1, Peroxiredoxin, 2-cys family AhpC	0.46	**<0.01**	1.16	0.34	1.61	**0.01**
Adhesion Proteins						
YP_006868804.1, Secreted protein with fasciclin repeat domain	1.33	**0.05**	1.29	0.84	1.07	0.76
YP_006865932.1, Cell surface protein, putative	1.41	**0.03**	0.89	0.52	1.36	0.30
YP_006866709.1, Secreted protein with fasciclin domain	5.94	**0.01**	0.58	0.30	1.15	0.74
YP_006867848.1, Secreted cell surface protein Por secretion system C-terminal sorting domain	1.18	0.41	0.55	0.10	0.44	0.06
YP_006865954.1, PEP-CTERM system TPR-repeat lipoprotein	3.76	**<0.01**	1.77	0.15	1.72	0.50
YP_006867733.1, Secreted protein with fasciclin domain	10.8	**<0.01**	0.56	**<0.01**	0.61	**0.01**

Proteins (Reference No./ Name)	17.5g·L^{-1}		52.5g·L^{-1}		70g·L^{-1}	
	FC	p value	FC	p value	FC	p value
YP_006868541.1, Secreted adhesin-like protein, cleaved adhesin superfamily, Por secretion system	1.45	0.34	1.54	**<0.01**	1.28	0.06
Gliding motility associated proteins						
YP_006866693.1, Gliding motility protien GldM	0.54	**<0.01**	0.63	**0.01**	0.56	**<0.01**
YP_006866692.1, Gliding motility protien GldN	2.31	**<0.01**	1.36	0.37	1.60	**0.01**
YP_006866694.1, Gliding motility protien GldL	0.62	0.07	1.20	**0.55**	1.30	0.21
YP_006868018.1, Gliding motility protien GldJ	1.65	**0.03**	1.60	**0.02**	1.23	0.56
Exopolysaccharide (EPS)						
YP_006867255.1, Glucose-1-phosphate thymidylyltransferase RfbA	1.44	**0.02**	0.87	0.22	0.95	0.64
YP_006867263.1, UDP-glucose dehydrogenase	1.82	**<0.01**	1.03	0.64	1.54	**0.04**
YP_006866124.1, N-acetyltransferase, NAT_SF superfamily	2.81	**<0.01**	1.77	0.06	2.03	**<0.01**
YP_006867241.1, Glycosyltransferase	1.94	**<0.01**	1.48	0.52	2.30	**0.03**
YP_006867236.1, Mannosyltransferase	1.41	0.06	1.33	0.41	1.61	0.08
YP_006867237.1, Galactosyltransferase	1.43	0.33	1.94	0.54	2.40	**0.03**
YP_006867247.1, Glycosyltransferase	0.55	**<0.01**	1.76	0.88	2.54	0.07
YP_006866298.1, UDP-D-glucuronic acid 4-epimerase	2.26	**<0.01**	0.71	**0.01**	0.88	0.42
YP_006867243.1, dTDP-D-Fucp3N transacetylase FdtB	1.24	**0.04**	1.87	**<0.01**	2.51	**<0.01**
YP_006867230.1, N-acetyl-L-fucosamine transferase	1.81	**0.02**	2.25	**0.01**	1.25	0.17
YP_006867237.1, Galactosyltransferase	1.14	0.21	1.75	**0.04**	1.45	0.11